Emulating randomized controlled trials using real-world evidence:

The case of long-acting insulins among patients with type 2 diabetes

Wanning Wang, BHSc

Department of Epidemiology, Biostatistics, and Occupational Health

McGill University, Montreal

June 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

ABSTRACT

Cardiovascular outcome trials (CVOTs) are a type of randomized controlled trial (RCT) used to inform decision making regarding the cardiovascular safety of antidiabetic medications for type 2 diabetes mellitus (T2DM). These trials have strict inclusion criteria that restrict their populations to individuals with established or at high risk of cardiovascular (CV) disease, leading to differences between the trial population and the real-world population who use the medications. Few studies have compared effect estimates between CVOTs and the emulation of RCTs using real-world data (RWD). To our knowledge, there are no studies examining potential differences between RCTs and RWD in patient characteristics and treatment effects of CV outcomes for long-acting insulin analogues. My thesis had two main objectives. The first was to synthesize the existing evidence on the emulation of CVOTs using RWD for newer antidiabetic medications. The second was to emulate the DEVOTE trial using RWD to determine the risk of major adverse cardiovascular events (MACE) among patients with T2DM taking one of two types of long-acting insulin analogues.

In the first manuscript, I systematically reviewed observational studies that emulated previous CVOTs for dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide 1 (GLP-1) receptor agonists, and sodium glucose cotransporter-2 (SGLT-2) inhibitors in patients with T2DM. I searched EMBASE, MEDLINE and Cochrane CENTRAL databases from inception to July 2023, for observational studies that focused on trial emulation or cross-sectional studies that reported the proportion of real-world patients eligible for completed CVOTs. Two independent reviewers screened articles, extracted data, and assessed study concordance with RCT results. Nineteen studies were included in our systematic review, including four cohort studies that emulated previous RCTs and 15 cross-sectional studies that evaluated trial eligibility. Results between RCTs and RWD were concordant for all drug classes in finding non-inferiority. The percent eligible ranged from 7% to 59% for SGLT-2 inhibitor trials and 6% to 54% for GLP-1 receptor agonist trials. These results suggest that, while RCTs and RWD are concordant in their estimates, the trials lack representativeness.

In the second manuscript, I emulated the DEVOTE trial using data from the United Kingdom's Clinical Practice Research Datalink (CPRD). I estimated the risk of MACE, a composite of myocardial infarction (MI), ischemic stroke, or CV death, among patients with

T2DM who initiated insulin glargine or insulin degludec. Two subpopulations were created based on whether patients met the inclusion criteria of the DEVOTE trial. I used Cox proportional hazards models with inverse probability of treatment weighting to estimate the hazard ratios (HRs) and corresponding bootstrapped 95% confidence intervals (CIs) for MACE, comparing new users of insulin degludec to insulin glargine in the overall population and the two subpopulations. There were 10,430 patients in the overall population, 5,280 patients in the DEVOTE eligible population, and 5,150 patients in the DEVOTE ineligible population. The overall CPRD (HR: 1.36, 95% CI: 0.83, 1.86) and DEVOTE eligible populations (HR: 1.07, 95% CI: 0.63, 1.58) were compatible with findings from the DEVOTE trial (HR: 0.91, 95% CI: 0.78, 1.06) for the risk of MACE. Due to the low number of outcomes, the DEVOTE ineligible population had deviations in point estimates and wide CIs (HR: 2.19, 95% CI: 0.30, 3.83).

My thesis suggests that the risk of MACE among patients with T2DM newly prescribed insulin degludec compared to insulin glargine was consistent between the overall population and the subpopulation eligible for the DEVOTE trial, while the DEVOTE ineligible population had discrepant point estimates. These results suggest further research on the generalizability of results from trials to a real-world population.

RÉSUMÉ

Les essais de résultats cardiovasculaires (ECVT) sont un type d'essai contrôlé randomisé (ECR) utilisé pour éclairer la prise de décision concernant la sécurité cardiovasculaire des médicaments antidiabétiques pour le diabète sucré de type 2 (T2DM). Ces essais ont des critères d'inclusion stricts qui limitent leur population à des personnes atteintes d'une maladie cardiovasculaire (CV) établie ou présentant un risque élevé, ce qui entraîne des différences entre la population de l'essai et la population réelle qui utilise les médicaments. Peu d'études ont comparé les estimations de l'effet entre les ECR et l'émulation des ECR à l'aide de données du monde réel (DMR) pour les ECVT. À notre connaissance, aucune étude n'a examiné les différences potentielles entre les ECR et les DMR en ce qui concerne les caractéristiques des patients et les effets du traitement sur les résultats CV pour les analogues de l'insuline à longue durée d'action. Ma thèse avait deux objectifs principaux. Le premier était de synthétiser les données existantes sur l'émulation des ECVT à l'aide de la DMR pour les nouveaux médicaments antidiabétiques. Le second était d'émuler l'essai DEVOTE en utilisant la DMR pour déterminer le risque d'événements cardiovasculaires majeurs (ECVM) chez les patients atteints de DT2 prenant l'un des deux types d'analogues de l'insuline à longue durée d'action.

Dans le premier manuscrit, j'ai procédé à un examen systématique des études observationnelles et transversales qui émulaient les ECVT antérieures pour les inhibiteurs de la dipeptidyl peptidase-4 (DPP-4), les agonistes des récepteurs du glucagon-like peptide 1 (GLP-1) et les inhibiteurs du sodium glucose cotransporter-2 (SGLT-2) chez les patients atteints de diabète de type 2. J'ai recherché dans trois bases de données, de la création à juillet 2023, des études observationnelles axées sur l'émulation d'essais ou des études transversales faisant état de la proportion d'éligibilité dans le monde réel pour les ECVT achevées. Deux examinateurs indépendants ont sélectionné les articles, extrait les données et évalué la concordance entre la DMR et les ECR. Dix-neuf études ont été incluses dans notre examen, dont quatre études de cohortes qui ont reproduit des ECR antérieurs et 15 études transversales qui ont évalué l'admissibilité à l'essai. En reproduisant les ECVT, les résultats entre les ECR et les DMR étaient concordants pour toutes les classes de médicaments en ce qui concerne la démonstration de la non-infériorité. Le pourcentage d'éligibilité pour les essais sur les inhibiteurs du SGLT-2 variait de 7 % à 59 % et pour les essais sur les agonistes du récepteur du GLP-1, de 6,2 % à 53,6 %. Ces résultats suggèrent que si les ECR et les DMR concordent dans leurs estimations, les essais manquent de représentativité.

Dans le deuxième manuscrit, j'ai imité l'essai DEVOTE en utilisant les données du Clinical Practice Research Datalink (CPRD) du Royaume-Uni. J'ai estimé le risque de ECVM, un composite d'infarctus du myocarde, d'accident vasculaire cérébral ischémique ou de décès CV, chez les patients atteints de DT2 qui étaient de nouveaux utilisateurs d'insuline glargine ou d'insuline degludec. Deux sous-populations ont été créées sur la base des critères d'inclusion de l'essai DEVOTE. J'ai utilisé des modèles de risques proportionnels de Cox avec pondération de la probabilité inverse de traitement pour estimer les rapports de risque (HR) et les intervalles de confiance à 95 % (IC) bootstrap correspondants pour le ECVM, en comparant les nouveaux utilisateurs d'insuline degludec à l'insuline glargine dans la population globale et dans les deux sous-populations. La population CPRD comptait au total 10 430 patients (9 618 utilisateurs d'insuline glargine contre 812 utilisateurs d'insuline dégludec), la population éligible DEVOTE comptait 5 280 patients (4 904 utilisateurs d'insuline glargine contre 376 utilisateurs d'insuline dégludec) et la population inéligible DEVOTE comptait 5 150 patients (4 714 utilisateurs d'insuline glargine contre 436 utilisateurs d'insuline dégludec). Les populations CPRD (HR : 1,36, 95% CI : 0,83, 1,86) et DEVOTE éligibles (HR : 1,07, 95% CI : 0,63, 1,58) étaient compatibles avec les résultats de l'essai DEVOTE (HR : 0,91, 95% CI : 0,78, 1,06) en ce qui concerne le risque de ECVM. La population non éligible de l'essai DEVOTE (HR : 2,19, 95% CI : 0,30, 3,83).

Dans cette thèse, mes résultats suggèrent que le risque de ECVM chez les patients atteints de DT2 à qui l'on a nouvellement prescrit de l'insuline degludec par rapport à l'insuline glargine était cohérent entre la population globale et la sous-population éligible à l'essai DEVOTE, tandis que la population non éligible à l'essai DEVOTE présentait des estimations ponctuelles divergentes. Ces résultats suggèrent de poursuivre les recherches sur la généralisation des résultats des essais à une population réelle.

CONTRIBUTION OF AUTHORS

Wanning Wang, BHSc

I wrote the entirety of this thesis. I led the development and execution of the systematic review, from developing the protocol, conducting the search, screening articles, quality assessment, synthesizing findings, interpreting results to drafting the manuscript. I also led the database study, including developing the protocol, designing the study, creating variable definitions, participating in statistical analyses, interpreting results, and drafting the manuscript.

Kristian B. Filion, PhD

Dr. Filion supervised the development of my thesis. He conceived the thesis topic and advised on the design and execution of the systematic review and database study. He provided feedback on study design, analytical approach, interpretation of results, and reviewed both manuscripts and my thesis for important intellectual content.

Michael Webster-Clark, PhD

Dr. Webster-Clark was a member of my thesis committee. He provided his methodological expertise for the study design and analytical approaches for the development of both manuscripts. He reviewed the entirety of my thesis for important intellectual content.

Oriana H.Y. Yu, MD MSc

Dr. Yu was a member of my thesis committee. She provided her clinical expertise to the study design and interpretation of both manuscripts. She reviewed the entirety of my thesis for important intellectual content.

Pauline Reynier, MSc

Pauline Reynier was the statistical analyst for the database study. She provided insight on the application of statistical methods, performed the data management, and conducted statistical analyses for the study. She also reviewed the second manuscript for important intellectual information.

William Tang, BSc

William was a collaborator on the systematic review and served as the second reviewer. He screened studies for inclusion, extracted data, assessed study quality, and reviewed the manuscript for intellectual content.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my thesis supervisor, Dr. Kristian Filion. Thank you, Kris, for your steadfast guidance throughout my thesis journey. Thank you for taking the time to answer all my questions, no matter how trivial. I appreciate all that you have done to support and mentor me throughout these past two years. I have grown as a researcher, an epidemiologist, and a student because of you.

I would like to thank my committee members, Dr. Michael Webster-Clark and Dr. Oriana Yu. Thank you, Michael, for your patience in repeatedly explaining methodological concepts to me and shaping my thesis to what it is. Thank you, Oriana, for your encouragement and clinical expertise. I appreciate the time you both took to provide feedback on my thesis.

I would like to thank Pauline Reynier for all your help with the data management and analyses. Thank you for your diligence in performing the analyses and then re-running them numerous times until we got it right. I appreciate your patience in responding to each of my requests and helping troubleshoot the roadblocks we encountered along the way. It has been a pleasure working with you.

I would like to thank my peers at the Lady Davis Institute (LDI) for their comradery. From journal clubs, to navigating the CPRD, and to helping me interpret my results. Thanks for making the journey a little less daunting.

I would like to thank the staff at the LDI and EBOH offices for ensuring that all the background administrative processes were running smoothly. Thanks for helping me with funding and (hopefully) letting me graduate.

Last but not least, thank you to my friends (old and new) and family for being a constant pillar of support and encouragement. Thank you for being there when life inevitably got hard. I would not be where and who I am without all of you. Love you lots!

Table of Contents

LIST OF TABLES AND FIGURES

Tables

Figures

Supplementary Tables

Supplementary Figures

ABBREVIATIONS

A1C	Glycated hemoglobin
ADA	American Diabetes Association
ASCVD	Atherosclerotic cardiovascular disease
APC	Admitted Patient Care
ACE	Angiotensin-converting-enzyme
BNF	British National Formulary
CADTH	Canada's Drug and Health Technology Agency
CI	Confidence interval
CKD	Chronic kidney disease
COPD	Chronic obstructive pulmonary disease
CPRD	Clinical Practice Research Datalink
CV	Cardiovascular
CVD	Cardiovascular disease
CVOT	Cardiovascular outcome trial
Dm+d	Dictionary of Medicines and Devices
DPP-4	Dipeptidyl peptidase 4
eGFR	Estimated glomerular filtration rate
EMA	European Medicines Agency
EU	European Union
FDA	Food and Drug Administration
FPG	Fasting plasma glucose
GIP	Glucose-dependent insulinotropic polypeptide
GLP-1	Glucagon-like peptide-1
GLP-1 RA	Glucagon-like peptide-1 receptor agonists
GP	General practitioner
HES	Hospital Episode Statistics
HF	Heart failure
HR	Hazard ratio

ICD	International Classification of Disease and Related Health Problems
IOW	Inverse odd weights
IPTW	Inverse probability of treatment weights
ITT	Intention-to-treat
MACE	Major adverse cardiovascular events
MeSH	Medical Subject Headings
MI	Myocardial infarction
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NPH	Neutral protamine Hagedorn
OGTT	Oral glucose tolerance test
ONS	Office of National Statistics
OPCS	Office of Population, Censuses and Surveys' Classification of Surgical Operations and Procedures
OR	Odds ratio
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PS	Propensity score
RCT	Randomized controlled trial
RWD	Real-world data
RWE	Real-world evidence
SGLT-2	Sodium-glucose-cotransporter-2
SWiM	Synthesis without meta-analysis
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
UK	United Kingdom
US	United States

CHAPTER 1: INTRODUCTION

1.1 Type 2 Diabetes Mellitus

1.1.1 Pathophysiology of Type 2 Diabetes Mellitus

Diabetes mellitus is a set of metabolic disorders categorized by higher-than-normal levels of glucose in the blood, also known as hyperglycemia. Although there are different types of diabetes, more than 90% of those diagnosed with diabetes have type 2 diabetes mellitus (T2DM)[1]. Generally, T2DM is caused by impaired insulin secretion, insulin resistance, or a combination of the two[2,3]. The relative contribution of insulin deficiency and insulin resistance vary for each person[1].

Pancreatic beta-cells are the main component in the maintenance of blood glucose levels in the body. They are responsible for the synthesis, storage, and release of insulin[4,5]. Insulin is the hormone that assists with glucose uptake, and it guides the transportation of glucose in the body for energy use or to the liver for storage[6]. When beta cells are dysfunctional, insulin secretion is reduced, limiting glucose uptake and resulting in higher than normal levels of glucose in the bloodstream[5]. One cause of beta cell dysfunction is excess nutritional states, such as obesity, hyperglycemia, and hyperlipidemia. In these states, beta cells are subject to toxic pressures and cellular stress, which ultimately lead to lack of cell integrity and increased cell death[7]. Outside of the pancreas, target tissue, such as skeletal muscle, adipose tissue, and the liver, may have an impaired insulin response. Disturbances in cell signaling pathways or receptor functionality reduce insulin intake by tissue cells, leading to insulin resistance, excess glucose in the bloodstream, and ultimately development of T2DM[5].

1.1.2 Epidemiology of T2DM

The global prevalence of T2DM is rising rapidly, making it one of the most important public health challenges worldwide[8]. The number of people worldwide with T2DM has more than doubled in the past thirty years[9]. Globally, as of 2017, there is an estimated 462 million individuals affected by T2DM, corresponding to 6.3% of the world's population[10]. By 2030, it is projected that the prevalence of T2DM worldwide will increase to 7.7%, with developed countries seeing a 20% increase in adults affected by T2DM[11]. In Canada, the prevalence of T2DM among those aged 20-79 years was estimated to be 7.5% in 2011[12]. In the United Kingdom (UK), the population used

for the database study included in this thesis, the prevalence of T2DM increased from 3.1% in 2004 to 5.3% in 2014[13].

1.1.3 Risk Factors of T2DM

The rise in T2DM parallels the global obesity epidemic and adoption of a more sedentary lifestyle[14]. Risk factors for T2DM include a complex combination of genetic, metabolic, and environmental influences[5]. Obesity is found to be the strongest risk factor for T2DM, with a pronounced effect in younger adults[15,16]. Epidemiological studies have shown that a healthy lifestyle involving physical activity, proper diet, and weight loss strongly prevents the development of the disease[17,18]. Physical inactivity has been strongly linked to the incidence of T2DM, independent of obesity, hypertension, and parental history of diabetes[19,20].

Predisposition to the disease is also influenced by unmodifiable risk factors such as genetics and ethnicity. Several recent genome-wide association studies of T2DM have shown that the disease is polygenic[21]. In the UK and United States (US), Black, Asian, Latino and Native Americans have a far higher incidence and prevalence of T2DM than Caucasians [22,23]. Differences in risk among ethnicities are postulated to be due to a combination of genetics, socioeconomic, and lifestyle factors.

1.1.4 Diagnosis of T2DM

The monitoring of prediabetes and the diagnosis of T2DM is determined by sustained hyperglycemic states that are measured using laboratory tests. These tests include fasting plasma glucose (FPG) tests, oral glucose tolerance test (OGTT) or glycated hemoglobin A1C (A1C) measures. FPG measures the blood glucose level after fasting for at least 8 hours. OGTT checks the blood glucose levels before and two hours after ingesting a liquid drink containing a number of grams of glucose to evaluate one's ability to process sugar. A1C is an average measure of one's blood glucose level for the past three months[2]. **Table 1.1** describes the typical clinical diagnostic thresholds for each test. The decision regarding which test to use for diabetes diagnosis is left to clinical judgement[1]. Each test has advantages and disadvantages as presented on **Table 1.2**. If a patient is not symptomatic and presents a single laboratory result in the diabetes range, a repeat confirmatory laboratory test must be done on another day. This is shown to confirm diagnosis of diabetes in approximately 40% to 90% of people with an initial positive test[24,25].

Table 1.1 Diabetes Canada guidelines for T2DM diagnosis.

FPG ≥7.0 mmol/L Fasting = no caloric intake for at least 8 hours or **A1C ≥6.5% (in adults)** Using a standardized, validated assay in the absence of factors that affect the accuracy of the A1C or **2hPG in a 75 g OGTT ≥11.1 mmol/L** or **Random PG ≥11.1 mmol/L** Random = any time of the day, without regard to the interval since the last meal
2hPG, 2-hour plasma glucose; *AlC*, glycated hemoglobin; *FPG*, fasting plasma glucose; OGTT, oral glucose tolerance test; *PG*, plasma glucose.

Adapted from: Diabetes Canada Clinical Practice Guidelines Expert Committee, Punthakee Z, Goldenberg R, Katz P. Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. Can J Diabetes. 2018 Apr;42 Suppl 1:S10–5. Reproduced with permission from Diabetes Canada

Table 1.2 Advantages and disadvantages of different T2DM diagnostic tests.

Parameter	Advantages	Disadvantages
FPG	• Established standard • Fast and easy • Single sample • Predicts microvascular complications	• Sample not stable • High day-to-day variability • Inconvenient (fasting) • Reflects glucose homeostasis at a single point in time
2hPG in a 75 g OGTT	• Established standard • Predicts microvascular complications	• Sample not stable • High day-to-day variability • Inconvenient • Unpalatable • Cost
A1C	• Convenient (measure any time of day) • Single sample • Predicts microvascular complications • Better predictor of CVD than FPG or 2hPG in a 75 g OGTT • Low day-to-day variability • Reflects long-term glucose concentration	• Cost • Misleading in various medical conditions (e.g. hemoglobinopathies, iron deficiency, hemolytic anemia, severe hepatic or renal disease) • Altered by ethnicity and aging • Standardized, validated assay required • Not for diagnostic use in children and adolescents† (as the sole diagnostic test), pregnant women as part of routine screening for gestational diabetes‡, those with cystic fibrosis or those with suspected type 1 diabetes
Abbreviations*: 2hPG*, 2-hour plasma glucose; *A1C*, glycated hemoglobin; *CVD*, cardiovascular disease; *FPG*, fasting plasma glucose; *OGTT*, oral glucose tolerance test.		

Adapted from: Diabetes Canada Clinical Practice Guidelines Expert Committee, Punthakee Z, Goldenberg R, Katz P. Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. Can J Diabetes. 2018 Apr;42 Suppl 1:S10–5. Reproduced with permission from Diabetes Canada.

1.1.5 Clinical Complications of T2DM

There are two main types of clinical complications that manifest from T2DM: microvascular and macrovascular complications. Microvascular complications include retinopathy, nephropathy, and neuropathy. Macrovascular complications include cardiovascular disease (CVD), cerebrovascular disease and peripheral artery disease with cardiovascular events such as, heart failure (HF), ischemic stroke, and sudden cardiac arrest[30]. These complications are common, with 54% of patients presenting with microvascular complications and 27% with macrovascular complications in an observational study of 28 countries in Asia, Africa, South America, and Europe from 2009 to 2010[31]. The risk of developing these complications increases with disease duration and severity[32]. Diabetic retinopathy is the most common microvascular complication in diabetes and is responsible for around 10,000 new cases of blindness every year in the US alone[32,33]. Diabetic nephropathy is the leading cause of renal damage in the US[34], while diabetic peripheral neuropathy accounted for 80% of major amputations related to foot ulcers in the US[35].

Atherosclerosis is a key mechanism behind many of the aforementioned macrovascular complications, in which the arterial walls throughout the body narrow due to chronic inflammation and injury in the peripheral or coronary vascular systems[32]. The precise mechanism through which diabetes increases atherosclerotic plaque formation is not completely understood, yet the association is strong[32]. These complications contribute greatly to the morbidity and mortality in diabetes, with CVD being the primary cause of death in patients with T2DM[36]. CVD is also one of the largest components of healthcare expenditure in the US among patients with T2DM[37].

The risk of premature death among people with T2DM is almost double that of individuals those without T2DM[38]. In addition to death from CVD, T2DM is associated with premature death from several cancers, renal disease, liver disease, infectious diseases, mental disorders, intentional self-harm, degenerative disorders, and chronic obstructive pulmonary disorder[39–43].

1.1.6 Clinical Management of T2DM

According to the American Diabetes Association (ADA) Standards of Care in Diabetes 2023, management of T2DM involves adoption of healthy lifestyle behaviors and subsequently the initiation of pharmacotherapy which considers each patient's person-centered treatment[44]. To facilitate positive health behaviors and wellbeing, these lifestyle management techniques are

recommended: diabetes self-management education and support, medical nutrition therapy, routine physical activity, tobacco cessation counselling when needed, health behavior counseling, and psychosocial care[45].

In terms of pharmacotherapy, metformin is the recommended first-line therapy[46,47]. It is effective, safe, inexpensive and has evidence of reducing the risk of death[48]. Compared with sulfonylurea, metformin as a first-line therapy has shown beneficial effects on A1C levels, weight loss, and cardiovascular mortality[49]. Based on each patient's comorbidities and treatment goals, clinicians should consider metformin, other agent(s), or combination therapy to achieve and maintain these goals. In adults with T2DM and established/high risk of atherosclerotic cardiovascular disease (ASCVD), HF, and/or chronic kidney disease (CKD), the treatment regimen should include agents that reduce cardiorenal risk. These may include sodium-glucose-cotransporter-2 (SGLT-2) inhibitors or glucagon-like peptide-1 receptor agonists (GLP-1 RA). Higher-efficacy approaches have greater likelihood of achieving glycemic goals, with the following considered to have very high efficacy for glucose lowering: the GLP-1 RA dulaglutide (high dose) and semaglutide, the gastric inhibitory peptide (GIP) agonist and GLP-1 RA tirzepatide, insulin, combination oral therapy, and combination injectable therapy[50]. Dipeptidyl peptidase 4 (DPP-4) inhibitors have shown intermediate efficacy. Weight management is also an important component to glycemic control in T2DM. Among drugs used to treat T2DM, very high efficacy for weight loss has been seen with semaglutide and tirezepatide[50]. Insulin is often recommended to patients later in disease progression who have difficulty controlling their blood glucose levels using oral antidiabetic drugs or GLP-1 RAs[51]

1.2 Cardiovascular Safety of Antidiabetic Medications

In 2008, the US Food and Drug Administration (FDA) issued recommendations for cardiovascular (CV) safety trials to be conducted for newly-marketed antidiabetic medications[48]. These recommendations were made in the wake of concerns over increased risk of adverse CV events associated with the thiazolidinedione rosiglitazone[49]. Broadly speaking, the FDA recommendations require the conduct of a placebo-controlled non-inferiority trial and set specific upper limit confidence interval (CI) thresholds (1.8 for pre-approval and 1.3 for post-approval) for the risk ratio in determining the CV risk for the newly marketed drug. The FDA also requires the inclusion of patients with advanced T2DM, a minimum of 2 years of CV safety data, and that all phase 2 and phase 3 trials examine CV events, one of which must be major adverse cardiovascular events (MACE), a composite endpoint that includes CV mortality, nonfatal myocardial infarction (MI), and nonfatal ischemic stroke[50,51]. Similar requirements have also been established by the European Medicines Agency (EMA)[52]. Since then, more than 13 cardiovascular outcome trials have been conducted on antidiabetic drugs for T2DM, including DPP-4 inhibitors, GLP-1 RAs, and SGLT-2 inhibitors[53]. Of note, these regulatory agencies do not require the conduct of CVOTs for insulins.

1.2.1 DPP-4 Inhibitors

The enzyme DPP-4 indirectly regulates blood glucose levels. It accepts incretin hormones, glucagon-like peptide-1 (GLP-1) and GIP, as substrates and degrades them into metabolites[54]. GLP-1 has many glycemic control benefits, including inducing insulin secretion and supressing glucagon secretion[55]. GIP has similar actions in increasing glucose-dependent insulin secretion as GLP-1[55]. In addition, GIP enhances glucagon responses to low glucose levels counteracting insulin-induced hypoglycemia[56]. Thus, DPP-4 inhibition decreases the catalytic activity of DPP-4 on these incretin substrates, allowing for increased concentration of these hormones in the body and improved glucose homeostasis[54].

In terms of glycemic control, a meta-analysis of 25 trials showed a reduction in A1C with DPP-4 inhibitors compared to placebo[57]. However, in a meta-analysis of 19 randomized controlled trials (RCTs), when compared to other antidiabetic medications such as metformin and second-line treatments (sulfonylureas and GLP-1 RAs), there was a smaller decline in A1C with DPP-4 inhibitors[58]. When added to metformin, DPP-4 inhibitors had greater weight loss than

sulfonylureas or pioglitazone alone. In addition, DPP-4 inhibitors resulted in lower discontinuation rates and fewer cases of diarrhea, vomiting, or nausea compared to patients receiving metformin or GLP-1 RAs[58]. DPP-4 inhibitors have been shown to be suitable for use in elderly, frail and/or vulnerable patients, with longstanding disease and multiple comorbidities[54,59].

1.2.1.1 DPP-4 Inhibitors: Cardiovascular Outcome Trials

There have been several completed CVOTs for DPP-4 inhibitors to date. These include, EXAMINE (alogliptin)[60], SAVOR-TIMI 53 (saxagliptin)[61], TECOS (sitagliptin)[62], CAROLINA (linagliptin)[63] and CARMELINA (linagliptin)[64], which were designed with the primary objective of evaluating non-inferiority. **Table 1.3** provides a summary of these trials.

The EXAMINE trial was the only one to enroll patients with T2DM after an acute coronary syndrome; all other trials included populations with elevated CVD risk or established CVD. Most studies compared the target drug to a placebo except for CAROLINA, which used a sulfonylurea as an active comparator[63]. All trials other than TECOS used a 3-point MACE outcome; TECOS used a 4-point MACE composite that included hospitalization for unstable angina. Duration of follow-up ranged from 1.5 years (EXAMINE) to 6.3 years (CAROLINA)[60,63]. All studies had a neutral effect, achieving non-inferiority.

In all but one trial, secondary outcomes of the components of MACE, death from any cause, death from cardiovascular causes or hospitalization due to HF did not show differences between the two groups[60,62–64]. Of note in the SAVOR-TIMI-53 trial, more patients randomized to saxagliptin were hospitalized for HF than among those randomized to placebo (hazard ratio [HR]: 1.27, 95% CI:1.07, 1.51)[61]. This safety signal was not corroborated in a subsequent multi-jurisdictional nested case-control study by Filion et al[65].

A meta-analysis that pooled data from SAVOR-TIMI 53, EXAMINE, TECOS, and CARMELINA showed no effect of DPP-4 inhibitors on MI (odds ratio [OR]: 1.01, 95% CI: 0.92, 1.10), stroke (OR: 0.99, 95% CI: 0.87, 1.13), and CV death (OR: 0.99, 95% CI: 0.91, 1.09) individually[66]. Taken together, DPP-4 inhibitors have demonstrated non-inferiority to placebo or active comparators for CV safety but have not shown any marked benefits on 'hard' clinical events.

Table 1.3 Summary of key features of DPP-4 inhibitor cardiovascular outcome trials

RCT	Inclusion Criteria	No. of patients	Follow up time	Target drug vs comparator	Primary Outcome	MACE HR (95% CI)	Secondary Outcomes HR (95% CI)
EXAMINE[60]	Patients with T2DM after an acute coronary syndrome event.	5,380	Median 1.5 years	Alogliptin vs placebo	3P MACE	0.96 (≤1.16)	• CV death: 0.79 (0.60, 1.04) • Death from any cause: 0.88 (0.71, 1.09) • Nonfatal MI: 1.08 (0.88, 1.33) • Nonfatal stroke: 0.91 (0.55, 1.50)
SAVOR-TIMI 53[61]	Patients with T2DM who had a history of, or were at risk for, cardiovascular events.	16,492	Median 2.1 years	Saxagliptin vs placebo	3P MACE	1.00 (0.89, 1.12)	• CV death: 1.03 (0.87, 1.22) • Death from any cause: 1.11 (0.96, 1.27) • HHF: 1.27 (1.07, 1.51) • MI*: 0.95 (0.80, 1.12) • Ischemic stroke: 1.11 (0.88, 1.39)
TECOS[62]	Patients with T2DM who had established cardiovascular disease or were over 50 years of age with A1C levels controlled by oral antihyperglycemic agents or insulin.	14,671	Median 3.0 years	Sitagliptin vs placebo	4P MACE	0.98 (0.89, 1.08)	• CV Death: 1.03 (0.89, 1.19) • Death from any cause: 1.01 (0.90, 1.14) • HHF: 1.00 (0.83, 1.20) • MI*: 0.95 (0.81, 1.11) • Stroke**: 0.97 (0.79, 1.19)
CAROLINA[63]	Patients with T2DM who were of elevated cardiovascular risk.	6,042	Median 6.3 years	Linagliptin vs glimepiride (sulfonylurea)	3P MACE	0.98 (0.84, 1.14)	• CV Death: 1.00 (0.81, 1.24) • Death from any cause: 0.91 (0.78, 1.06) • HHF: 1.21 (0.92, 1.59) • MI*: 1.03 (0.82, 1.29) • Stroke**: 0.86 (0.66, 1.12)
CARMELINA[64]	Patients with T2DM with high cardiovascular or renal risk were included.	6,979	Median 2.2 years	Linagliptin vs placebo	3P MACE	1.02 (0.89, 1.17)	• CV Death: 0.96 (0.81, 1.14) • Death from any cause: 0.98 (0.84, 1.13) • HHF: 0.90 (0.74, 1.108) • MI*: 1.12 (0.90, 1.40) • Stroke**: 0.91 (0.67, 1.23)

*Fatal or non-fatal MI; ** Fatal or nonfatal stroke. Abbreviations: HR: Hazard ratio; 3P MACE: 3-point major adverse cardiovascular events, composite of myocardial infarction, cardiovascular death, or stroke; 4P MACE: 4-point major adverse cardiovascular event, composite of myocardial infarction, cardiovascular death, stroke, or hospitalization for unstable angina; CV: Cardiovascular; HHF: Hospitalization due to heart failure; MI: Myocardial infarction

1.2.2 GLP-1 Receptor Agonists

GLP-1 has many physiological and pharmacological actions that enhance the body's natural response to food and reduce glucose levels[67,68]. Most notably, GLP-1 increases insulin release only in the context of hyperglycemia and therefore, does not induce hypoglycemia[69]. It also supresses glucagon release from pancreatic alpha cells, controlling the release of glucose into the bloodstream[70]. In fasting patients with T2DM, the insulin and glucagon activity of GLP-1 contribute equally to reducing blood glucose levels[71]. After a meal, glucose control is thought to be largely mediated through GLP-1 delaying gastric emptying, which slows the entry of nutrients such as glucose into the bloodstream[72]. GLP-1, with the ability to decrease glucagon secretion, decrease hepatic gluconeogenesis, improve insulin sensitivity, and delay gastric emptying, is thought to potentially promote central satiety and reduce overall caloric intake[73]. These mechanisms have shown to be effective in reducing blood sugar levels and managing obesity.

GLP-1 RAs mimic the effect of GLP-1 on the body. There are two types in production, short-acting and long-acting. Short-acting GLP-RAs activate the GLP-1 receptor for only 6 hours after each injection[68]. They are recommended before a meal and have a larger effect in slowing gastric emptying than reducing fasting glucose levels[68]. Conversely, long-acting GLP-1 RAs remain in the bloodstream between doses. They provide better glycemic control than short-acting GLP-1 RAs and rely more on increasing postprandial insulin concentrations than decreasing gastric emptying[74,75].

A meta-analysis of 33 studies of GLP-1 RAs demonstrated a decrease in A1C and FPG levels, with greater mean reductions in FPGs among long-acting GLP-1 RAs compared to short-acting ones[76]. In addition, body weight substantially decreases in a dose-dependent manner with all of the long-acting GLP-1 RAs[76,77]. Results varied from 1.66 kg (liraglutide once daily) to 2.41kg (exenatide once a week), decrease in mean weight[76]. Gastrointestinal effects, including nausea and vomiting, are the most frequently reported adverse effects seen with long-acting GLP-1 RAs. Side effects occur early during treatment but tend to be transient[78–80].

1.2.2.1 GLP-1 RAs: Cardiovascular Outcome Trials

To date there have been 7 CVOTs for long-acting GLP-1 RAs. These include LEADER (liraglutide)[81], ELIXA (lixisenatide)[82], SUSTAIN-6 (semaglutide)[83], EXSCEL (exenatide)[84], HARMONY Outcomes (albiglutide)[85], REWIND (dulaglutide)[86], and PIONEER-6 (oral

semaglutide)[87]. These trials were designed with the primary objective of assessing non-inferiority. **Table 1.4** provides a summary of the trials.

The trials varied in patient population with all including older adults with CV risk or previous CV events. Each trial compared their drug of interest to placebo with a median follow-up ranging from 1.3 years (PIONEER-6) to 5.4 years (REWIND)[86,87]. All studies examined 3-point MACE except ELIXA, which included hospitalization for unstable angina in their 4-point MACE outcome[81–87]. All studies demonstrated noninferiority with many also demonstrating superiority. For secondary outcomes of the individual components of MACE, hospitalization from heart failure and all-cause mortality, results varied across trials.

A systematic review and meta-analysis of the 7 GLP-1 RA CVOTs demonstrated a 12% reduction in MACE with GLP1 RAs versus placebo (HR: 0.88, 95% CI: 0.82, 0.94)[88]. There were reductions in CV death (HR: 0.88, 95% CI: 0.81, 0.96), fatal or non fatal stroke (HR:0.84, 95% CI: 0.76, 0.93), fatal or non-fatal MI (HR: 0.91, 95% CI: 0.84, 1.00), all-cause mortality (HR:0.88, 95% CI: 0.83, 0.95), and hospitalization due to HF (HR: 0.91, 95% CI: 0.83, 0.99)[88]. There were also reductions in composite kidney outcomes (HR:0.83, 95% CI: 0.78, 0.89)[88]. Taken together, GLP-1 RAs have shown strong evidence of cardiovascular and renal safety.

Table 1.4 Summary of key features of GLP-1 RA cardiovascular outcome trials

RCT	Inclusion Criteria	No of patients	Follow up time	Target drug vs comparator	Primary Outcome	MACE HR (95%CI)	Secondary Outcomes HR (95% CI)
LEADER[81]	Patient with T2DM who were aged 50 or over with known cardiovascular disease or aged 60 years or older with multiple cardiovascular risk factors	9,340	Median 3.8 years	Liraglutide vs placebo	3P MACE	0.87 (0.78, 0.97)	• CV death: 0.78 (0.66, 0.93) • Death from any cause: 0.85 (0.74, 0.97) • HHF: 0.87 (0.73, 1.05) • MI*: 0.86 (0.73, 1.00) • Stroke**: 0.86 (0.71, 1.06)
ELIXA[82]	Patient with T2DM who had sustained an acute coronary event within 180 days before randomization	6,068	Median 2.1 years	Lixisenatide vs placebo	4P MACE	1.02 (0.89, 1.17)	• CV death: 0.98 (0.78, 1.22) • Death from any cause: 0.94 (0.78, 1.13) • HHF: 0.96 (0.75, 1.23) • MI*: 1.03 (0.87, 1.22) • Stroke**: 1.12 (0.79, 1.58)
SUSTAIN-6[83]	Patient with T2DM who were at high CV risk	3,297	Median 2.1 years	Semaglutide vs placebo	3P MACE	0.74 (0.58, 0.95)	• CV Death: 0.98 (0.65, 1.48) • Death from any cause: 1.05 (0.74, 1.50) • HHF: 1.11 (0.77, 1.61) • MI*: 0.74 (0.51, 1.08) • Nonfatal stroke: 0.61 (0.38, 0.99)
EXSCEL[84]	Patient with T2DM with and without CV risk factors	14,752	Median 3.2 years	Exenatide vs placebo	3P MACE	0.91, (0.83, 1.00)	• CV Death: 0.88 (0.76, 1.02) • Death from any cause: 0.86 (0.77, 0.97) • HHF: 0.94 (0.78, 1.13) • MI*: 0.97 (0.85, 1.10) • Stroke**: 0.85 (0.70, 1.03)
HARMONY Outcomes[85]	Patient with T2DM who were 40 years of age and had CV disease	9,643	Median 1.5 years	Albiglutide vs placebo	3P MACE	0.78 (0.68, 0.90)	• CV Death: 0.93 (0.73, 1.19) • Death from any cause: 0.95 (0.79, 1.16) • MI*: 0.75 (0.61, 0.90) • Stroke**: 0.86 (0.66, 1.14)

REWIND[86]	Patient with T2DM who were at least 50 years old and had either a previous CV event or CV risk factors were included in the trial	9,901	Median of 5.4 years	Dulaglutide vs placebo	3P MACE	0.88 (0.79, 0.99)	• CV Death: 0.91 (0.78, 1.06) • Death from any cause: 0.90 (0.80, 1.01) • HHF: 0.93 (0.77, 1.12) • MI*: 0.96 (0.79, 1.16) • Stroke**: 0.76 (0.62, 0.94)
PIONEER-6[87]	Patient with T2DM who were at high CV risk	3,183	Median 1.3 years	Oral semaglutide vs placebo	3P MACE	0.79 (0.57, 1.11)	• CV Death: 0.49 (0.27, 0.92) • Death from any cause: 0.51 (0.31, 0.84) • HHF: 0.86 (0.48, 1.55) • Nonfatal MI: 1.18 (0.73, 1.90) • Nonfatal stroke: 0.74 (0.35, 1.57)

*Fatal or non-fatal MI; ** Fatal or nonfatal stroke. Abbreviations: HR: Hazard ratio; 3P MACE: 3-point major adverse cardiovascular events, composite of myocardial infarction, cardiovascular death, or stroke; 4P MACE: 4-point major adverse cardiovascular event, composite of myocardial infarction, cardiovascular death, stroke, or hospitalization for unstable angina; CV: Cardiovascular; HHF: Hospitalization due to heart failure; MI: Myocardial infarction

1.2.3 SGLT-2 inhibitors

Sodium-glucose co-transporters play a key role in glucose regulation. In the kidney, SGLT-1 and SGLT-2 actively transport glucose across cells to be reabsorbed by the bloodstream[93]. SGLT-2 accounts for approximately 90% of reabsorbed glucose, and when the receptors are saturated, the extra glucose is excreted in the urine[94]. In T2DM, due to the presence of hyperglycemia, SGLT receptors are at their limit, resulting in excess glucose in the urine. In addition, SGLT-2 is expressed at higher levels in patients with T2DM, leading to increased renal uptake of glucose into the body[95]. The goal of SGLT-2 inhibitors is to supress glucose reabsorption. This attenuation increases glucose excretion and reduces blood glucose levels[94]. As this mode of action behaves independently of insulin and pancreatic beta-cell function, there is limited loss of potency in SGLT-2 inhibitors upon the deterioration of beta-cells, which is often observed with other types of glucose-lowering agents[96]. Furthermore, since SGLT-2 inhibitors do not directly interfere with glucose production nor stimulates insulin release, they do not increase the risk of hypoglycemia[96].

Several meta-analyses have shown the clinical efficacy of SGLT-2 inhibitors for glycemic control. In one meta-analysis of 17 double-blind RCTs, patients randomized to SGLT-2 inhibitors alone or in combination with other antidiabetic drugs saw substantial reductions in mean changes in A1C, FPG, body weight, and blood pressure compared to placebo or placebo with other antidiabetic drugs[97]. Other meta-analyses of placebo and active comparator trials of SGLT-2 drugs have shown similar results[98,99]. SGLT-2 inhibitors were also shown to be associated with an increased occurrence of adverse events such as urinary tract and genital tract infections[98]. In 2015, the FDA issued warnings to labels of SGLT-2 inhibitors to include the increased risk of euglycemic diabetic ketoacidosis and serious urinary tract infections[100]. Currently marketed SGLT-2 inhibitors are indicated as monotherapy for patients with T2DM and inadequate glycemic control from diet and exercise (US and European Union [EU] indications), who are unable to use metformin (EU specific), and as an add-on therapy with other glucose-lowering agents, including insulin (EU and Canada)[96,101,102].

1.2.3.1 SGLT-2 Inhibitors: Major Cardiovascular Outcome Trials

There have been several CVOTs evaluating the safety of SGLT-2 inhibitors. These include EMPA-REG Outcome (Empagliflozin)[103], CANVAS (Canagliflozin) [104], DECLARE (Dapagliflozin)[105], CREDENCE (Canagliflozin)[106], VERTIS-CV (Ertugliflozin)[107] which are

summarized in **Table 1.5**. These trials were conducted with the primary objective of evaluating non-inferiority.

Each trial had specific inclusion criteria based on cardiovascular disease or risk factors, or renal risk factors, and participants were followed up for a median or mean of 2 to 4 years for 3-point MACE. All trials demonstrated non-inferiority and were thus considered safe in comparison to placebo. For secondary outcomes, many of the trials saw reductions in CV death, hospitalization due to HF, and all cause mortality. There has been documented risk of diabetic ketoacidosis and the CANVAS study also indicated an increase in amputations[104,108].

In a recent meta-analysis of the six cardiovascular outcome trials for SGLT-2 inhibitors, there was a reduction of risk of MACE with SGLT-2 inhibitor use (HR: 0.90, 95% CI: 0.85, 0.95)[109]. In addition, there were consistent benefits for hospitalization due to HF across the trials (HR 0.68; 95% CI: 0.61, 0.76)[109]. This evidence suggests that SGLT-2 inhibitors reduce the risk of cardiovascular outcomes in patients with T2DM.

Table 1.5 Summary of key features of SGLT-2 inhibitor cardiovascular outcome trials

RCT	Inclusion Criteria	No of patients	Follow up time	Target drug vs comparator	Primary Outcome	MACE HR (95%CI)	Secondary Outcomes HR (95% CI)
EMPA-REG Outcome[99]	Patients with T2DM with known cardiovascular disease	7,020	Median 3.1 years	Empagliflozin vs placebo	3P MACE	0.86 (0.74, 0.99)	• CV death: 0.62 (0.49, 0.77) • Death from any cause: 0.68 (0.57, 0.82) • HHF: 0.65 (0.50, 0.85) • MI*: 0.87 (0.70, 1.09) • Stroke**: 1.18 (0.89, 1.56)
CANVAS[100]	Patients with T2DM with symptomatic atherosclerotic cardiovascular disease, or without known history of cardiovascular disease but with significant risk factors	10,142	Median 2.4 years	Canagliflozin vs placebo	3P MACE	0.86 (0.75, 0.97)	• CV death: 0.87 (0.72, 1.06) • Death from any cause: 0.87 (0.74, 1.01) • HHF: 0.67 (0.52, 0.87) • MI*: 0.89 (0.73, 1.09) • Stroke**: 0.87 (0.69, 1.09)
DECLARE[101]	Patients with T2DM with established cardiovascular disease or cardiovascular risk factors	17,160	Median 4.2 years	Dapagliflozin vs placebo	3P MACE	0.93 (0.84, 1.03)	• CV Death: 0.98 (0.82, 1.17) • Death from any cause: 0.93 (0.82, 1.04) • HHF: 0.73 (0.61, 0.88) • MI*: 0.89 (0.77, 1.01) • Ischemic stroke: 1.01 (0.84, 1.21)
CREDENCE[102]	Patients with T2DM with chronic kidney disease	4,401	Median 2.62 years	Canagliflozin vs placebo	3P MACE	0.80, (0.67, 0.95)	• CV Death: 0.78 (0.61, 1.00) • Death from any cause: 0.83 (0.68, 1.02) • HHF: 0.61 (0.47, 0.80)
VERTIS-CV[103]	Patients with T2DM with atherosclerotic cardiovascular disease involving the coronary, cerebrovascular or peripheral arterial systems	8,246	Mean 3.5 years	Ertugliflozin vs placebo	3P MACE	0.97 (0.85, 1.11)	• CV Death: 0.92 (0.77, 1.11) • Death from any cause: 0.93 (0.80, 1.08) • HHF: 0.70 (0.54, 0.90) • MI*: 1.04 (0.86, 1.27) • Non-fatal stroke: 1.00 (0.76, 1.32)

*Fatal or non-fatal MI; ** Fatal or nonfatal stroke. Abbreviations: HR: Hazard ratio; 3P MACE: 3-point major adverse cardiovascular events, composite of myocardial infarction, cardiovascular death, or stroke; 4P MACE: 4-point major adverse cardiovascular event, composite of myocardial infarction, cardiovascular death, stroke, or hospitalization for unstable angina; CV: Cardiovascular; HHF: Hospitalization due to heart failure; MI: Myocardial infarction

1.2.4 Long-acting Insulins

With prolonged diabetes, the pancreatic cells responsible for producing endogenous insulin begin to fail. Exogenous insulin may be administered to achieve and maintain glycemic control[106]. There are two broad classifications of insulin: human insulin and insulin analogues[107,108]. These types of insulin have varying duration of effects on the body. Rapid- and short-acting insulin are forms of bolus insulin administered before a meal. They control the increase of glucose with the ingestion of a meal and have strong but transient effects on glucose absorption. Whereas intermediate and long-acting insulins are absorbed slowly and maintain a constant level of insulin in the body[109]. Both human and analog insulins bind to insulin receptors on target tissue to facilitate glucose absorption[47]. The following thesis will focus on intermediate-acting insulin (human: neutral protamine Hagedorn [NPH]), long-acting insulins (analogue: glargine, detemir) and ultra-long-acting insulins (analogue: degludec), which will be combined and termed long-acting insulins from this point forward.

Long-acting insulin analogues and NPH insulin have demonstrated comparable A1C control[110]. In a meta-analysis of RCTs comparing glycemic control of insulin glargine to NPH, the two treatment groups were similar with respect to the proportion of patients reaching target A1C $\leq$ 7.0% (30.8% and 32.1% respectively)[111]. A Cochrane review further supported these findings[112]. However, a meta-analysis of 14 RCTs by Monami et al. demonstrated that NPH was superior (by 0.1%) at reducing A1C over insulin detemir, but not over insulin glargine[113].

Hypoglycemia is the most common adverse effect of insulin therapy[109]. Abnormally low blood glucose levels could lead to negative health outcomes progressing from sweating and palpitations to cognitive dysfunction, seizures, coma, and death[114]. Several meta-analyses have compared the hypoglycemia risk of various types of insulin analogues[111,113,115–118]. One meta-analysis of 4 RCTs found that insulin glargine resulted in reduced overall symptomatic hypoglycemia, nocturnal hypoglycemia, and severe hypoglycemia[111]. In two meta-analyses of RCTs comparing the risk of hypoglycemia in patients using insulin degludec versus insulin glargine, both studies found that insulin degludec was associated with a reduction in the risk of all confirmed hypoglycemia events[117,118]. However, compared to oral antidiabetic drugs, NPH and long-acting insulins were associated with increased risks of hypoglycemic events[119].

Insulin is typically prescribed when patients are unable to control their blood glucose level using oral antidiabetic medications or GLP-1 RAs. These patients are usually older and are in an advanced disease state, with insulin therapy administered 10-15 years after diagnosis[47]. European, American, and Canadian guidelines recommend treating patients needing constant insulin control with either NPH insulin or long-acting insulin[42,46,120].

1.2.4.1 Long-Acting Insulins: Cardiovascular Outcome Trials

There are only two trials designed to evaluate the cardiovascular effects of new long-acting insulins, ORIGIN[121] and DEVOTE[122], as insulin therapy was exempt from the CVOT requirement of the FDA[123]. The ORIGIN trial commenced before the establishment of the FDA guidelines but nevertheless provided information on CV risk of insulin glargine. The DEVOTE study was conducted as part of the FDA preapproval requirements for insulin degludec based on the results of its phase 2 and 3 meta-analysis[124,125]. Both trials aimed to assess non-inferiority. **Table 1.6** provides a summary of the trials.

The ORIGIN trial randomized 12,537 participants with elevated cardiovascular risk factors and prediabetes or T2DM to receive insulin glargine or standard of care[121]. The DEVOTE trial compared the long-acting insulins degludec and glargine among 7,637 participants[122] with T2DM and risk factors for CVD, established CVD or CKD. Follow up was 6.2 years for the ORIGIN trial and 2 years for the DEVOTE trial. In the ORIGIN trial, there was no association between insulin glargine and MACE (HR: 1.02. 95% CI: 0.94, 1.11)[121]. In the DEVOTE trial, a similar risk of MACE was observed between the two groups (HR: 0.91, 95% CI: 0.78, 1.06)[122]. For both trials, the individual components of MACE and hospitalization due to HF did not show any difference between treatment groups[121,122].

Table 1.6 Summary of key features of long-acting insulin cardiovascular outcome trials

RCT	Inclusion Criteria	No of patients	Follow up time	Target drug vs comparator	Primary Outcome	MACE HR (95%CI)	Secondary Outcomes HR (95% CI)
ORIGIN[121]	Patients with elevated cardiovascular risk factors and prediabetes or T2DM	12,537	Median 6.2 years	Insulin glargine vs standard of care	3P MACE	1.02 (0.94, 1.11)	• CV death: 1.00 (0.89, 1.13) • Death from any cause: 0.98 (0.90, 1.08) • HHF: 0.90 (0.77, 1.05) • MI*: 1.02 (0.88, 1.19) • Stroke**: 1.03 (0.89, 1.21)
DEVOTE[122]	T2DM patients with risk factors for cardiovascular disease, established cardiovascular disease or chronic kidney disease	7637	Median 1.99 years	Insulin degludec vs insulin glargine	3P MACE	0.91 (0.78, 1.06)	• CV death: 0.96 (0.76, 1.21) • Death from any cause: 0.91 (0.76, 1.11) • Nonfatal MI: 0.85 (0.68, 1.06) • Nonfatal stroke: 0.90 (0.65, 1.23)

*Fatal or non-fatal MI; ** Fatal or nonfatal stroke. Abbreviations: HR: Hazard ratio; 3P MACE: 3-point major adverse cardiovascular events, composite of myocardial infarction, cardiovascular death, or stroke; CV: Cardiovascular; HHF: Hospitalization due to heart failure; MI: Myocardial infarction

1.3 Comparison of Real-World Evidence and RCTs for Cardiovascular Safety of Antidiabetic Medications

1.3.1 RCTs and External Validity

RCTs are the gold-standard in determining the efficacy of medical interventions. Evidence from RCTs is used by regulatory decision-makers for drug approval, by healthcare providers to guide clinical decisions, and by policy makers to support recommendations for the adoption of new therapies in clinical practice[126]. RCTs are typically designed to maximize internal validity to reduce the potential of bias regarding the effect of an intervention. They are conducted in highly controlled settings with rigorous adherence to structured protocols, strict inclusion/exclusion criteria and participant randomization[127]. For the results of the RCTs to be useful to everyday clinical practice, external validity (or generalizability) should also be considered to ensure that findings from the trial can be applicable to a relevant, definable patient population in a specific healthcare setting in routine practice[128]. However, it is challenging to balance both internal and external validity; most current RCTs prioritize the former at the expense of the latter. To address external validity, pragmatic trials have emerged to evaluate the effectiveness of a medical interventions in a setting more representative of routine clinical practice[129].

Rothwell outlines concerns of generalizability throughout the trial design[128]. At the trial setting level, differences in health-care system may affect the quality and timeliness of care delivered to patients, impacting the time of randomization and intervention, especially for outcomes that are time sensitive. Even before consideration of eligibility, the setting in which the trial is conducted may differ from where typical patients are seen for the disease. For example, trials are typically done in large academic centers, whereas care for diseases like hypertension and diabetes is usually provided at primary care centers. Those who are seeking care from large hospital centers may be different than those who access care from primary care providers[128]. When selecting patients, concerns also arise as only a small proportion of patients with a specific disease participate in a particular trial. Differences may exist beyond the already restrictive nature of eligibility criteria. Patients recruited into RCTs differ from those who are eligible but not recruited in terms of age, sex, race, severity of disease, educational status, social class, and place of residence. Outcomes and follow up time may also impact generalizability[128].

With these important differences, a literature review of 46 studies from the fields of cardiology, oncology, and mental health evaluated the representativeness of RCT populations in the real-world[127]. The review found that 71% of the studies had population differences which may have a relevant impact on the external validity of the RCT findings. They identified implicit and explicit factors that influenced the external validity of RCTs. Explicit factors were the use of restrictive inclusion and exclusion criteria, where high-risk patients were excluded from RCTs. Implicit factors were the ability to obtain informed consent, patient-related factors such as study participation, beliefs and attitudes regarding the safety of trial medications, cultural factors, level of satisfaction with current treatment, and willingness to participate[127].

1.3.2 Current use of Real-World Evidence

Given the strict nature of RCTs, there is a need to evaluate the safety and effectiveness of medical interventions in real-world populations. Large amounts of real-world data (RWD) on drug exposure and health outcomes are becoming readily available with the adoption of electronic health records and administrative healthcare databases to generate real-world evidence (RWE). Using RWE eases concerns over external validity with the inclusion of broader patient populations, expanded sample sizes, and longer follow-up periods, while providing perspectives from health care providers, patients, and caregivers on issues related to drug accessibility, acceptability, and preferences[130]. However, the causal inference from RWD studies may be impacted as RWD are limited by their susceptibility to bias and confounding. Treatment assignment is based upon physician judgement rather than random assignment, which may lead to differences in treatment assignment based on key risk factors for the targeted outcome. In addition, RWE can vary largely in quality. Thus, reaching appropriate conclusions from RWE requires transparent reporting and careful interpretation of the RWD source, study design and methods[130].

With these strengths and limitations in mind, regulatory and health technology assessment agencies are increasingly using RWE to complement evidence generated by RCTs for the evaluation of the effectiveness and safety of medical interventions. Health Canada has launched an initiative to integrate RWE throughout the lifecycle of drugs, and the Canada's Drug and Health Technology Agency (CADTH) has recently created guidelines for reporting RWE[130,131]. In response to CADTH's guidelines, Health Canada stated that they are open to relying on RWD/RWE in certain situations such as expanding evidence-based indications for populations often excluded from

clinical trials (e.g., children, older adults and expectant mothers), addressing diseases where clinical trials are not feasible (e.g., rare diseases) or responding to emergencies where clinical trials are unethical[132]. Given the unique considerations required for regulatory decision-making, the usefulness of RWD/RWE in regulatory decision-making will be determined case by case. Decisions will be based on the methodologies used to generate evidence, as well as the reliability and relevance of the RWE[132]. Similarly, in the FDA's August 2023 guidelines on using RWE for drug approvals, they encourage sponsors to consult the agency early in the drug development process if they are planning to use RWE to evaluate the appropriateness and potential challenges of such an approach[133].

Historically, the FDA has used RWE post-market to evaluate safety and on rare occasions, to inform decisions about effectiveness. The FDA's RWE Program, established in December 2018, aims to examine the potential use of RWD/RWE to support regulatory decisions about product effectiveness[134]. The program evaluated the potential use of RWE to support changes to labeling about drug product effectiveness, including adding or modifying an indication, such as a change in dose, dose regiment or route of administration; adding a new population; or adding comparative effectiveness or safety information. Some ways that RWD are being evaluated by the FDA is for pragmatic trials with broader patient populations based on randomizing individuals from large administrative databases, for non-randomized single armed trials using RWD for an external control arm, and for observational studies[134]. They highlight the need to replicate results of RCTs using rigorously designed observational studies to provide insight into the opportunities and limitations of using these designs in regulatory decisions[134]. Taken together, RWE is slowly being considered as a supplement to RCTs for regulatory approvals and large agencies are currently proceeding with caution.

1.3.3 Generalizability of Cardiovascular Outcome Trials

The CVOTs described above are prone to concerns of external validity as they all had specific inclusion and exclusion criteria to capture individuals at elevated risk of cardiovascular events. Although these study design decisions increased the event rate of the trial (and thus decreased the required number of participants to reach a given statistical power), they raise concern of whether the trial's results apply to real-world populations[135]. For example, Black individuals and those of lower socioeconomic status are disproportionately burdened by CVD and CKD and

are at a higher risk of T2DM[136]. However, these populations are consistently under-represented in trials[137]. The lack of data for these groups may influence trends in prescribing patterns, which show that these populations are less likely to receive SGLT-2 inhibitors or GLP-1 RA in both the US and Denmark despite potential benefits [136–138]. These prescribing patterns raise concerns of widening the future disparities in cardiovascular outcomes.

Meanwhile, the UK's National Institute for Health and Care Excellence (NICE) updated their T2DM guidance in 2022 to propose offering SGLT-2 inhibitor therapy to people with established ASCVD or HF, and considering SGLT-2 inhibitor therapy for those at risk of CVD defined as a 10-year cardiovascular risk of >10% using the QRISK2 algorithm (an algorithm used to determine the risk of CVD)[139,140]. A cross-sectional study evaluating the impact of the new guideline found that 93.1% of the UK population currently on antidiabetic treatment in the Clinical Practice Research Database (CPRD), a large primary care database, are now recommended or considered for SGLT-2 inhibitor therapy, with 60% eligible based on the QRISK2 criteria and 33% based on ASCVD or HF[141]. Under these new guidelines, the majority of those eligible for SGLT-2 inhibitors goes beyond the population evaluated in the SGLT2 inhibitor trials, where 57-85% of the population had ASCVD[142].

Given these concerns, more knowledge is required to understand the overlap between patient characteristics and outcomes of those participating in CVOT and patients with T2DM taking antidiabetic medication in a real-world setting. To my knowledge, no study has replicated CVOTs for long-acting insulin analogues in a real-world population of T2DM patients. In addition, no study has assessed the eligibility of the real-world population for long-acting insulin RCTs for T2DM. Thus, with this thesis, I aim to explore the generalizability of CVOTs for long-acting insulin analogues in a real-world population of individuals with T2DM.

1.4 Thesis work

1.4.1 Thesis Objectives

This thesis contains three primary objectives:

1. To synthesize the existing evidence regarding the successful replication of CVOTs of newer antidiabetic medications using RWD and to examine the proportion of patients seen in real-world settings that would have been eligible for these cardiovascular outcome trials.

2. To compare patient characteristics among patients with T2DM in a real-world population to the original DEVOTE trial.
3. To compare the estimated treatment effect of insulin degludec vs insulin glargine on CV outcomes in a real-world population of patients with T2DM and in subpopulations eligible and ineligible for the DEVOTE trial.

1.4.2 Thesis Overview

Chapter 2 consists of the first manuscript of the thesis which is a systematic review of existing observational studies that have assessed the generalizability of previous CVOT among patients with T2DM. Chapter 3 is a bridging chapter between the systematic review and my database study. Chapter 4 includes a detailed description of the methods used in the database study of the thesis. Chapter 5 is the second manuscript of the thesis, describing a retrospective cohort study aiming to replicate the DEVOTE trial to compare cardiovascular outcomes in patients with T2DM patients receiving insulin degludec to insulin glargine. Chapter 6 provides a discussion on the main findings and implications of the thesis as well as directions for future research. Chapter 7 concludes the thesis.

CHAPTER 2: SYSTEMATIC REVIEW

2.1 Preface to the systematic review

Given that cardiovascular outcome trials are a crucial component to evaluating the cardiovascular effects of newer antidiabetic medications for patients with T2DM, there is a need to assess their generalizability to real-world populations. A precursory search suggested limited existing literature in the area with substantial heterogeneity in drug classes and methodology. Due to the limited literature, we aimed to perform a systematic review rather than a meta-analysis to determine the generalizability of previous CVOTs for the newer antidiabetic medications of SGLT-2 inhibitors, GLP-1 RAs, and DPP-4 inhibitors in real-world populations of patients with T2DM. To our knowledge, this is the first systematic review to examine the external validity of studies that have replicated previous RCTs using RWD and to assess the proportion of real-world populations eligible for previous RCTs for patients with T2DM. The results of the systematic review are reported using best practices as described by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Synthesis without meta-analysis (SWiM).

2.2 Replicating cardiovascular outcome trials of medications used to type 2 diabetes using real-world data: A systematic review of observational studies

Wanning Wang BHSc[1,2], William Tang BSc[1,2], Michael Webster-Clark PhD[2,3], Oriana HY Yu MD MSc[2,3,4], Kristian B. Filion PhD[1,2,3]

[1] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

[2] Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Quebec, Canada

[3] Department of Medicine, McGill University, Montreal, Quebec, Canada

[4] Division of Endocrinology and Metabolism, Jewish General Hospital/McGill University, Montreal, Quebec, Canada

Correspondence to:
Kristian B. Filion, PhD
Associate Professor and William Dawson Scholar
Departments of Medicine and of Epidemiology, Biostatistics and Occupational Health
McGill University
3755 Cote Ste-Catherine Road, Suite H410.1
Montreal, Quebec, Canada H3T 1E2

Phone: (514) 340-8222 x 28394
Fax: (514) 340-7564
Email: kristian.filion@mcgill.ca

Keywords: Cardiovascular Outcome Trials, Generalizability, Real-world data, Systematic review

Conflict of interest disclosure: None

Sources of Support: Dr. Filion is supported by a William Dawson Scholar award from McGill University. Wanning Wang is supported by a McGill Internal Studentship Award and and a CANadian Consortium of Clinical Trial TRAINing platform (CANTRAIN) studentship; CANTRAIN is funded by the Canadian Institutes of Health Research.

Word count: 3541; **Abstract:** 239; **Tables:** 4; **Figures:** 1; **Supplementary tables:** 5; **Supplementary figures:**7

Twitter: @KB_Filion

2.2.1 Abstract

Background: Cardiovascular outcome trials (CVOTs) are mandated by the Food and Drug Administration to assess the cardiovascular safety of new antidiabetic medications that enter the market. However, they are often selective and may not generalize to real-world settings. Our study aimed to synthesize observational studies in the area to assess the generalizability of CVOTs to real-world settings.

Methods: We systematically reviewed observational studies that emulated previous CVOTs for dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide 1 (GLP-1) receptor agonists, and sodium glucose cotransporter-2 (SGLT-2) inhibitors among patients with T2DM. We searched the MEDLINE, EMBASE and Cochrane databases for observational studies that focused on trial emulation or cross-sectional studies that reported the proportion of real-world patients eligible for completed CVOTs. Two independent reviewers screened articles, extracted data, and assessed study concordance with RCT results.

Results: Nineteen studies were included in our systematic review, including four cohort studies that emulated previous RCTs and 15 cross-sectional studies that evaluated trial eligibility. Results between RCTs and RWD were concordant for all drug classes in finding non-inferiority. The median eligibility percentage ranged from 13% to 31% for SGLT-2 inhibitor trials and 12% to 43% for GLP-1 receptor agonist trials. No included studies evaluated trial eligibility for DPP-4 inhibitors.

Conclusions: These results suggest that, while RCTs and RWD are concordant in their estimates, the trials lack representativeness. More research is needed on the replication of CVOTs using RWD to understand how different methods may impact findings.

2.2.2 INTRODUCTION

Randomized controlled trials (RCTs) are the current gold standard for assessing drug efficacy and for evidence-based regulatory decision making. However, due to their strict inclusion and exclusion criteria, trial populations may differ substantially from the real-world population of interest[1]. RCTs are also resource and time intensive. With the adoption of electronic health records, large amounts of real-world data (RWD) on prescription drug use and health outcomes are becoming readily available. In situations where it may not be suitable to conduct RCTs such as rare diseases or in populations ineligible for RCTs (e.g., children, pregnant women, the elderly), RWD play an important role in the generation of evidence to evaluate the effectiveness and safety of medical interventions.

In 2008, the FDA issued recommendations regarding the conduct of cardiovascular outcome trials (CVOT) to establish that antidiabetic medications have acceptable cardiovascular risk profiles[2]. These recommendations were made in the wake of concerns over the increased risk of myocardial infarction (MI) of rosiglitazone, a thiazolidinedione, among patients with type 2 diabetes mellitus (T2DM)[3]. Since then, more than 13 CVOTs have been conducted on antidiabetic drugs, including dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide 1 receptor agonists (GLP-1 RAs), and sodium glucose cotransporter-2 (SGLT-2) inhibitors. Results from these trials demonstrated cardiovascular safety, with many showing superiority[4]. However, these trials were highly selective and may not be generalizable to patients seen in real-world settings. For example, many of the trials required the presence of high cardiovascular risk factor levels to increase the number of events to achieve greater statistical power.

With a growing need to better understand the complementary nature of RWD to RCT evidence as regulatory agencies increasing consider RWD in its decision-making[2,5,6], there is an urgent need to examine the generalizability of RCT data. This is particularly true for antidiabetic medications given the preference for large CVOTs with strict inclusion criteria for evidence-based decision-making. Our objective was therefore to synthesize the information in this area to understand what proportion of patients treated in real-world setting are eligible for participation in CVOTs and compare patient characteristics and estimated treatment effects using RWD and the respective CVOTs via systematic review of observational studies.

2.2.3 METHODS

This systematic review was conducted following guidelines described in the Cochrane Handbook[7] and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)[8] and Synthesis Without Meta-analysis (SWiM) guidelines[9]. The study protocol was registered in the Open Science Framework platform (10.17605/OSF.IO/G8ERW).

2.2.3.1 Search strategy

We searched MEDLINE, EMBASE, and the Cochrane databases from inception to July 31st, 2023, to identify studies that used RWD to examine the proportion of patients eligible for CVOTs that emulated these trials. Our predefined search strategy utilized Medical Subject Headings (MeSH) in OVID MEDLINE, EMTREE terms in EMBASE, and keywords in all databases. The search strategy was defined in consultation with a medical librarian and is described in detail in **Supplementary Tables 2.1-2.3**. Briefly, we searched using the following concepts: observational study, RCTs, major adverse cardiovascular outcomes (MACE), T2DM, emulation/replication/comparability. There were no restrictions on language or geographic location. We also conducted a hand search of references included in studies (backward search), studies that have referenced identified key studies (forward search) and previous reviews not captured by our initial database search. We also hand-searched the grey literature using Google Scholar (first 10 pages).

2.2.3.2 Inclusion and exclusion criteria

We included published, peer-reviewed observational studies (e.g., cohort or nested-case controls studies) that compared the risk of cardiovascular outcomes in real-world patients with T2DM aged 18+ years to that observed in CVOTs. To reduce bias, cross-sectional studies were only used to examine eligibility and patient characteristics; they were not included in the assessment of clinical outcomes. The interventions of interest were DPP-4 inhibitors, GLP-1 receptor agonists, and SGLT-2 inhibitors. A list of completed CVOTs and their drug class and molecule can be found in **Supplementary Table 2.4**. There were no restrictions on comparators; active comparators such as other antidiabetic medications, standard of care, or lifestyle interventions were eligible for inclusion.

We excluded RCTs, meta-analyses, case reports, case series, letters-to-the-editor, editorials, and commentaries. In addition, we excluded conference abstracts as they often present

preliminary results and typically do not undergo rigorous peer review. Full texts published in a language other than English were also excluded.

The primary outcome of interest was MACE, which included cardiovascular death, MI, and ischemic stroke. Unstable angina was also included if it was reported as part of the MACE definition. Secondary outcomes were the individual components of MACE, hospitalization for heart failure, and all-cause mortality. We examined the proportion of the study population eligible for a referenced CVOT as reported by the authors. We also compared the reported patient characteristics of the real-world population and the subpopulation eligible for the CVOTs.

2.2.3.3 Study selection

Citations generated from the electronic database search were exported to Covidence and duplicate citations were removed. Two independent reviewers (WW and WT) screened each study title and abstract for potential inclusion. Any study deemed potentially eligible by either reviewer proceeded to full-text review, where the full-text of each study was evaluated independently by both reviewers. Disagreements were resolved by consensus or if needed, with input from a third author (KBF).

2.2.3.4 Data extraction

We extracted data using a Microsoft Excel workbook that had been designed and pilot-tested using relevant studies from a prior search. The extraction was done by the same two independent reviewers (WW and WT), with disagreements resolved via consensus or if needed, with the aid of a third author (KBF).

2.2.3.5 Assessment of agreement between RCTs and RWE

To assess generalizability in the RWD studies that estimated treatment effects in addition to comparing RCT and real-world outcomes, we used agreement statistics as established by the DUPLICATE study to compare effect estimates of primary outcomes between the restricted RWD population and the referenced CVOT population[10,11]. *Full statistical significance agreement* was considered to have occurred when the RWD and RCT estimates and 95% confidence intervals (CIs) were on the same side of the null. *Partial significance agreement* was considered to have occurred when the prespecified noninferiority criteria was met, even if the RWD study indicated superiority[11]. It was categorized as yes, no, or partial. *Estimate agreement* was considered to have

occurred when the effect estimate of the RWD study fell within the 95% CI of the RCT effect estimate. *Standardized difference agreement* was defined by standardized differences $|Z| < 1.96$.

$$Z = \frac{\hat{\theta}_{RWE} - \hat{\theta}_{RCT}}{\sqrt{\hat{\sigma}^2_{RWE} + \hat{\sigma}^2_{RCT}}}$$ $\hat{\theta}$ are effect estimates and $\hat{\sigma}^2$ are variances

As per the FDA, all the major CVOTs were designed for non-inferiority with an upper CI limit of 1.3[12]. For our statistical significance agreement analysis, we first assessed whether the trial was able to demonstrate non-inferiority. We then evaluated whether the RWD study was able to replicate the non-inferiority finding within the same margin. If both studies demonstrated non-inferiority, then full statistical significance agreement was established. If the trial achieved non-inferiority but the RWD study achieved superiority, partial statistical significance agreement was established. In addition to non-inferiority, if superiority was established in the trial, we assessed if the RWD study also found superiority to achieve full statistical significance agreement. Finally, we captured the methods used to adjust the RWE population to emulate the RCTs (i.e., weighting, matching, restriction) to examine potential sources of heterogeneity.

2.2.3.6 Data synthesis

Due to the high level of clinical heterogeneity among studies and interventions, meta-analysis was not feasible. Thus, we followed SWiM reporting guidelines[9]. The main summary measure was the adjusted hazard ratio (HR), as RCTs estimated HRs to assess MACE outcomes. Studies were grouped by drug class and further stratified by RCT emulated. Data were synthesized based on the treatment groups and are presented in tabular format. Heterogeneity and certainty of evidence were assessed by examining the methods in which the study attempted to replicate the RCT. We captured the exposure definition (intention-to-treat vs on-treatment), active comparators, restriction of the population, follow-up duration, and methods used to reduce confounding.

We conducted several exploratory, post-hoc analyses to examine the potential association of patient characteristics and the proportion of the RWD population eligible for the trial. First, we used scatterplots to identify trends between patient characteristics to the percentage of the population that would have been eligible for their reference CVOT. Second, we calculated standardized mean differences of patient characteristics between the RWD populations and the trial population for select studies based on data completeness.

2.2.4 RESULTS

2.2.4.1 Search results

We identified 5,522 potentially relevant publications (**Figure 2.1)**. After removing duplicates, 2,073 articles underwent title and abstract review, and 41 underwent full-text review. A total of 19 studies were included in the systematic review[13–31].

2.2.4.2 Study characteristics

Table 2.1 summarizes the characteristics of the included studies stratified by drug class. There were four cohort studies[10,20,30,31] that examined ten unique populations for the outcome of interest. We identified 15 cross-sectional studies that examined the proportion eligible and patient characteristics[13,15–19,21–29]. All included studies were conducted between 2009 and 2020. Most studies were conducted in the US (9) or Europe (7), with three out of 19 studies conducted in Asia. A total of ten studies assessed GLP-1 RAs[10,15–17,20,22,24,26,29,30], 12 assessed SGLT-2 inhibitors[10,13,17,19–21,23,25–28,31], and two assessed DPP-2 inhibitors[10,20]. Within the exposure class, 13 studies evaluated multiple CVOTs (range: 1 to 7 CVOTs; median: 3 per class).

2.2.4.3 MACE Outcomes in Emulated RCTs

There were four studies that emulated RCTs using RWD (**Table 2.3**). Sciannameo et al.[20] used odds weighting based on RCT subgroup-specific HRs to transport results onto their RWD population; consequently, there was no real-world exposure or comparator for these analyses. As the exposures and outcomes in these analyses were directly from the relevant RCTs, we did not calculate the agreement statistics for this study. For the other three studies, exposure definitions of intention-to-treat (ITT) or on-treatment were used with an active comparator. These studies also imitated the inclusion criteria and exclusion criteria of the emulated RCTs. They also ensured follow-up was the same duration as the RCT and used propensity score matching to reduce confounding. MACE was defined as a composite endpoint that included cardiovascular or all-cause death, stroke and MI; Franklin et al. emulated TECOS by including angina in their MACE definition[10].

For studies that investigated SGLT-2 inhibitors, there were three RCTs examined: EMPA-REG Outcome[32], CANVAS[33] and DECLARE[34]. Treatment matched the SGLT-2 inhibitors evaluated in the RCTs. However, while the RCTs used a placebo comparator, active comparators (DPP-4 inhibitors) were used in the RWD studies. There was strong agreement between RWD and

RCTs, with the three studies achieving full statistical significance agreement and estimate agreement, with HRs from the RWD in the same direction as those from the RCTs and within the RCT's 95% CIs.

There were three analyses of GLP-1 RAs, which examined six RCTs (EXSCEL[35], LEADER[36], PIONEER-6[37], REWIND[38], SUSTAIN-6[39]). The RWD studies used ITT and on-treatment exposure definitions. For the emulation of the LEADER trial, liraglutide was compared with sulfonylureas, second-to-third-line antidiabetic drugs, and DPP-4 inhibitors. Across the three different analyses, all showed partial statistical significance agreement. All studies except Abrahami et al. evaluating liraglutide compared to sulfonylureas demonstrated estimate agreement[30].

There were five analyses of DPP-4 inhibitors, which examined three RCTs (CARMELINA[40], SAVOR-TIMI-53[41], and TECOS[42]). RCT odds weighting was used in two of the five analyses. In the other three analyses, an on-treatment exposure definition was used, and second-generation sulfonylureas were used as the comparator. All analyses were able to replicate the non-inferiority findings of the RCTs, and all but one demonstrated estimate agreement.

2.2.4.4 Secondary Clinical Outcomes in Emulated RCTs

Table 2.4 summarizes the results for secondary clinical outcomes in the RWD studies and the emulated RCTs. Jang et al, which emulated the EMPA-REG Outcome trial by comparing empagliflozin with sitagliptin, examined select individual components of MACE (MI and stroke), all-cause mortality, and hospitalization due to heart failure (HHF)[31]. All-cause mortality, MI, and HHF had either partial or full statistical significance agreement and estimate agreement. Empagliflozin was found to be non-inferior for stroke in RWD but not the RCT; thus, statistical significance agreement was not met for this outcome. However, there was estimate agreement[31]. Sciannameo et al. and Franklin et al. examined HHF and/or cardiovascular death when replicating the DECLARE trial[14,20]. Franklin et al. found both regulatory agreement for superiority and estimate agreement.

2.2.4.5 Eligibility

Table 2.4 summarizes the findings when examining the proportion of RWD population eligible (including cross-sectional studies) for CVOTs. In total, there were 59 populations across CVOTs evaluating SGLT-2 inhibitors and GLP-1 RAs. Populations varied in composition, with

some from the general population, others only from select populations (inpatient vs outpatient), and others restricted to users of the drug class of interest.

For SGLT-2 inhibitors, RWD populations were compared to a total of four RCTs (CANVAS[33], DECLARE[34], EMPA-REG Outcome,[32] and VERTIS-CV[43]). The median percentage eligible from the trials ranged between 12.6% to 30.7%. Of note, the Jang et al. study that demonstrated full statistical, estimate, and standardized difference agreement with the EMPA-REG Outcome trial had an 12.6% eligible for the RCT from a population of patients with T2DM who were new-users of empagliflozin or sitagliptin. For GLP-1 RAs, RWD populations were compared to eight different RCT (ELIXA[44], EXSCEL[35], FREEDOM-CVO[45], HARMONY[46], LEADER[36], PIONEER-6[37], REWIND[38], SUSTAIN-6[39]). Median percentage eligible from the trials ranged from 11.8% to 42.6%. No included studies evaluated trial eligibility for DPP-4 inhibitors.

We conducted post-hoc analyses for patient characteristics which can be found in **Supplementary Figures 2.1-2.7**. There were no association between age or sex in RWD populations and percentage of individuals who were eligible for the RCTs. Moreover, there were substantial differences in patient characteristics such as age, sex, use of antidiabetic medication between the RWD populations compared to the populations examined in the respective RCTs.

2.2.5 DISCUSSION

Our study was designed to synthesize observational studies that emulated CVOTs in patients with T2DM and summarized cross-sectional studies to assess the proportion of real-world populations eligible for previous CVOTs. We found there was relatively strong agreement between RWD and RCTs for SGLT-2 inhibitors, GLP-1 RAs, and DPP-4 inhibitors. Individual components of MACE, HHF, and cardiovascular death also showed consistency between the RWD studies[14,20,31] and the SGLT-2 inhibitor CVOTs EMPA-REG Outcome[32] and DECLARE[34]. There was considerable heterogeneity in the percentage of individuals who would have been eligible for the CVOTs of SGLT-2 inhibitors and GLP-1 RAs. Patient characteristics also varied between RWD and RCTs, with no association between patient characteristics and real-world proportion eligible for RCTs.

Our systematic review highlighted the heterogeneity of methods used to emulate RCTs using RWD. As this is an emerging area of research, the studies included in our study had different exposure definitions, active comparators, and methods to reduce confounding. The lack of

agreement between Abrahami et al's study and the LEADER trial highlights how differences in comparators from sulfonylureas to broader 2nd or 3rd line antidiabetic drugs influenced the direction of the effect estimate[30]. In addition, some studies utilized an ITT approach where exposure was defined at cohort entry until end of follow-up while others used an on-treatment approach which censored for treatment discontinuation or switching. As these changes may impact the research question and effect estimate, more research is needed to better understand the strengths and limitations of using different methods for replicating RCTs with RWD.

Our study has shown that studies emulating RCTs using RWD to estimate the effects of newer antidiabetic medications on MACE generally demonstrated agreement. However, there was some heterogeneity in our results, which may be due to the types of methods used for replication such as choice of exposure definition (ITT vs on-treatment), choice of active comparator, and methods to adjust for confounding. In addition, there were some differences between RCTs and RWD which may be due to inherent differences between the two designs, such as the source population, exposure (active comparator vs placebo), follow-up duration, and how events are captured and measured. The results may have also differed due to chance. Overall, the studies included in our systematic review demonstrated that similar conclusions can be obtained when using RWD to emulate RCTs when efforts are made in study design to closely mimic RCTs. While RCTs are the gold-standard for the generation of evidence, the regulatory bodies of several regions such as Canada, USA and the European Union are evaluating the use of RWD in decision-making[2,6,47]. Our results provide reassurance that with the proper methods, RWD can similarly generate quality evidence.

The proportion of the real-world population that would have been eligible for the RCTs varied across studies but was almost always less than 50%. The lack of generalizability of these RCT populations is driven in part by their design. These trials predominantly included older individuals who possessed one or more cardiovascular disease (CVD) risk factors and generally had longer durations of disease. These factors were necessary for event driven CVOTs to ensure that that a sufficient number of events could be accrued in a short time frame[4]. This requirement explains in part the observed heterogeneity (and large standardized mean differences for patient characteristics) between the RWD and the RCT populations. For example, the RWD populations consistently had lower percentages of established CVD. Given these differences, it is important to

utilize studies from routine practice settings to complement findings from highly controlled RCTs to account for differences in patients and practice for clinical decision making[4]. There are opportunities to use RWD for Phase IV trials and pragmatic trials as highlighted by the FDA[2].

Previous reviews in the area of RCT emulation using RWD examined the reporting of RWD studies that aimed to replicate RCTs[48,49]. They found inconsistencies in the reporting of key elements in these studies. For example, in a systematic review of 200 studies aiming to emulate RCTs using observational data, Hansford et al. reported that 43% did not describe all key elements of how the target trial was emulated[48]. In addition, only 37% of studies reported potential unmeasured confounders. A scoping review of 96 studies aiming to emulate a target trial across medical fields also identified unmeasured confounding as the most commonly stated limitation[49]. We found that the cohort studies included in our systematic review all reported the key elements of eligibility criteria, treatment strategy, assignment procedures, outcome, follow-up, causal contrast, analyses plan and specification of time-zero. However, they did not comment on potential unmeasured confounders. Our study builds upon these previous studies by examining agreement statistics between RWD and RCTs, the proportion of RWD patients who would have been eligible for RCTs, and differences in patient characteristics. The recent availability of guidelines for the reporting of target trial emulation may improve reporting in this area[50].

Our study has many strengths. It included a comprehensive, systematic search that was constructed with the assistance of an experienced librarian. To our knowledge, this is the first systematic review to examine the emulation of CVOTs for newer antidiabetic drugs using RWD. In addition, we assessed the agreement between the RWD studies and RCTs based on regulatory standards. Furthermore, our study is the first to synthesize the generalizability of CVOTs in the real-world by examining the proportion eligible of real-world populations for these CVOTs. We also compared patient characteristics between the CVOTs and RWD to better understand differences in study populations.

Our study has several limitations. First, there was a limited number of studies emulating RCTs using RWD among patients with T2DM. These studies also had important heterogeneity in study design, drug class, and analytical approach. Consequently, we conducted a systematic review without formal meta-analysis. Second, as an emerging area of research, there are currently no established guidelines on how to assess the quality of these studies for external validity. While we

used the DUPLICATE Team's agreement statistics[11,14] to assess the estimates achieved from the emulation in comparison to the RCT that was being emulated, these statistics have their own limitations. For example, it is difficult to replicate studies that have shown a null effect, as observational studies often have more precision. This issue is well illustrated by our results for DPP-4 inhibitors where RWD studies achieved superiority while the RCTs reported a null effect[14,20]. Moreover, the estimate agreement is contingent on the precision of the RCT. If the RCT has a wider CI, there is a higher probability that the RWD estimate will fall within this interval. Third, to our knowledge, there are no guidelines or standards on reporting or assessing external validity of RCT. Finally, we chose to include cross-sectional studies for our secondary objectives examining the proportion eligible of real-world populations for RCTs and their characteristics, however, these studies had considerable missing data and are at an inherently higher risk of bias.

2.2.6 CONCLUSION

In this systematic review of studies using real-world populations to emulate CVOTs, we found that all studies included were able to replicate either superiority or non-inferiority findings for MACE with SGLT-2 inhibitors, GLP-1 receptor agonists, and DPP-4 inhibitors among patients with T2DM. The proportion of real-world patients eligible for the CVOTs was small and had high degrees of variability across studies. These results complement existing RCTs and RWDs to assist with regulatory decision-making for cardiovascular safety in antidiabetic drugs and highlight the lack of generalizability of current CVOTs.

2.2.7 ACKNOWLEDGMENTS

WW received funding from a Faculty of Medicine Internal Studentship and a CANadian Consortium of Clinical Trial TRAINing platform (CANTRAIN) studentship; CANTRAIN is funded by the Canadian Institutes of Health Research. KBF holds a William Dawson Scholar award from McGill University.

2.2.8 DISCLOSURES

None to declare.

2.2.9 AUTHORS' CONTRIBUTIONS

WW led the protocol development, conducted the analysis, and drafted the manuscript. WW and WT extracted data. All authors contributed to protocol development, were involved in data interpretation, critically reviewed the manuscript for important intellectual content, and approved the final manuscript. KBF supervised the study and is the guarantor.

2.2.10 REFERENCES

1. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-World Evidence — What Is It and What Can It Tell Us? N Engl J Med. 2016 Dec 8;375(23):2293–7.

2. Rick Turner J. FDA's New Guidance for Industry Addressing Evaluation of the Safety of New Drugs for Improving Glycemic Control: A Case Study in Regulatory Science. Ther Innov Regul Sci. 2021 Jan 1;55(1):1–5.

3. Nissen SE, Wolski K. Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes. New England Journal of Medicine. 2007 Jun 14;356(24):2457–71.

4. Eckel RH, Farooki A, Henry RR, Koch GG, Leiter LA. Cardiovascular Outcome Trials in Type 2 Diabetes: What Do They Mean for Clinical Practice? Clin Diabetes. 2019 Oct;37(4):316–37.

5. CADTH. Guidance for Reporting Real-World Evidence [Internet]. Canada: Canada's Drug and Health Technology Agency; 2023 May p. 66. (CADTH Methods and Guidelines). Available from: https://www.cadth.ca/sites/default/files/RWE/MG0020/MG0020-RWE-Guidance-Report.pdf

6. Canada H. Regulatory review of drugs and devices: Strengthening the use of real world evidence for drugs [Internet]. 2018 [cited 2023 Jun 19]. Available from: https://www.canada.ca/en/health-canada/corporate/transparency/regulatory-transparency-and-openness/improving-review-drugs-devices/strengthening-use-real-world-evidence-drugs.html

7. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. [cited 2023 Oct 19]. Available from: https://training.cochrane.org/handbook/current

8. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;372:n71.

9. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020 Jan 16;368:l6890.

10. Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, et al. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. Clinical Pharmacology & Therapeutics. 2020;107(4):817–26.

11. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376–85.

12. Center for Drug Evaluation and Research. Type 2 Diabetes Mellitus: Evaluating the Safety of New Drugs for Improving Glycemic Control Guidance for Industry [Internet]. FDA; 2020 [cited 2023 Sep 8]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/type-2-diabetes-mellitus-evaluating-safety-new-drugs-improving-glycemic-control-guidance-industry

13. McGovern A, Feher M, Munro N, de Lusignan S. Sodium-Glucose Co-transporter 2 (SGLT2) Inhibitor: Comparing Trial Data and Real-World Use. Diabetes Ther. 2017 Apr;8(2):365–76.

14. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. Circulation. 2021 Mar 9;143(10):1002–13.

15. Romera I, Artime E, Ihle K, Díaz-Cerezo S, Rubio de-Santos M, De Prado A, et al. A Retrospective Observational Study Examining the Generalizability of Glucagon-Like Peptide 1 Receptor Agonist Cardiovascular Outcome Trials to the Real-World Population with Type 2 Diabetes in Spain: The REPRESENT Study. Adv Ther. 2022 Aug;39(8):3589–601.

16. Webb J, Mount J, von Arx LB, Rachman J, Spanopoulos D, Wood R, et al. Cardiovascular risk profiles: A cross-sectional study evaluating the generalizability of the glucagon-like peptide-1 receptor agonist cardiovascular outcome trials REWIND, LEADER and SUSTAIN-6 to the real-world type 2 diabetes population in the United Kingdom. Diabetes Obes Metab. 2022 Feb;24(2):289–95.

17. Wittbrodt E, Chamberlain D, Arnold SV, Tang F, Kosiborod M. Eligibility of patients with type 2 diabetes for sodium-glucose co-transporter-2 inhibitor cardiovascular outcomes trials: An assessment using the Diabetes Collaborative Registry. Diabetes, Obesity and Metabolism. 2019 Aug;21(8):1985–9.

18. Eric T. Wittbrodt P, James M. Eudicone MS, Kelly F. Bell P, Devin M. Enhoffer P, Keith Latham P, Jennifer B. Green MD. Generalizability of Glucagon-Like Peptide-1 Receptor Agonist Cardiovascular Outcome Trials Enrollment Criteria to the US Type 2 Diabetes Population. 2018 Apr 13 [cited 2023 Oct 19];24. Available from: https://www.ajmc.com/view/generalizability-of-glucagon-peptide1-agonist-cv-outcome-trials

19. Pintat S, Fenici P, Hammar N, Ji L, Khunti K, Medina J, et al. Eligibility of patients with type 2 diabetes for sodium–glucose cotransporter 2 inhibitor cardiovascular outcomes trials: a global perspective from the DISCOVER study. BMJ Open Diab Res Care. 2019 Mar;7(1):e000627.

20. Sciannameo V, Berchialla P, Avogaro A, Fadini GP, Consoli A, Formoso G, et al. Transposition of cardiovascular outcome trial effects to the real-world population of patients with type 2 diabetes. Cardiovascular Diabetology. 2021 May 10;20(1):103.

21. Nicolucci A, Candido R, Cucinotta D, Graziano G, Rocca A, Rossi MC, et al. Generalizability of Cardiovascular Safety Trials on SGLT2 Inhibitors to the Real World: Implications for Clinical Practice. Adv Ther. 2019 Oct;36(10):2895–909.

22. Boye KS, Riddle MC, Gerstein HC, Mody R, Garcia-Perez LE, Karanikas CA, et al. Generalizability of glucagon-like peptide-1 receptor agonist cardiovascular outcome trials to the overall type 2 diabetes population in the United States. Diabetes, Obesity and Metabolism. 2019 Jun;21(6):1299–304.

23. Birkeland KI, Bodegard J, Norhammar A, Kuiper JG, Georgiado E, Beekman-Hendriks WL, et al. How representative of a general type 2 diabetes population are patients included in cardiovascular outcome trials with SGLT2 inhibitors? A large European observational study. Diabetes Obes Metab. 2019 Apr;21(4):968–74.

24. Fan W, Tong C, Wong ND. LEADER Trial Eligibility and Preventable Cardiovascular Events in US Adults with Diabetes: the National Health and Nutrition Examination Surveys 2007–2016. Cardiovasc Drugs Ther. 2020 Dec;34(6):737–43.

25. Zhou YL, Zhang YG, Zhang R, Zhou YL, Li N, Wang MY, et al. Population diversity of cardiovascular outcome trials and real-world patients with diabetes in a Chinese tertiary hospital. Chinese Medical Journal. 2021 Jun 5;134(11):1317–23.

26. Arnold SV, Inzucchi SE, Tang F, McGuire DK, Mehta SN, Maddox TM, et al. Real-world use and modeled impact of glucose-lowering therapies evaluated in recent cardiovascular outcomes trials: An NCDR® Research to Practice project. Eur J Prev Cardiolog. 2017 Oct;24(15):1637–45.

27. Shao SC, Lin YH, Chang KC, Chan YY, Hung MJ, Yang YHK, et al. Sodium glucose co-transporter 2 inhibitors and cardiovascular event protections: how applicable are clinical trials and observational studies to real-world patients? BMJ Open Diabetes Research and Care. 2019 Dec 1;7(1):e000742.

28. Hinton W, Feher M, Munro N, Joy M, de Lusignan S. Sodium–glucose co-transporter-2 inhibitor cardiovascular outcome trials and generalizability to English primary care. Diabet Med. 2020 Sep;37(9):1499–508.

29. Sciannameo V, Berchialla P, Orsi E, Lamacchia O, Morano S, Querci F, et al. Enrolment criteria for diabetes cardiovascular outcome trials do not inform on generalizability to clinical practice: The case of glucagon-like peptide-1 receptor agonists. Diabetes Obes Metab. 2020 May;22(5):817–27.

30. Abrahami D., Pradhan R., Yin H., Honig P., Baumfeld Andre E., Azoulay L. Use of Real-World Data to Emulate a Clinical Trial and Support Regulatory Decision Making: Assessing the Impact of Temporality, Comparator Choice, and Method of Adjustment. Clin Pharmacol Ther. 2021;109(2):452–61.

31. Jang HY, Kim IW, Oh JM. Using real-world data for supporting regulatory decision making: Comparison of cardiovascular and safety outcomes of an empagliflozin

randomized clinical trial versus real-world data. Frontiers in Pharmacology [Internet]. 2022 [cited 2023 Jul 25];13. Available from: https://www.frontiersin.org/articles/10.3389/fphar.2022.928121

32. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. N Engl J Med. 2015 Nov 26;373(22):2117–28.

33. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. N Engl J Med. 2017 Aug 17;377(7):644–57.

34. Furtado RHM, Bonaca MP, Raz I, Zelniker TA, Mosenzon O, Cahn A, et al. Dapagliflozin and Cardiovascular Outcomes in Patients With Type 2 Diabetes Mellitus and Previous Myocardial Infarction. Circulation. 2019 May 28;139(22):2516–27.

35. Drucker DJ, Buse JB, Taylor K, Kendall DM, Trautmann M, Zhuang D, et al. Exenatide once weekly versus twice daily for the treatment of type 2 diabetes: a randomised, open-label, non-inferiority study. The Lancet. 2008 Oct 4;372(9645):1240–50.

36. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. N Engl J Med. 2016 Jul 28;375(4):311–22.

37. Husain M, Birkenfeld AL, Donsmark M, Dungan K, Eliaschewitz FG, Franco DR, et al. Oral Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. New England Journal of Medicine. 2019 Aug 29;381(9):841–51.

38. Gerstein HC, Colhoun HM, Dagenais GR, Diaz R, Lakshmanan M, Pais P, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. The Lancet. 2019 Jul 13;394(10193):121–30.

39. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. N Engl J Med. 2016 Nov 10;375(19):1834–44.

40. Rosenstock J, Perkovic V, Johansen OE, Cooper ME, Kahn SE, Marx N, et al. Effect of Linagliptin vs Placebo on Major Cardiovascular Events in Adults With Type 2 Diabetes and High Cardiovascular and Renal Risk. JAMA. 2019 Jan 1;321(1):69–79.

41. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B, et al. Saxagliptin and Cardiovascular Outcomes in Patients with Type 2 Diabetes Mellitus. New England Journal of Medicine. 2013 Oct 3;369(14):1317–26.

42. Green JB, Bethel MA, Armstrong PW, Buse JB, Engel SS, Garg J, et al. Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes. New England Journal of Medicine. 2015 Jul 16;373(3):232–42.

43. Cannon CP, Pratley R, Dagogo-Jack S, Mancuso J, Huyck S, Masiukiewicz U, et al. Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. New England Journal of Medicine. 2020 Oct 8;383(15):1425–35.

44. Pfeffer MA, Claggett B, Diaz R, Dickstein K, Gerstein HC, Køber LV, et al. Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome. New England Journal of Medicine. 2015 Dec 3;373(23):2247–57.

45. Ruff CT, Baron M, Im K, O'Donoghue ML, Fiedorek FT, Sabatine MS. Subcutaneous infusion of exenatide and cardiovascular outcomes in type 2 diabetes: a non-inferiority randomized controlled trial. Nat Med. 2022 Jan;28(1):89–95.

46. Hernandez AF, Green JB, Janmohamed S, D'Agostino RB, Granger CB, Jones NP, et al. Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. The Lancet. 2018 Oct 27;392(10157):1519–29.

47. Use of real-world evidence in regulatory decision making – EMA publishes review of its studies | European Medicines Agency [Internet]. [cited 2024 May 30]. Available from: https://www.ema.europa.eu/en/news/use-real-world-evidence-regulatory-decision-making-ema-publishes-review-its-studies

48. Hansford HJ, Cashin AG, Jones MD, Swanson SA, Islam N, Douglas SRG, et al. Reporting of Observational Studies Explicitly Aiming to Emulate Randomized Trials: A Systematic Review. JAMA Network Open. 2023 Sep 27;6(9):e2336023.

49. Zuo H, Yu L, Campbell SM, Yamamoto SS, Yuan Y. The implementation of target trial emulation for causal inference: a scoping review. Journal of Clinical Epidemiology. 2023 Oct 1;162:29–37.

50. Hansford HJ, Cashin AG, Jones MD, Swanson SA, Islam N, Dahabreh IJ, et al. Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. BMJ Open. 2023 Sep 12;13(9):e074626.

51. Barakat S., Habib H.A., Chun C., Oh E. Linagliptin vs. Sitagliptin for Type 2 Diabetes Inpatient Glycemic Management and Cardiovascular Outcomes-Is There a Difference between DPP-4 Inhibitors? Diabetes [Internet]. 2022;71(Supplement 1). Available from: https://diabetesjournals.org/diabetes/article/71/Supplement_1/130-LB/146731/130-LB-Linagliptin-vs-Sitagliptin-for-Type-2

52. Sheu WHH, Gantz I, Chen M, Suryawanshi S, Mirza A, Goldstein BJ, et al. Safety and Efficacy of Omarigliptin (MK-3102), a Novel Once-Weekly DPP-4 Inhibitor for the Treatment of Patients With Type 2 Diabetes. Diabetes Care. 2015 Nov;38(11):2106–14.

Table 2.1 Summary of study characteristics for systematic review of observational studies replicating cardiovascular outcome trials among patients with type 2 diabetes

Study	Study design	Data origin	Study Period	Exposure class	RCT referenced	Number of particpants in study	Patient population
Abrahami, 2020[30]	Cohort Study	UK	2009-2013	GLP-1 Receptor Agonists	LEADER[36]	63 297[a]	Clinical Practice Research Datalink (CPRD)
Franklin, 2020[14]	Cohort study	USA	2004-2019	GLP-1 Receptor Agonists	LEADER[36]	168 690	US Healthcare Claims: OPTUM, IBM, Medicare
Sciannameo, 2021[20]	Reanalyses of RCT[b]	Italy	2015-2016	GLP-1 Receptor Agonist	PIONEER 6[37], REWIND[38], SUSTAIN-6[39], EXSCEL[35], LEADER[36]	139 700	DARWIN- T2DM database
Franklin, 2020[14]	Cohort study	USA	2004-2019	SGLT-2 inhibitors	DECLARE[34], EMPA-REG[32], CANVAS[33]	305 744[c]	US Healthcare Claims: OPTUM, IBM, Medicare
Jang, 2022[31]	Cohort Study	South Korea	2011-2020	SGLT-2 inhibitors	EMPA-REG[32]	23 126	Korean Health Insurance Review and Assessment Service
Sciannameo, 2021[20]	Reanalyses of RCT[b]	Italy	2015-2016	SGLT-2 inhibitors	DECLARE[34], EMPA-REG[32]	139 700	DARWIN- T2DM database
Franklin, 2020[14]	Cohort study	USA	2004-2019	DPP-4 inhibitors	CARMELINA[40], TECOS[42], SAVOR-TIMI[41]	633 432[d]	US Healthcare Claims: OPTUM, IBM, Medicare
Sciannameo, 2021[20]	Reanalyses of RCT[b]	Italy	2015-2016	DPP-4 Inhibitors	SAVOR-TIMI[41], TECOS[42]	139 700	DARWIN- T2DM database
Arnold, 2017[26]	Cross-sectional	USA	2017	GLP-1 Receptor Agonists	LEADER[36]	87 601	Diabetes Collective Registry (DCR)
Boye, 2018[22]	Cross-sectional	USA	2012-2017	GLP-1 Receptor Agonists	EXSCEL[35], LEADER[36], REWIND[38], SUSTAIN-6[39]	26 110 573	IQIVIA Real World Data Adjudicated Claims linked with EMR and US National Health and Nutrition Examination Surveys (NHANES)
Fan, 2020[24]	Cross-sectional	USA	2007- 2016	GLP-1 Receptor Agonists	LEADER[36]	29 629	NHANES
Romera, 2022[15]	Cross-sectional	Spain	2013- 2019	GLP-1 Receptor Agonists	LEADER[36], REWIND[38], SUSTAIN-6[39]	24 268	IQIVIA EMR database in Spain
Sciannameo, 2020[20]	Cross-sectional	Italy	2015- 2016	GLP-1 Receptor Agonists	LEADER[36], SUSTAIN-6[39], EXSCEL[35], REWIND[38], PIONEER-6[37], HARMONY[46]	130 380	DARWIN-T2DM database

Webb, 2021[16]	Cross-sectional	UK	2018	GLP-1 Receptor Agonists	REWIND[38], LEADER[36], SUSTAIN-6[39]	33 118	CPRD GOLD
Wittbrodt, 2018[17]	Cross-sectional	USA	2009-2012	GLP-1 Receptor Agonist	EXSCEL[35], REWIND[38], FREEDOM-CVO[45], LEADER[36], SUSTAIN-6[39], HARMONY[46], ELIXA[44]	20 142	NHANES
Arnold, 2017[26]	Cross-sectional	USA	2017	SGLT-2 inhibitors	EMPA-REG[32]	47 872	Diabetes Collective Registry (DCR)
Birkeland, 2018[23]	Cross-sectional	Norway, Sweden, Germany, Netherlands	2014-2015	SGLT-2 inhibitors	CANVAS[33], EMPA-REG[32], VERTIS-CV[43]	803 836	Germany: Health Insurance funds database Netherlands: Electronic health database Norway and Sweden: National Public Health System
Hinton, 2020[28]	Cross-sectional	England	2016	SGLT-2 inhibitors	CANVAS[33], DECLARE[34], EMPA-REG[32], VERTIS-CV[43]	1 595 445	Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC)
McGovern, 2017[13]	Cross-sectional	England	2016	SGLT-2 inhibitors	EMPA-REG[32]	60 327	RCGP RSC
Nicolucci, 2019[21]	Cross-sectional	Italy	2016	SLGT-2 inhibitors	EMPA-REG[32], CANVAS[33], DECLARE[34], VERTIS-CV[43]	455 622	Italian Association of Diabetologists (AMD) database
Pintat, 2019[19]	Cross-sectional	38 Countries in Discover study	2017[e]	SGLT-2 inhibitors	CANVAS[33], DECLARE[34], EMPA-REG[32] VERTIS- CV[43]	11 385	DISCOVER prospective observational study
Shao, 2019[27]	Cross-sectional	Taiwan	2018-2019	SGLT-2 inhibitors	CANVAS[33], DECLARE[34], EMPA-REG[32], VERTIS CV[43]	11 650	Chang Gung Research Database
Wittbrodt, 2019[17]	Cross-sectional	USA	2013- 2016	SGLT-2 inhibitors	EMPA-REG[32], CANVAS[33], DECLARE[34], VERTIS-CV[43]	172 643	DCR
Zhou, 2020[25]	Cross-sectional	China	2011-2019	SGLT-2 inhibitors	EMPA-REG[32]	214 963	EMR from West China Hospital

T2DM: Type 2 Diabetes Mellitus; EMR: Electronic Medical Records; CPRD: Clinical Practice Research Datalink; DCR: Diabetes Collective Registry; NHANES: National Health and Nutrition Examination Surveys; RCGP: Royal College of General Practitioners; RSC: Research and Surveillance Centre; AMD: Italian Association of Diabetologists; RCT: Randomized Controlled Trial; [a]Total including liraglutide, sulfonylureas and second-to-third line antidiabetic medications populations; [b]Weighting RCT stratum-specific treatment effects according to proportions of a given characteristics in the target population; [c]Total from DECLARE, EMPA-REG Outcome and CANVAS RWE populations; [d]Total from CAREMLINA, TECOS and SAVOR-TIMI RWE population ; [e]Taken from DISCOVER stud

Table 2.2 Summary of MACE outcomes for RWD and RCT studies grouped by exposure class

Study	RCT	Exposure definition	Exposure	Comparator	Exposure N	Comparator N	HR	95% CI	RCT HR	95% CI	Observed Agreement[a]		
											SA	EA	SD
SGLT-2 inhibitors													
Jang, 2022	EMPA-REG[32]	ITT	Empagliflozin	Sitagliptin	11 563	11 563	0.87	0.79, 0.96	0.86	0.74, 0.99	Y	Y	Y
Franklin, 2020	EMPA-REG[32]	ITT	Empagliflozin	DPP-4i	4 687	2 333	0.83	0.73, 0.94	0.86	0.74, 0.99	Y	Y	Y
Franklin, 2020	CANVAS[33]	On treatment	Canagliflozin	DPP-4i	5 795	4 347	0.77	0.70, 0.85	0.86	0.75, 0.97	Y	Y	Y
Sciannameo, 2021[b]	DECLARE[34]	N/A	N/A	N/A	N/A	N/A	0.94	0.84, 1.04	0.93	0.84, 1.03	-	-	-
Sciannameo, 2021[b]	EMPA-REG[32]	N/A	N/A	N/A	N/A	N/A	0.88	0.74, 1.03	0.86	0.74, 0.99	-	-	-
GLP-1 Receptor Agonists													
Abrahami, 2020	LEADER[36]	ITT	Liraglutide	SU	1 868	25 895	1.03	0.82, 1.30	0.87	0.78, 0.97	N	N	Y[d]
Abrahami, 2020	LEADER[36]	ITT	Liraglutide	2nd – 3rd line[c]	1 864	32 899	0.97	0.78, 1.22	0.87	0.78, 0.97	N	Y	Y[d]
Franklin, 2020	LEADER[36]	On treatment	Liraglutide	DPP-4i	4 668	4 672	0.82	0.76, 0.87	0.87	0.78, 0.97	Y	Y	Y
Sciannameo, 2021[b]	EXSCEL[35]	N/A	N/A	N/A	N/A	N/A	0.92	0.82, 1.02	0.91	0.83, 1.00	-	-	-
Sciannameo, 2021[b]	LEADER[36]	N/A	N/A	N/A	N/A	N/A	0.88	0.77, 0.99	0.87	0.78, 0.97	-	-	-
Sciannameo, 2021[b]	PIONEER-6[37]	N/A	N/A	N/A	N/A	N/A	0.76	0.41, 1.10	0.79	0.57, 1.11	-	-	-
Sciannameo, 2021[b]	REWIND[38]	N/A	N/A	N/A	N/A	N/A	0.87	0.76, 0.98	0.88	0.79, 0.99	-	-	-
Sciannameo, 2021[b]	SUSTAIN-6[39]	N/A	N/A	N/A	N/A	N/A	0.73	0.47, 0.99	0.74	0.58, 0.95	-	-	-
DPP-4 Inhibitors													
Franklin, 2020	CARMELIN[40]	On treatment	Linagliptin	2nd – gen SU	3 494	3 485	0.90	0.84, 0.96	1.02	0.89, 1.17	P	Y	Y
Franklin, 2020	SAVOR-TIMI[41]	On treatment	Saxagliptin	2nd – gen SU	8 280	8 212	0.81	0.76, 0.86	1.00	0.89, 1.12	P	N	Y
Franklin, 2020[e]	TECOS[42]	On treatment	Sitagliptin	2nd – gen SU	7 257	7 266	0.89	0.86, 0.91	0.98	0.88, 1.09	P	Y	Y
Sciannameo, 2021[b]	SAVOR-TIMI[41]	N/A	N/A	N/A	N/A	N/A	0.99	0.87, 1.10	1.00	0.89, 1.12	-	-	-
Sciannameo, 2021[b]	TECOS[42]	N/A	N/A	N/A	N/A	N/A	0.97	0.87, 1.06	0.98	0.88, 1.09	-	-	-

Abbreviations: HR: Hazard Ratio; CI: Confidence Interval; ITT: intention-to-treat, DPP-4i: Dipeptidyl peptidase 4 inhibitors; SU: sulfonylureas; 2nd – gen SU: Second generation sulfonylureas; RCT: Randomized Controlled Trial;

[a]SA: Full statistical significance agreement (Y/N/P) – determines if the RWD and RCT have estimates and CIs on the same side of the null; P: partial significance agreement- met the prespecified noninferiority criteria even though the database study may have indicated superiority, EA: Estimate agreement (Y/N) – determines if the effect estimate of the RWD falls within the 95% confidence interval of the RCT effect estimate. SD: Standardized difference agreement is defined by standardized differences $|Z| < 1.96$ (Y = yes, N = no).

[b] Weighting RCT stratum-specific treatment effects according to proportions of a given characteristics in the target population; [c]second to third line antidiabetic drugs; [d] Calculated by hand; [e]MACE including angina;

Table 2.3 Summary of secondary outcomes of all-cause mortality, MI, Stroke, HHF and CV death for RWD and RCT studies for SGLT-2 Inhibitors

Study	RCT	Exposure definition	Exposure	Comparator	Exposure N	Comparator N	Outcome	HR	95% CI	RCT HR	RCT 95% CI	Observed Agreement [a]		
												SA	EA	SD
Jang, 2022	EMPA-REG[32]	ITT	Empagliflozin	Sitagliptin	11 563	11 563	All cause death	0.78	0.67, 0.91	0.68	0.57, 0.82	Y	Y	Y
Jang, 2022	EMPA-REG[32]	ITT	Empagliflozin	Sitagliptin	11 563	11 563	MI	0.91	0.76, 1.08	0.87	0.70, 1.09	Y	Y	Y
Jang, 2022	EMPA-REG[32]	ITT	Empagliflozin	Sitagliptin	11 563	11 563	Stroke	0.89	0.75, 1.05	1.18	0.89, 1.56	N	Y	Y
Jang, 2022	EMPA-REG[32]	ITT	Empagliflozin	Sitagliptin	11 563	11 563	HHF	0.85	0.75, 0.95	0.65	0.50, 0.85	Y	Y	Y
Franklin, 2020	DECLARE[34]	ITT	Dapagliflozin	DPP-4i	8 582	8 578	HHF and CV death	0.70	0.59, 0.82	0.83	0.73, 0.95	Y	Y	Y[b]
Sciannameo, 2021[c]	DECLARE[34]	N/A	N/A	N/A	N/A	N/A	HHF and/or CV death	0.86	0.73, 0.99	0.83	0.73, 0.95	-	-	-

Abbreviations: HR: Hazard Ratio; CI: Confidence Interval; ITT: intention-to-treat; HHF: hospitalization due to heart failure; CV: cardiovascular; RCT: Randomized Controlled Trial; DPP-4i: Dipeptidyl peptidase 4 inhibitors

[a] SA: Full statistical significance agreement (Y/N/P) – determines if the RWD and RCT have estimates and CIs on the same side of the null; P: partial significance agreement- met the prespecified noninferiority criteria even though the database study may have indicated superiority, EA: Estimate agreement (Y/N) – determines if the effect estimate of the RWD falls within the 95% confidence interval of the RCT effect estimate. SD: Standardized difference agreement is defined by standardized differences $|Z| < 1.96$ (Y = yes, N = no).

[b] Calculated by hand; [c]Weighting RCT stratum-specific treatment effects according to proportions of a given characteristics in the target population

Table 2.4 Summary of proportion of real-world patients eligible for cardiovascular outcome trials examining newer antidiabetic drugs.

Study	RCT	Population Name	Population Size	Eligible for RCT (%)	Median (25th, 75th percentile)
SGLT-2 Inhibitors					
Birkeland, 2018	CANVAS[33]	Total population with T2DM	803,836	34.0	
Hinton, 2020	CANVAS[33]	Total population with T2DM	84,394	17.0	
Nicolucci, 2019	CANVAS[33]	Total population with T2DM	45,566	29.4	
Pintat, 2019	CANVAS[33]	Total population with T2DM	11,385	19.9	
Shao, 2019	CANVAS[33]	Canagliflozin new users	1,091	57.3	
Wittbrodt, 2019	CANVAS[33]	Total population with T2DM	172,643	32.0	
Summary	CANVAS[33]				30.7 (22.3, 33.5)
Birkeland, 2018	DECLARE[34]	Total population with T2DM	803,836	59.0	
Hinton, 2020	DECLARE[34]	Total population with T2DM	84,394	27.0	
Nicolucci, 2019	DECLARE[34]	Total population with T2DM	45,566	55.9	
Pintat, 2019	DECLARE[34]	Total population with T2DM	11,385	40.5	
Shao, 2019	DECLARE[34]	Dapagliflozin new users	4,748	50.4	
Wittbrodt, 2019	DECLARE[34]	Total population with T2DM	172,643	44.0	
Summary	DECLARE[34]				47.2 (41.4, 54.5)
Arnold, 2017	EMPA-REG[32]	SGLT-2 inhibitor users	2,497	5.2	
Arnold, 2017	EMPA-REG[32]	Total population with T2DM	182,525	26.2	
Birkeland, 2018	EMPA-REG[32]	Total population with T2DM	803,836	21.0	
Hinton, 2020	EMPA-REG[32]	Total population with T2DM	84,394	7.0	
Jang, 2022	EMPA-REG[32]	Empagliflozin, sitagliptin new uesers	384,579	12.6	
McGovern, 2017	EMPA-REG[32]	RCGP RSC Total T2DM	60,327	15.7	
McGovern, 2017	EMPA-REG[32]	RCGP RSC SGLT2-users	1,642	11.1	
Nicolucci, 2019	EMPA-REG[32]	Total population with T2DM	45,566	11.7	
Pintat, 2019	EMPA-REG[32]	Total population with T2DM	11,385	7.1	
Shao, 2019	EMPA-REG[32]	Empagliflozin new users	11,650	18.7	
Wittbrodt, 2019	EMPA-REG[32]	Total population with T2DM	172,643	26.0	
Zhou, 2020	EMPA-REG[32]	Inpatients	48,257	17.4	
Zhou, 2020	EMPA-REG[32]	Outpatients	36,857	7.2	
Summary	EMPA-REG[32]				12.6 (7.2, 18.7)
Birkeland, 2018	VERTIS-CV[43]	Total population with T2DM	803,836	17.0	
Hinton, 2020	VERTIS-CV[43]	Total population with T2DM	84,394	7.0	
Nicolucci, 2019	VERTIS-CV[43]	Total population with T2DM	45,566	12.8	
Pintat, 2019	VERTIS CV[43]	Total population with T2DM	11,385	7.2	
Shao, 2019	VERTIS-CV[43]	Total population with T2DM	33,118	19.2	
Wittbrodt, 2019	VERTIS CV[43]	Total population with T2DM	172,643	27.0	
Summary	VERTIS CV[43]				14.9 (8.6, 18.7)
GLP-1 Receptor Agonists					
Wittbrodt, 2018	ELIXA[44]	Adults likely to have T2DM	23,941,512	6.4	
Boye, 2018	EXSCEL[35]	Total population with T2DM	26,110,573	15.9	
Sciannameo, 2020	EXSCEL[35]	Total population with T2DM	16,544	13.4	
Wittbrodt, 2018	EXSCEL[35]	Adults likely to have T2DM	24,062,453	47.2	
Summary	EXSCEL[35]				15.9 (14.7, 31.6)
Wittbrodt, 2018	FREEDOM-CVO[45]	Adults likely to have T2DM	23,941,512	15.5	

Sciannameo, 2020	HARMONY[46]	Total population with T2DM	10,208	9.5	
Wittbrodt, 2018	HARMONY[46]	Adults likely to have T2DM	24,062,453	8.0	
Abrahami, 2020	LEADER[36]	New users of liraglutide	159,660	1.2	
Arnold, 2017	LEADER[36]	GLP-1 RA users	5,249	6.0	
Arnold, 2017	LEADER[36]	Total population with T2DM	182,525	48.0	
Boye, 2018	LEADER[36]	Total population with T2DM	26,110,573	12.9	
Fan, 2020	LEADER[36]	Total population with T2DM	800	15.4	
Romera, 2022	LEADER[36]	Total population with T2DM	24,268	10.1	
Sciannameo, 2020	LEADER[36]	Total population with T2DM	10,061	9.4	
Webb, 2021	LEADER[36]	Total population with T2DM	33,118	13.3	
Wittbrodt, 2018	LEADER[36]	Adults likely to have T2DM	23,941,512	12.8	
Summary	LEADER[36]				12.8 (9.4, 13.3)
Sciannameo, 2020	PIONEER-6[37]	Total population with T2DM	39,726	34.1	
Boye, 2018	REWIND[38]	Total population with T2DM	26,110,573	42.6	
Romera, 2022	REWIND[38]	Total population with T2DM	24,268	53.6	
Sciannameo, 2020	REWIND[38]	Total population with T2DM	37,574	35.8	
Webb, 2021	REWIND[38]	Total population with T2DM	33,118	44.4	
Wittbrodt, 2018	REWIND[38]	Adults likely to have T2DM	23,941,512	22.4	
Summary	REWIND[38]				42.6 (35.8, 44.4)
Boye, 2018	SUSTAIN-6[39]	Total population with T2DM	26,110,573	13.0	
Romera, 2022	SUSTAIN-6[39]	Total population with T2DM	24,268	10.4	
Sciannameo, 2020	SUSTAIN-6[39]	Total population with T2DM	9,942	10.1	
Webb, 2021	SUSTAIN-6[39]	Total population with T2DM	33,118	13.5	
Wittbrodt, 2018	SUSTAIN-6[39]	Adults likely to have T2DM	23,941,512	11.8	
Summary	SUSTAIN-6[39]				11.8 (10.4, 13.0)

Abbreviations: RCT: Randomized Controlled trial T2DM: Type 2 Diabetes Mellitus; SGLT-2: Sodium-Glucose Transport Protein 2;
GLP-1 RA: Glucagon-Like Peptide-1 Receptor Agonist

2.2.11 Figure Legend

Figure 2.1 PRISMA flow diagram of study inclusion for systematic review of replication of CVOTs using RWD in patients with T2DM

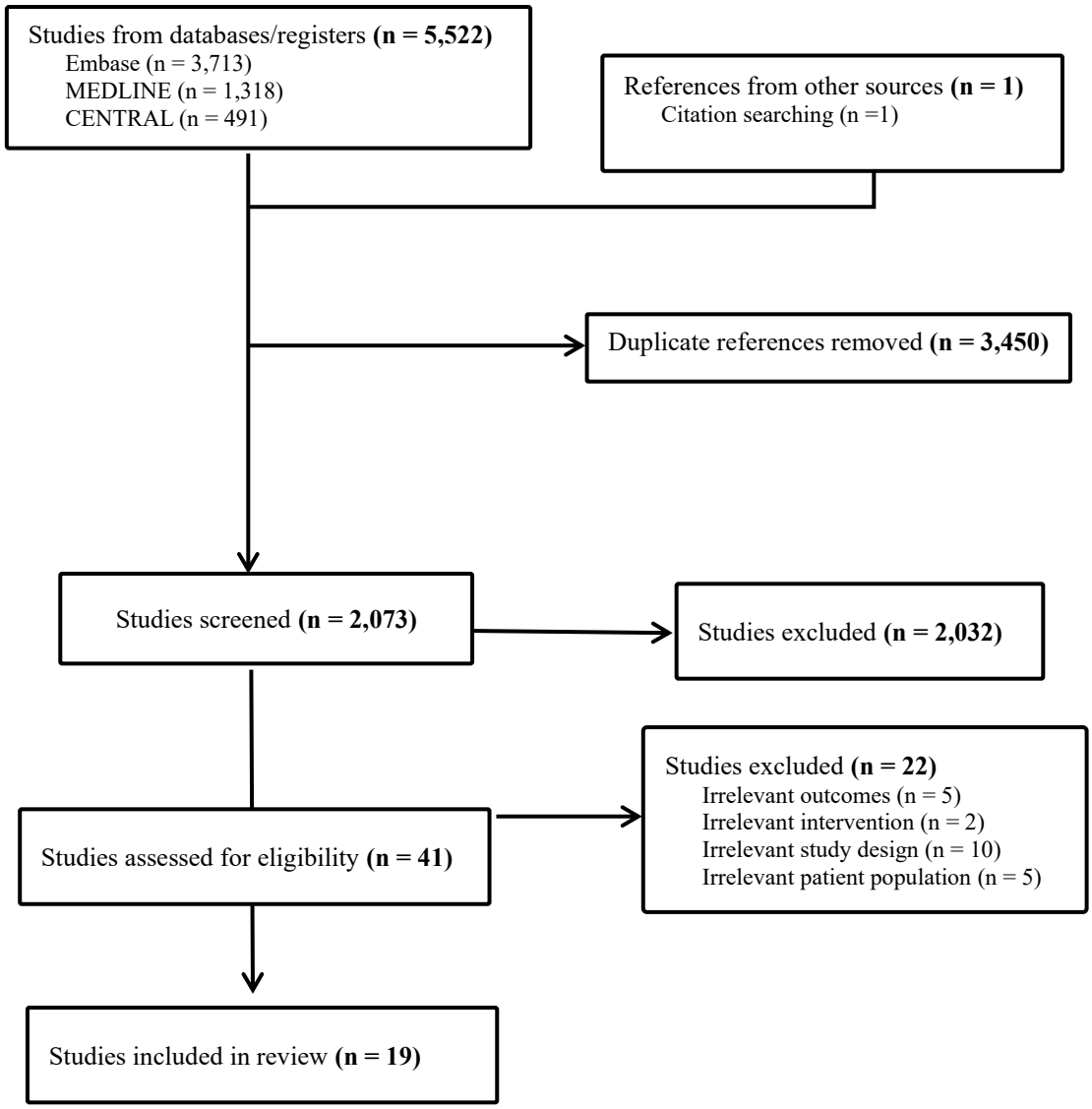

Studies from databases/registers (n = 5,522)
Embase (n = 3,713)
MEDLINE (n = 1,318)
CENTRAL (n = 491)
References from other sources (n = 1)
Citation searching (n =1)
Duplicate references removed (n = 3,450)
Studies screened (n = 2,073)
Studies excluded (n = 2,032)
Studies excluded (n = 22)
Irrelevant outcomes (n = 5)
Irrelevant intervention (n = 2)
Irrelevant study design (n = 10)
Irrelevant patient population (n = 5)
Studies assessed for eligibility (n = 41)
Studies included in review (n = 19)

Figure 2.1

Supplementary Table 2.1 Search strategy for systematic review of observational studies assessing generalizability of CVOTS using OVID MEDLINE database

	Search Entry
1.	Observational study.pt. or cohort studies/ or case-control studies/ or cross-sectional studies/ or (real-world or observation* or nonrandomi#ed or non randomi#ed or case control or cohort stud* or cross sectional or (emulat* adj3 trial?)).mp.
2.	Randomized Controlled Trial/ or Randomized Controlled Trials as Topic/ or Clinical Trial/ or Clinical Trials as Topic/ or trial?.mp.
3.	(major adverse cardiovascular event* or (cardiovascular adj3 outcome*) or all cause mortality).mp.
4.	Diabetes Mellitus/ or Diabetes Mellitus, Type 2/ or (antidiabet* or diabet* or t2d or t2dm).mp.
5.	(emul* or compar* or applicab* or generali#ab* or transportability).mp.
6.	1 AND 2 AND 3 AND 4 AND 5

Supplementary Table 2.2 Search strategy for systematic review of observational studies assessing generalizability of CVOTS using OVID EMBASE + EMBASE CLASSIC database

	Search Strategy
1.	Observational study.pt. or cohort analysis/ or case control study/ or cross-sectional study/ or (real-world or observation* or nonrandomi#ed or non randomi#ed or case control or cohort stud* or cross sectional or (emulat* adj3 trial?)).mp.
2.	Randomized Controlled Trial/ or Clinical Trial/ or trial?.mp.
3.	(major adverse cardiovascular event* or (cardiovascular adj3 outcome*) or all cause mortality).mp.
4.	Diabetes Mellitus/ or non insulin dependent diabetes mellitus/ or (antidiabet* or diabet* or t2d or t2dm).mp.
5.	(emul* or compar* or applicab* or generali#ab* or transportability).mp.
6.	1 AND 2 AND 3 AND 4 AND 5

Supplementary Table 2.3 Search strategy for systematic review of observational studies assessing generalizability of CVOTS using Cochrane CENTRAL database

	Search Strategy
1.	"observational study":pt OR [mh ^"cohort studies"] OR [mh ^"case-control studies"] OR [mh ^"cross-sectional studies"] OR (real-world OR observation* OR nonrandomi#ed OR "non randomi")
2.	[mh ^"Randomized Controlled Trial"] OR [mh ^"Randomized Controlled Trials as Topic"] OR [mh ^"Clinical Trial"] OR [mh ^"Clinical Trials as Topic"] OR trial?:ti,ab,kw
3.	(("major adverse cardiovascular" NEAR/2 event*) OR (cardiovascular NEAR/3 outcome*) OR "all cause mortality"):ti,ab,kw
4.	[mh ^"Diabetes Mellitus"] OR [mh ^"Diabetes Mellitus, Type 2"] OR (antidiabet* OR diabet* OR t2d OR t2dm):ti,ab,kw
5.	(emul* OR compar* OR applicab* OR generali#ab* OR transportability):ti,ab,kw
6.	1 AND 2 AND 3 AND 4 AND 5

Supplementary Table 2.4 A list of completed cardiovascular outcome trials and their drug class and molecule.

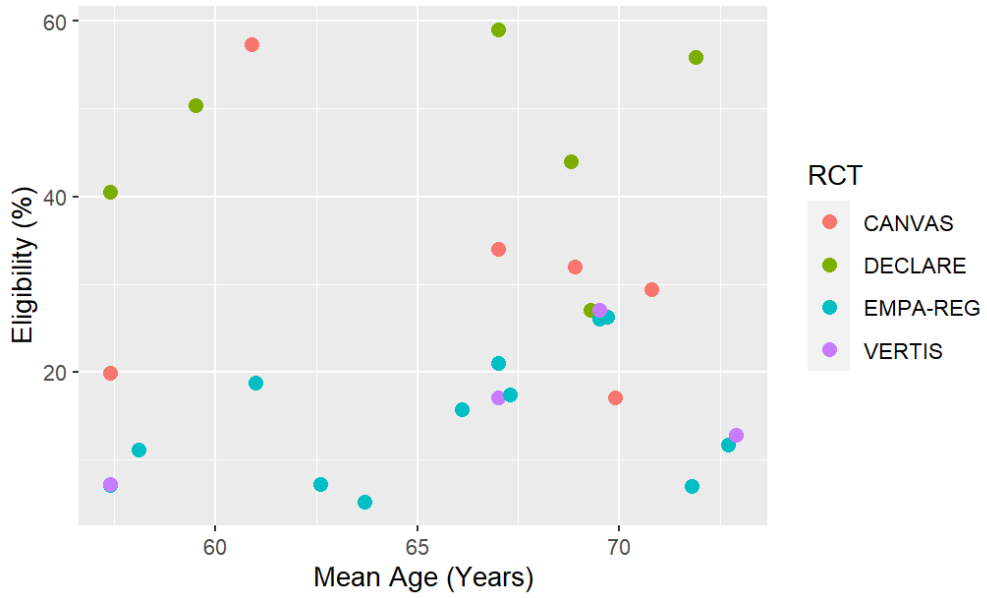
Drug Class	Trial Name	Drug Molecule
DPP-4 Inhibitor	SAVOR-TIMI[41]	Saxaglibptin
	EXAMINE	Alogliptin
	TECOS[42]	Sitagliptin
	CAROLINA[51]	Lingaliptin
	CARMELINA[40]	Linagliptin
	MK-3102[52]	Omarigliptin
GLP-1 receptor agonist	ELIXA[44]	Lixisenatide
	LEADER[36]	Liraglutide
	SUSTAIN-6[39]	Semaglutide
	PIONEER-6[37]	Semaglutide
	EXSCEL[35]	Exenatide-pragmatic trial
	REWIND[38]	Dulaglutide
	FREEDOM-CVO[45]	Exenatide in DUROS
	HARMONY[46]	Albiglutide
SGLT-2 inhibitor	EMPA-REG[32]	Empagliflozin
	CANVAS[33]	Canagliflozin
	DECLARE[34]	Dapagliflozin
	VERTIS-CV[43]	Ertugliflozin

Supplementary Table 2.5 Standardized differences computed for observational study compared to RCT hazard ratios

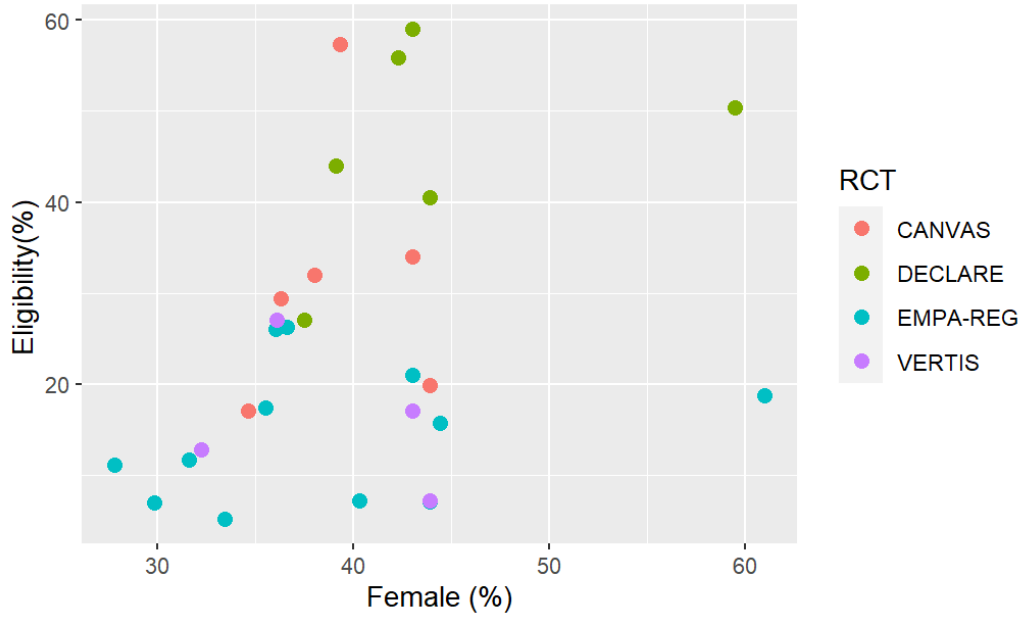
Study	RCT	Sample Size		RCT				
		Exposure	Comparator	HR	95% CI	HR	95% CI	SD
Sciannameo, 2021	EMPA-REG[32]	N/A	N/A	0.88	0.74, 1.03	0.86	0.74, 0.99	0.20
Sciannameo, 2021	DECLARE[34]	N/A	N/A	0.94	0.84, 1.04	0.93	0.84, 1.03	0.14
Sciannameo, 2021	EXSCEL[35]	N/A	N/A	0.92	0.82, 1.02	0.91	0.83, 1.00	0.15
Abrahami, 2020	LEADER[36]	1 868	25 895	1.03	0.82, 1.30	0.87	0.78, 0.97	1.09
Abrahami, 2020	LEADER[36]	1 864	32 899	0.97	0.78, 1.22	0.87	0.78, 0.97	0.73
Franklin, 2020	LEADER[36]	4 668	4 672	0.82	0.76, 0.87	0.87	0.78, 0.97	-0.88
Sciannameo, 2021	LEADER[36]	N/A	N/A	0.88	0.77, 0.99	0.87	0.78, 0.97	0.13
Sciannameo, 2021	PIONEER-6[37]	N/A	N/A	0.76	0.41, 1.10	0.79	0.57, 1.11	-0.13
Sciannameo, 2021	REWIND[38]	N/A	N/A	0.87	0.76, 0.98	0.88	0.79, 0.99	-0.12
Sciannameo, 2021	SUSTAIN-6[39]	N/A	N/A	0.73	0.47, 0.99	0.74	0.58, 0.95	-0.06
Sciannameo, 2021	SAVOR-TIMI[41]	N/A	N/A	0.99	0.87, 1.10	1.00	0.89, 1.12	-0.12
Sciannameo, 2021	DECLARE[34]-HHF/CV Death	N/A	N/A	0.86	0.73, 0.99	0.83	0.73, 0.95	0.30

Abbreviations: RCT: Randomized Controlled Trials; HR: Hazard Ratio; CI: Confidence Interval; SD: Standardized Difference; HHF: Hospitalization due to heart failure; CV: Cardiovascular

Supplementary Figure 2.1 Distribution of percent eligible for inclusion in CVOT of SGLT-2 inhibitors in real-world populations by mean age (i) and by sex (ii)

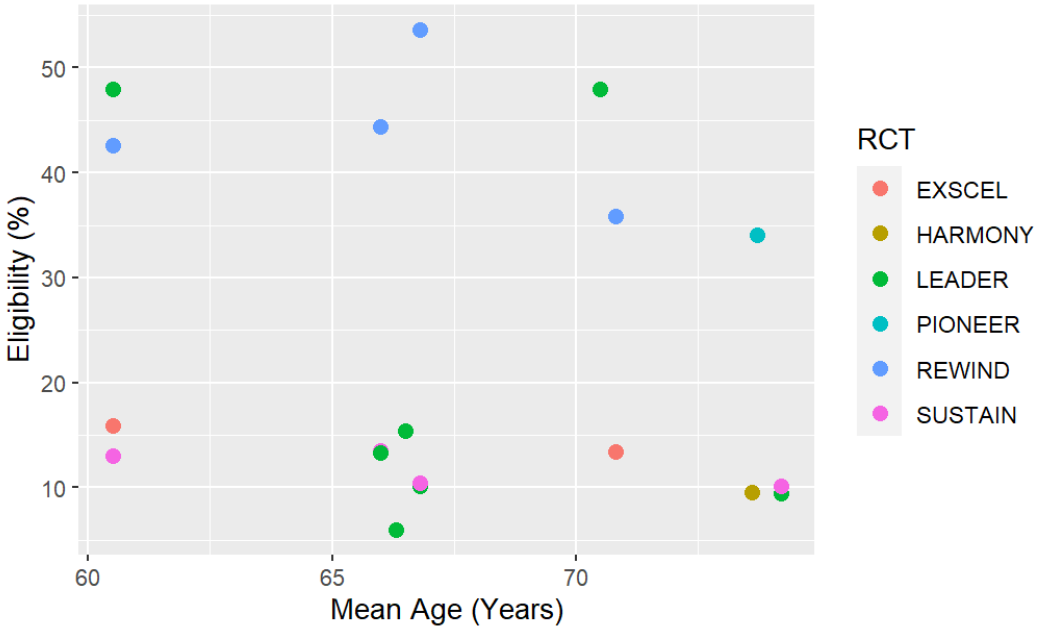


i.

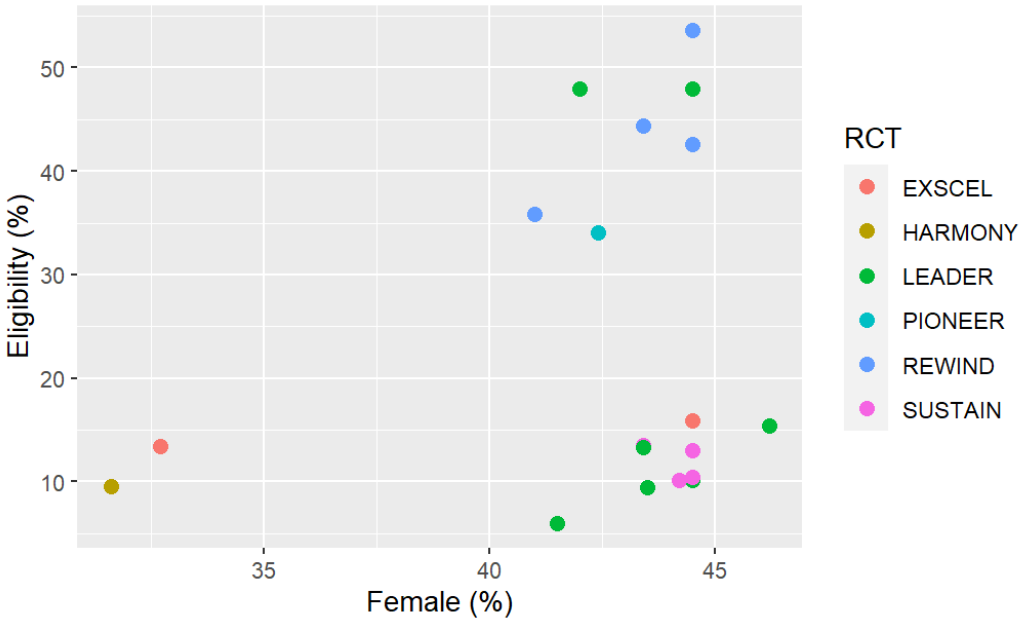


ii.

Supplementary Figure 2.2 Distribution of percent eligible for inclusion in CVOT of GLP-1 Receptor Agonists in real-world populations by mean age (i) and by sex (ii)

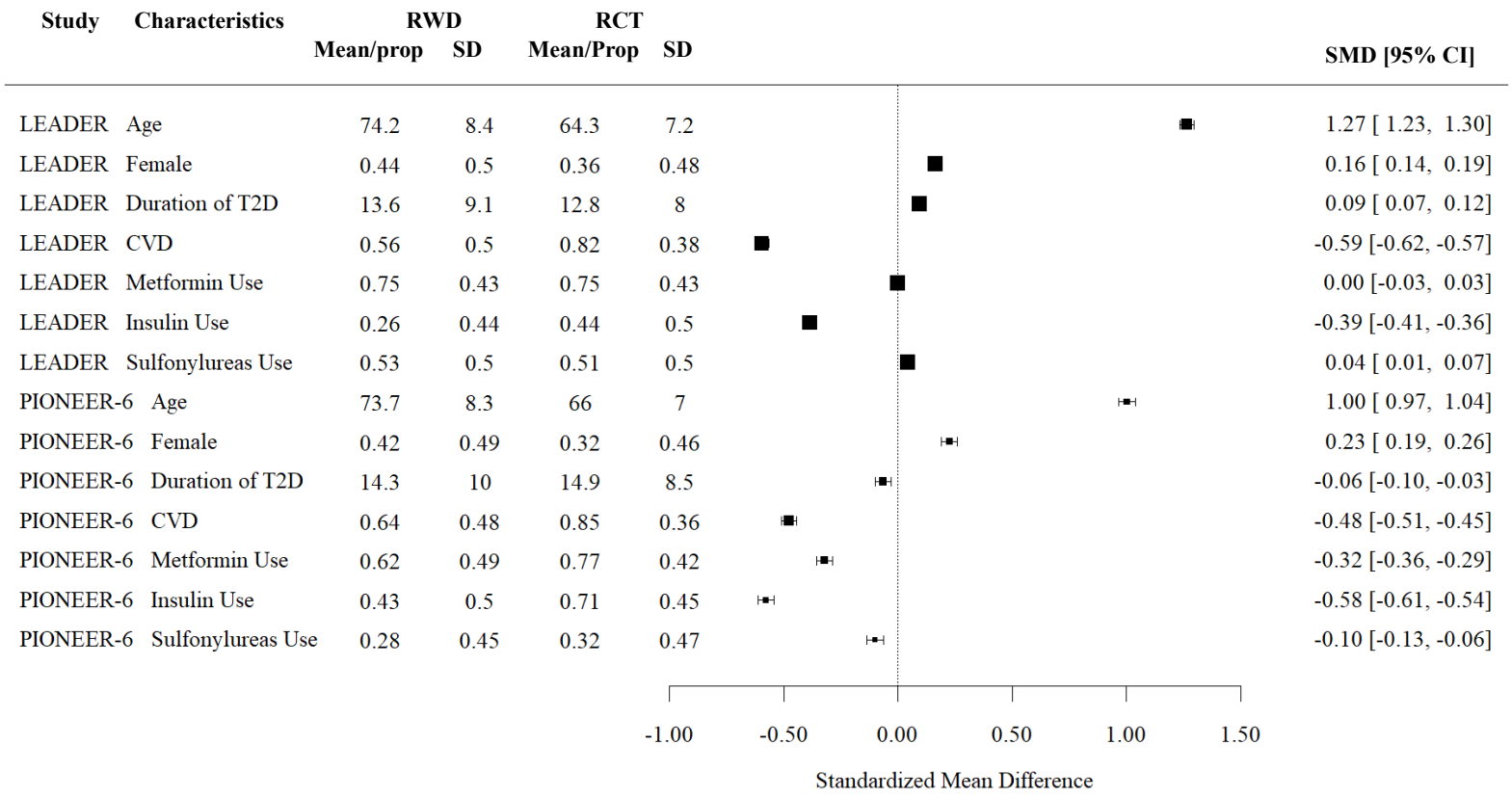


i.



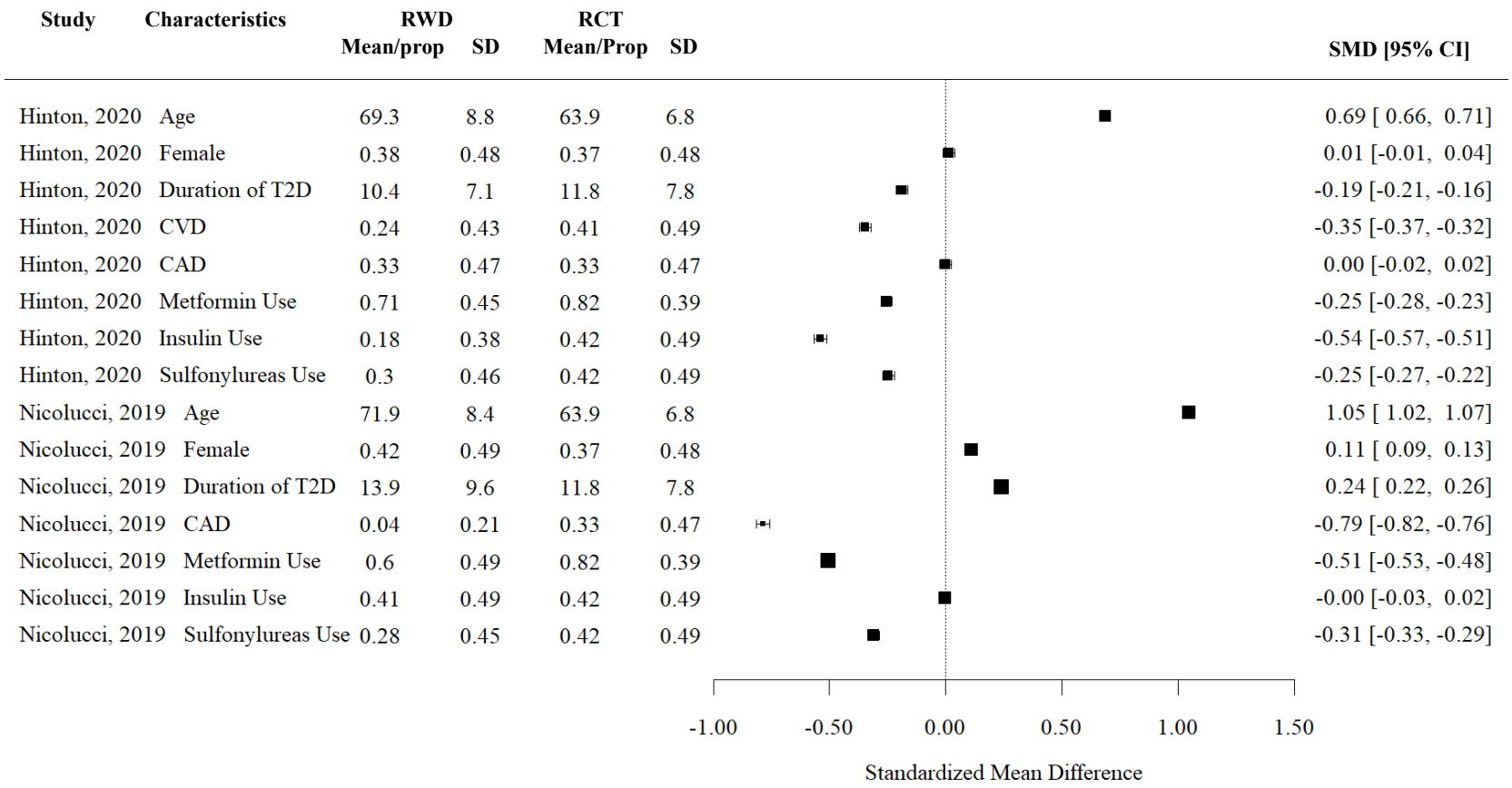
ii.

Supplementary Figure 2.3 Standardized Mean Differences of patient characteristics between GLP-1 Receptor Agonist CVOTs and the study population from the RWD study from Sciannameo, 2020.



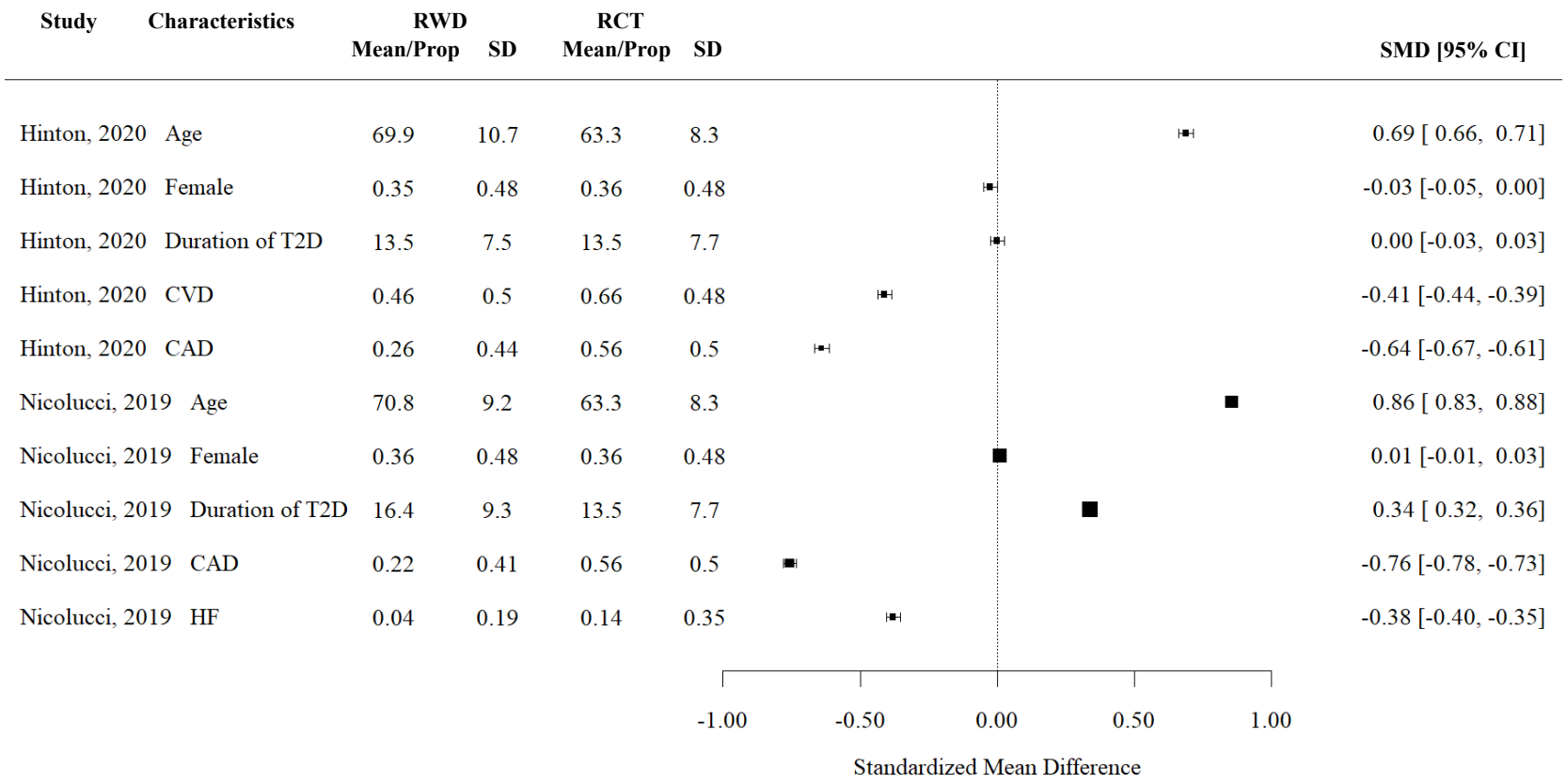
Abbreviations: CVD: cardiovascular disease, CVOT: cardiovascular outcome trial, Prop: Proportion; RCT: randomized controlled trial, RWE: real-world evidence, SD: standard deviation, SMD: standardized mean difference, T2D: Type 2 diabetes,

Supplementary Figure 2.4 Select standardized mean difference of patient characteristics for CVOT and RWD for the DECLARE trial



Abbreviations: CAD: coronary artery disease, CVD: cardiovascular disease, CVOT: cardiovascular outcome trial, Prop: Proportion, RCT: randomized controlled, RWE: real-world evidence, RWD: real-world data, SD: standard deviation, trial, SMD: standardized mean difference, T2D: Type 2 diabetes

Supplementary Figure 2.5 Select standardized mean difference of patient characteristics for CVOT and RWD for CANVAS trial



Abbreviations: CAD: coronary artery disease, CVD: cardiovascular disease, CVOT: cardiovascular outcome trial, HF: heart failure, RCT: randomized controlled, RWE: real-world evidence, RWD: real-world data, SD: standard deviation, trial, SMD: standardized mean difference, T2D: Type 2 diabetes

CHAPTER 3: TRANSITION

Our systematic review identified 15 observational studies examining the generalizability of RCTs of antidiabetic medications and their cardiovascular treatment effects among individuals with T2DM using RWD. We found that effect estimates were mostly concordant between observational studies and the trials, with many coming to the same regulatory conclusion as the trials they were emulating. We also found that the proportion of real-world patients that would have been eligible for the trials varied dramatically. Our systematic review examined the new antidiabetic medications of SGLT-2 inhibitors, GLP-1 receptor agonists, and DPP-4 inhibitors. Although insulins are exempt from the FDA's CVOT requirement, the cardiovascular effects of insulin analogues were compared in the DEVOTE trial[122]. However, to our knowledge, there has been no study that has emulated this trial. Given that patients using insulin often have longer duration and severity of disease, it is important to know how the trial results may generalize the to real-world population taking the drugs. To address this knowledge gap, I conducted a database study emulating the DEVOTE trial. I used the Clinical Practice Research Database (CPRD) linked to Hospital Episode Statistics Admitted Patient Care (HES) and Office for National Statistics Vital Statistics (ONS) databases to define my population of new users of insulin degludec and insulin glargine who have T2DM. Using the DEVOTE trial inclusion criteria, I created two subpopulations from the CPRD of DEVOTE eligible and DEVOTE ineligible patients, chapter 5 contains the corresponding manuscript.

Chapter 4: Emulating randomized controlled trials of long-acting insulins and cardiovascular events using real-world data for patients with type 2 diabetes

Wanning Wang BHSc[1,2], Pauline Reynier[1,4] MSc, Michael Webster-Clark PhD[2,3], Oriana HY Yu MD MSc[2,3,4], Kristian B. Filion PhD[1,2,3]

[1] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

[2] Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Quebec, Canada

[3] Department of Medicine, McGill University, Montreal, Quebec, Canada

[4] Division of Endocrinology and Metabolism, Jewish General Hospital/McGill University, Montreal, Quebec, Canada

Correspondence to:
Kristian B. Filion, PhD
Associate Professor and William Dawson Scholar
Departments of Medicine and of Epidemiology, Biostatistics and Occupational Health
McGill University
3755 Cote Ste-Catherine Road, Suite H410.1
Montreal, Quebec, Canada H3T 1E2

Phone: (514) 340-8222 x 28394
Fax: (514) 340-7564
Email: kristian.filion@mcgill.ca

Keywords: Long-acting insulin analogues, Real-world data, Cardiovascular outcomes

Conflict of interest disclosure: None

Sources of Support: Dr. Filion is supported by a William Dawson Scholar award from McGill University. Wanning Wang is supported by a McGill Internal Studentship Award and a CANadian Consortium of Clinical Trial TRAINing platform (CANTRAIN) studentship; CANTRAIN is funded by the Canadian Institutes of Health Research.

Word count: 3580; **Abstract word count:** 252**; Tables**: 4 **; Figures:** 1**; Supplementary tables:** 5**; Supplementary figures:** 2

Twitter: @KB_Filion

4.1 ABSTRACT

Introduction: Randomized controlled trials (RCTs) have high internal validity but often have limited generalizability. To our knowledge, there are no studies examining potential differences between RCTs and real-world data (RWD) in patient characteristics and risk of cardiovascular outcomes among patients with type 2 diabetes (T2DM) treated with long-acting insulin analogues.

Methods: We emulated the DEVOTE trial of insulin degludec vs glargine among patients with T2DM using data from the UK's Clinical Practice Research Datalink. Two subpopulations were created based on DEVOTE eligibility. Cox proportional hazards models with inverse probability of treatment weighting were used to estimate hazard ratios (HRs) and corresponding confidence intervals (CIs) for major adverse cardiovascular events (MACE) comparing new users of insulin degludec to new users of insulin glargine overall and in the two subpopulations.

Results: There were 10,430 patients in the overall population, 5,280 in the DEVOTE eligible population, and 5,150 in the DEVOTE ineligible population. The overall (HR: 1.36, 95% CI: 0.83, 1.86) and DEVOTE eligible populations (HR: 1.07, 95% CI: 0.63, 1.58) were compatible with findings from the DEVOTE trial (HR: 0.91, 95% CI: 0.78, 1.06) for the risk of MACE. Due to a low number of outcomes, the DEVOTE ineligible population had deviations in point estimates and wider CIs (HR: 2.19, 95% CI: 0.30, 3.83).

Conclusion: The risk of MACE among patients with T2DM newly prescribed insulin degludec compared to insulin glargine was consistent between the overall population and the subpopulation eligible for the DEVOTE trial, while the DEVOTE ineligible population had discrepant point estimates.

4.2 INTRODUCTION

Insulin is prescribed to patients with type 2 diabetes mellitus (T2DM) when other antidiabetic drugs have failed to achieve or maintain glycemic control and to prevent micro- (retinopathy, neuropathy and nephropathy) and macrovascular (cardiovascular) complications[1,2]. There are two main types of insulin, human insulin (often neutral protamine hagedorn [NPH]) and long-acting insulin analogues (glargine, degludec, detemir)[3]. The efficacy of long-acting insulin analogues for the prevention of adverse cardiovascular events has been examined in two large randomized controlled trials (RCTs). However, the inclusion criteria of these trials were highly selective and may have resulted in trial populations that do not represent the characteristics of patients using long-acting insulin analogues in everyday clinical practice. For example, the ORIGIN trial selected individuals with pre-diabetes only[4]. In the DEVOTE trial[5], which compared cardiovascular safety of insulin degludec and insulin glargine, the inclusion criteria were based on blood glucose levels and the presence of elevated cardiovascular risk factors; while this increased the number of events (and thus reduced the sample size required to achieve adequate statistical precision), it also makes the generalizability of its findings to patients without elevated cardiovascular risk unclear.

As real-world data (RWD) generated from electronic health records, claims data, and registries are becoming readily available, there has been a push by regulatory and health technology assessment agencies to utilize real-world evidence (RWE) in conjunction with RCTs to evaluate the effectiveness and safety of medical interventions[6–8]. Some studies examining the safety and efficacy of treatments of T2DM have produced concordant findings between RCTs and observational studies[9]. However, there is currently no literature comparing differences in patient characteristics and outcomes between RCTs and real-world populations among patients with T2DM using long-acting insulin analogues. The objective of this study was to assess the generalizability and representativeness of DEVOTE by examining the effect of long-acting insulin analogues on cardiovascular outcomes among patients with T2DM using RWD overall and in subpopulations defined by DEVOTE trial eligibility.

4.3 METHODS

4.3.1 Data Source

We conducted an observational study emulating a RCT using population-based data from the United Kingdom's Clinical Practice Research Database (CPRD) Aurum. CPRD Aurum

consists of primary care provider records of over 41 million patients in the UK, accounting for ~20% of the population[10]. It contains longitudinal routinely collected electronic health records, with information on demographic characteristics, diagnoses and symptoms, drug exposures, vaccination history, laboratory tests, and referrals to hospital and specialist care. The CPRD is representative of the English population regarding geographical spread, socioeconomic deprivation, age, and gender[11]. Given the rich primary care data, CPRD Aurum is well suited for studies of people with T2DM since in the UK, T2DM is predominantly treated by general practitioners[12]. CPRD diagnoses and non-prescription information are coded using SNOMED CT (UK Edition), Read Version 2, and local EMIS Web ® codes. Drug and device prescriptions are coded using the Dictionary of Medicines and Devices and stored in the EMIS Web ® electronic medical record[11]. CPRD Aurum data were linked to Hospital Episode Statistics (HES) Admitted Patient Care (APC) and Office of National Statistics (ONS) death registration data. HES APC provides hospitalization data in the UK, including date of admission, date of discharge and diagnoses made during hospital stay[13]. ONS death registration data contain records of all deaths in the UK, with the deceased's age, sex, and cause of death. HES APC and ONS diagnoses and causes of death are coded using the International Classification of Disease and Related Health Problems (ICD) 10th revision. Linkage of HES APC and ONS is available for 93% of the patients in the CPRD Aurum and has been previously validated[14].

The protocol of this study (19_217) was approved by the CPRD's Independent Scientific Advisory Committee (ISAC) and by the CIUSSS West-Central Montreal Research Ethics Board (Montréal, Canada).

4.3.2 Study population

The study population included patients diagnosed with T2DM who were prescribed insulin degludec or insulin glargine between March 1st, 2013 (when degludec became available in the UK) and November 31st, 2018 (end of data availability). The cohort was restricted to patients newly treated with insulin glargine or degludec, including those who switched from an oral antidiabetic medication or from another basal insulin to one of the treatments of interest. The date of the first prescription of insulin glargine or degludec defined study cohort entry. We excluded patients that had 1) < 1 year of recorded medical history (to capture comorbidities); 2) age <18 years; 3) previous diagnosis of type 1 diabetes (to ensure we capture only patients with T2DM); 4) diagnosis of gestational diabetes in the previous year (to ensure we captured only patients with

T2DM); 5) zero days of follow-up (to capture outcomes); and 6) patients prescribed both long-acting insulin analogues on the day of cohort entry (to prevent mixing effects). Patients were followed until they experienced an outcome of interest (described below), censoring due to death of any cause, end of registration with CPRD, or the end of study period (November 31st, 2018), whichever occurred first. The population was limited to those with linkage between CPRD, HES, and ONS.

Using the cohort described above, we created two sub-populations according to the inclusion criteria of the DEVOTE trial. Those who do not meet the inclusion criteria composed our DEVOTE ineligible subpopulation. **Supplementary Table 4.1** outlines the operationalization of the RCT inclusion criteria for the CPRD and the number of patients that met each criterion.

4.3.4 Exposure

In the DEVOTE trial, participants were randomized to insulin degludec or insulin glargine. We replicated this comparison, using glargine as an active comparator. In addition, we used an intention-to-treat (ITT) approach to define exposure in our study to reflect the study design and analysis used in the DEVOTE trial. An ITT approach is the suggested framework for studies that emulate RCT with observational data to ensure a useful treatment effect to compare to RCTs[15,16]. To mimic the ITT analyses of a trial, we used a time-fixed exposure definition where patients were classified into one of two mutually-exclusive categories (new user of insulin degludec or new user of insulin glargine) from study entry until the end of follow-up regardless of treatment switching or discontinuation. For all exposure categories, use of other antidiabetic drugs including other insulins were permitted.

4.3.5 Outcome

Our primary outcome was the occurrence of major adverse cardiovascular events (MACE), a composite endpoint of myocardial infarction (MI), ischemic stroke, or cardiovascular death. For our secondary outcomes, we examined the individual components of MACE, all-cause mortality, hospitalization for heart failure, and hospitalization for hypoglycemia. Hospitalizations due to heart failure and hypoglycemia were not reported in the DEVOTE trial but were included as they were important safety signals from previous CVOTs of other antidiabetic medications[20]. We used the following ICD-10 codes in HES and ONS to define the outcomes of interest: MI (ICD-10: I21.x), ischemic stroke (ICD-10: I63.x-I64.x), and hospitalization for heart failure (ICD-10: I11.0,

I13.0, I13.2, I50.x). The event date was defined by the date of admission for HES-defined events and date of death for ONS-defined events. Death from cardiovascular disease was defined using the underlying cause of death in ONS (ICD-10: I00.x-I78.x [except 146.9]). All-cause mortality was defined using CPRD, HES, and ONS, with the earliest recorded date of death defining the event date. Hypoglycemia was defined by a relevant ICD-10 code in HES (E16.0, E16.1, E16.2).

4.3.6 Covariates

We included three types of baseline covariates in our propensity score (PS) model: demographics, comorbidities, and medications used at cohort entry. Covariates measured at baseline include year of cohort entry, age, sex, race, Index of Multiple Deprivation deciles, duration of treated diabetes (time since first prescription for an antidiabetic drug and cohort entry date), smoking status, previous history of alcohol related disorders, atrial fibrillation, previous diagnosis of cancer (not including non-melanoma skin cancer), chronic obstructive pulmonary disease, acute kidney injury, chronic kidney disease, retinopathy, neuropathy, dialysis, cerebrovascular disease, body mass index (BMI)(last measurement before cohort entry), A1C (last measured before cohort entry), and estimated glomerular filtration rate (eGFR). The medications assessed at baseline were ACE inhibitors, angiotensin II receptor blockers, beta blockers, diuretics, statins, acetylsalicylic acid, non-steroidal anti-inflammatory drugs, fibrates, and use of other diabetic medications (metformin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, alpha-glucosidase inhibitors, meglitinides, SGLT-2 inhibitors, and other types of insulin). For our MACE models, we further adjusted for coronary artery disease, hyperlipidemia, hypertension, peripheral vascular disease, stroke, MI, coronary revascularization, systolic blood pressure (latest measure before cohort entry), diastolic blood pressure (latest measure before cohort entry), total cholesterol, and use of calcium channel blockers, oral anticoagulants, and antiplatelets. For the hypoglycemia models, we also included history of hypoglycemia in the year prior to cohort entry, thyroid disease (including hypothyroidism and hyperthyroidism), liver cirrhosis, previous medication usages of acetaminophen, opioids, and glucagon.

We used multiple imputation by chained equations (MICE) to impute missing values for the variables of race, Index of Multiple Deprivation decile, smoking status, A1C, eGFR, BMI, and systolic and diastolic blood pressure.

4.3.7 Statistical Analyses:

We compared the patient characteristics between those eligible and ineligible for the DEVOTE trial and our overall population to the DEVOTE trial using absolute values of the standardized difference with standardized differences greater than 0.1 considered important[17]. We used proportions with 95% confidence intervals (CIs) to calculate the real-world populations who would have been eligible for the trial.

We used PS to account for differential baseline covariates between study groups. We included all covariates mentioned above in our model using multivariable logistic regression, with the inverse of the computed PS subsequently used to weigh exposure groups via inverse probability of treatment weights (IPTW). We estimated absolute values of standardized differences to compare the characteristics of each exposure group before and after weighting. After stabilized IPTW, we truncated weights at 10 to minimize the impact of extreme weights[18]. We used Poisson regression to calculate crude incidence rates and 95% CIs for each subpopulation for the primary and secondary outcomes. Using the exposure groups created after IPTW, time-fixed Cox proportional hazards models were used to estimate hazard ratios (HRs) and bootstrapped 95% CIs for MACE in our three study populations.

We performed two secondary analyses. First, we repeated our analysis for individual components of MACE as well as all-cause mortality, hospitalization due to heart failure, and hospitalization due to hypoglycemia. Second, similar to the DUPLICATE studies[9,19], we estimated three agreement statistics (statistical agreement, estimate agreement and standardized difference agreement) to compare our RWD to RCTs for all three of our populations. The definitions for these agreement statistics can be found in the **Supplementary Methods.** We conducted six sensitivity analyses to examine the robustness of our results; these analyses are described in the **Supplementary Methods.**

4.4 RESULTS

Figure 4.1 describes our study cohort composition. Our CPRD cohort consisted of 10,430 patients with T2DM, of which 812 initiated insulin degludec and 9,618 initiated insulin glargine. From our CPRD cohort, 5,280 patients were eligible for the DEVOTE trial (51% [95% CI: 49%, 52%] of the total population). The mean follow-up time was 1.5 years among insulin degludec initiators and 2.1 years among insulin glargine initiators. **Table 4.1** and **Supplementary Table**

4.2 describe baseline patient characteristics for each exposure group. Across all subpopulations, degludec users were younger, had higher BMI, and greater use of other anti-diabetic medications. **Table 4.2** and **Supplementary Table 4.3** describe absolute standardized mean differences between populations after multiple imputation and weighting. After imputation and weighting, most characteristics had a standardized difference of <0.1. **Supplementary Figure 4.1** and **Supplementary Figure 4.2** illustrate the absolute standardized mean differences between the DEVOTE eligible and DEVOTE ineligible subpopulations and between the trial population and the RWD populations, respectively, demonstrating large degrees of heterogeneity in patient characteristics.

The results of our primary analyses are reported in **Table 4.3**. In the overall CPRD population, the incidence rate of MACE was 39.4 events per 1000 person-years (95% CI: 29.6, 52.4) in the insulin degludec group and 45.5 events per 1000 person-years (95% CI: 42.6, 48.6) in the insulin glargine group. The incident rates for the DEVOTE eligible population were 71.4 events per 1000 person-years (95% CI: 52.2, 97.8) and 79.7 events per 1000 person-years (95% CI: 74.2, 85.6), respectively. The DEVOTE ineligible patients had lower incidence rates for both exposure groups (insulin degludec: 12.3 events per 1000 person-years, 95% CI: 6.2, 24.7; insulin glargine: 15.3 events per 1000 person-years, 95% CI: 13.1, 17.8). In the overall CPRD population, the adjusted HR for MACE with degludec versus glargine was 1.36 (95% CI 0.83, 1.86). The adjusted HR was 1.07 (95% CI: 0.63, 1.58) for the DEVOTE eligible population and 2.19 (95% CI: 0.30, 3.83) for DEVOTE ineligible population. The DEVOTE trial had a HR of 0.91 (95% CI: 0.78, 1.06).

All the primary analyses failed to reject the null at a p-value of 0.05, which is consistent with the results of the DEVOTE trial. However, due to our wide CIs, we were unable to achieve the FDA non-inferiority upper limit of 1.3 and thus did not achieve statistical agreement with the DEVOTE trial. Estimate agreement was unable to be met as our effect estimates fell outside the CI of the DEVOTE trial. Only the DEVOTE eligible population was able to achieve standardized difference agreement with the trial results.

Table 4.4 presents the results of our secondary analyses of the individual components of MACE, hospitalization due to heart failure, all-cause mortality, and hospitalization due to hypoglycemia. For MI, HR estimates were compatible with the findings of the trial across all three observational populations. For stroke, due to low sample size, HR estimates varied greatly. For

CV death, HR estimates were compatible among the RWD populations; however, they were on the opposite side of the null compared to the DEVOTE trial. For hospitalization due to heart failure, HR estimates were compatible among the three populations. Hospitalization due to hypoglycemia saw differences between the three populations, with the DEVOTE ineligible sub-population effect estimate below the null. For all-cause mortality, the CIs overlapped for the three populations, with all of them achieving standardized difference agreement with the original trial. The overall CPRD population achieved estimate agreement and the DEVOTE eligible achieved full significance agreement with the DEVOTE trial.

Supplementary Table 4.4 reports the results from our sensitivity analyses including using inverse odds weighting (IOW) to examine transportability. These results from our sensitivity analyses were consistent with those from our primary analysis. **Supplementary Table 4.5** describes reasons for censoring, of note, adherence was low in our study, with 80% of our study population discontinuing their treatment before the end of follow-up.

4.5 DISCUSSION

Our study demonstrated that there is a lack of generalizability of the DEVOTE trial to patients with T2DM treated in a real-world setting. Only half of the real-world population would have been eligible for the DEVOTE trial, with notable differences in patient characteristics such as age, use of comedications, and comorbidities between the DEVOTE eligible and ineligible populations. There were also large differences in patient characteristics such as duration of diabetes, comedication use, and comorbidities between the trial and the real-world population in our study. Our findings for the risk of MACE among patients prescribed insulin degludec compared to insulin glargine were consistent across the three populations examined. The DEVOTE eligible subpopulation had an estimate compatible with the original DEVOTE trial, with both suggesting no clinically important difference in risk of MACE between degludec and glargine. In the DEVOTE ineligible population, on the other hand, the low number of events resulted in extremely imprecise estimates. For individual components of MACE, the treatment estimates for MI had the highest degree of agreement with the original DEVOTE trial for all three populations. The other components of MACE, all-cause mortality, hospitalization due to heart failure, hospitalization due to hypoglycemia had varying levels of compatibility among the three populations in the study and, when reported, with the DEVOTE trial. Sensitivity analyses revealed that results were robust to study assumptions.

Trials are known to have limited generalizability to real-world populations, which is evident from our findings. This is, in part, by design. The FDA guidelines for cardiovascular outcome trials target older individuals with T2DM who are at elevated cardiovascular risk which is beneficial in increasing the number of events and reducing the required sample size[21]. Although insulins are exempt from the FDA CVOT requirement[21], a similar approach was used in DEVOTE. Due to this trial eligibility criteria, the incidence rate was higher in the DEVOTE eligible population compared to the DEVOTE ineligible or the overall CPRD population in our study. While this selection process resulted in the inclusion of higher-risk patients and a more precise estimate, half of the real-world population was excluded. The increasing use of pragmatic trial designs will hopefully increase the generalizability of future trials in this area.

We found that the DEVOTE eligible population often had the highest level of agreement with the DEVOTE trial. This finding provides important reassurance that RWD can be used to complement existing RCT evidence when rigorous methods are used, and attempts are made to emulate RCTs as close as possible. This is especially pertinent as there has been an effort made by regulators in the US, Canada, and European Union to incorporate the use of RWD in their decision making[6,7,22].

We conducted sensitivity analyses to assess transportability such as using IOW to transport our observational study to the trial population based on the distribution of MI in the DEVOTE trial. However, we were limited to aggregate trial-level data. Future research using more granular data will allow for more robust comparisons between patient characteristics of trial and RWD populations using PS and probability distributions[23]. Utilizing patient level trial data for IOW would allow for the inclusion of additional covariates to account for potential effect modifiers and to see how effect estimates may differ based on distribution of covariates that more accurately reflect the trial population[24].

To our knowledge, there have been no previous studies that emulated the DEVOTE trial with RWD. Previous studies have emulated other cardiovascular outcome trials for antidiabetic medications[19,25–27]. For example, the DUPLICATE Team emulated 32 RCTs using non-randomized data from US health care claims data with eight trials evaluating cardiovascular outcomes of antidiabetic medication[19]. All eight emulations achieved either full or partial statistical significance agreement, two did not achieve estimate agreement, and one did not have standardized difference

agreement[19]. In a meta-analysis of their 32 trial emulations, they found that the heterogeneity between RCT and RWD effect estimates was mainly attributed to three emulation differences: treatment started in hospital, discontinuation of some baseline treatment at randomization, and delayed onset of drug effects[28]. While our data availability did not include in-hospital drug data and may miss treatments that started in hospital, diabetes treatment often is monitored by the general practitioner in the UK, and we therefore expect subsequent prescriptions to be captured within the CPRD. In terms of discontinuation of some baseline treatment at randomization, the DEVOTE trial allowed for concomitant use of other antidiabetic medications. Our study may be prone to delayed onset of drug effects due to the short duration of medication persistence in clinical practice. Our study adds to this existing body of work as it evaluates an RCT that was not previously studied. It used a primary care database with richer clinical data than administrative claims data.

Our study has many strengths. First, we used a large primary care database that is representative of the real-world population. The CPRD provides clinical laboratory measures such as BMI and A1C to account for clinically relevant covariates. Second, we are the first study to replicate the DEVOTE trial to evaluate the risk of MACE for patients with T2DM prescribed insulin degludec and glargine. Compared to ORIGIN[4], the other cardiovascular outcome trial in the area, the DEVOTE trial allowed for a stronger emulation. DEVOTE selected for patients who were more likely to receive the treatment in a real-world setting and included an active comparator, reducing time-lag bias and potential confounding by indication. ORIGIN included patients with prediabetes and used standard of care, which is difficult to replicate and may introduce residual confounding. Third, we used IPTW to balance our population and then stabilized and truncated the weights to reduce the influence of extreme weights. Fourth, to our knowledge, we are also the first study emulating CVOTs to compare the patient characteristics and risk of cardiovascular outcomes of the eligible vs non-eligible populations to ascertain differences between those who were and were not eligible to enroll in a trial.

There are several potential limitations to our study. First, due to the small sample size, our study had a low number of events, leading to a lack of precision in our estimates. Second, even after stabilized IPTW weighting, there were some differences in covariate distributions between exposure groups. Residual confounding is thus possible. We did not further adjust for these differences, as we used bootstrapping to adjust our variance, and the current literature recommends

against doing so in this setting[29]. Double adjustment only leads to doubly robust estimates if both models are known to be accurate, which we cannot guarantee for our outcome model[29]. Third, given the amount of treatment discontinuation present in our cohort and our use of an ITT approach, exposure misclassification may have biased our results towards the null. Our on-treatment sensitivity analysis saw a shift of our effect estimate towards that reported in the DEVOTE trial.

4.6 CONCLUSIONS

Our study demonstrated that the DEVOTE trial lacked generalizability to a real-world population of patients with T2DM taking long-acting insulin analogues. Only half of the real-world population would be eligible for the DEVOTE trial, with notable differences in patient characteristics between those eligible and ineligible for the trial. The risk of MACE among patients with T2DM prescribed insulin degludec compared to insulin glargine was consistent for the overall CPRD, DEVOTE eligible, and DEVOTE ineligible populations. The DEVOTE eligible population had effect estimates most compatible with the original DEVOTE trial. However, risks of individual components of MACE, heart failure, and hospitalization due to hypoglycemia varied across the three populations. Additional research is required to further understand the complementary nature of RCTs and RWD when assessing the safety and effectiveness of long-acting insulin analogues among patients with T2DM.

4.7 ACKNOWLEDGMENTS

KBF is supported by a William Dawson Scholar award from McGill University. WW is supported by a McGill Internal Studentship Award and a CANadian Consortium of Clinical Trial TRAINing platform (CANTRAIN) studentship; CANTRAIN is funded by the Canadian Institutes of Health Research

4.8 DISCLOSURES

None to declare.

4.9 AUTHORS' CONTRIBUTIONS

WW led the protocol development and drafted the manuscript. PR conducted the data management and analyses. All authors contributed to protocol development, were involved in data interpretation, critically reviewed the manuscript for important intellectual content, and approved the final manuscript. KBF conceived of the study idea, supervised the study, and is the guarantor.

4.10 REFERENCES

1. Howard-Thompson A, Khan M, Jones M, George CM. Type 2 Diabetes Mellitus: Outpatient Insulin Management. afp. 2018 Jan 1;97(1):29–37.

2. American Diabetes Association. 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes—2019. Diabetes Care. 2018 Dec 7;42(Supplement_1):S90–102.

3. Tibaldi JM. Evolution of Insulin: From Human to Analog. The American Journal of Medicine. 2014 Oct 1;127(10):S25–38.

4. The ORIGIN Trial Investigators. Basal Insulin and Cardiovascular and Other Outcomes in Dysglycemia. N Engl J Med. 2012 Jul 26;367(4):319–28.

5. Marso SP, McGuire DK, Zinman B, Poulter NR, Emerson SS, Pieber TR, et al. Efficacy and Safety of Degludec versus Glargine in Type 2 Diabetes. N Engl J Med. 2017 Aug 24;377(8):723–32.

6. U.S. Food & Drug Administration. Framework for FDA's Real-World Evidence Program [Internet]. 2018. Available from: https://www.fda.gov/media/120060/download

7. Canada H. Regulatory review of drugs and devices: Strengthening the use of real world evidence for drugs [Internet]. 2018 [cited 2023 Jun 19]. Available from: https://www.canada.ca/en/health-canada/corporate/transparency/regulatory-transparency-and-openness/improving-review-drugs-devices/strengthening-use-real-world-evidence-drugs.html

8. CADTH. Guidance for Reporting Real-World Evidence [Internet]. Canada: Canada's Drug and Health Technology Agency; 2023 May p. 66. (CADTH Methods and Guidelines). Available from: https://www.cadth.ca/sites/default/files/RWE/MG0020/MG0020-RWE-Guidance-Report.pdf

9. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies. Circulation. 2021 Mar 9;143(10):1002–13.

10. Clinical Practice Research Datalink. CPRD Aurum March 2022 [Internet]. [cited 2023 May 4]. Available from: https://www.cprd.com/cprd-aurum-march-2022-dataset

11. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. Int J Epidemiol. 2019 Dec;48(6):1740–1740g.

12. Hobbs FDR, Bankhead C, Mukhtar T, Stevens S, Perera-Salazar R, Holt T, et al. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007-14. Lancet. 2016 Jun 4;387(10035):2323–30.

13. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). Int J Epidemiol. 2017 Aug;46(4):1093–1093i.

14. Persson R, Vasilakis-Scaramozza C, Hagberg KW, Sponholtz T, Williams T, Myles P, et al. CPRD Aurum database: Assessment of data quality and completeness of three important comorbidities. Pharmacoepidemiology and Drug Safety. 2020;29(11):1456–64.

15. Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, et al. Observational Studies Analyzed Like Randomized Experiments: An Application to Postmenopausal Hormone Therapy and Coronary Heart Disease. Epidemiology. 2008 Nov;19(6):766.

16. Toh S, Manson JE. An Analytic Framework for Aligning Observational and Randomized Trial Data: Application to Postmenopausal Hormone Therapy and Coronary Heart Disease. Stat Biosci. 2013 Nov;5(2):344–60.

17. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009 Nov 10;28(25):3083–107.

18. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol. 2009 Jul 28;9:57.

19. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376–85.

20. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B, et al. Saxagliptin and Cardiovascular Outcomes in Patients with Type 2 Diabetes Mellitus. New England Journal of Medicine. 2013 Oct 3;369(14):1317–26.

21. Regier EE, Venkat MV, Close KL. More Than 7 Years of Hindsight: Revisiting the FDA's 2008 Guidance on Cardiovascular Outcomes Trials for Type 2 Diabetes Medications. Clin Diabetes. 2016 Oct;34(4):173–80.

22. Use of real-world evidence in regulatory decision making – EMA publishes review of its studies | European Medicines Agency [Internet]. [cited 2024 May 30]. Available from: https://www.ema.europa.eu/en/news/use-real-world-evidence-regulatory-decision-making-ema-publishes-review-its-studies

23. Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The Use of Propensity Scores to Assess the Generalizability of Results from Randomized Trials. Journal of the Royal Statistical Society Series A: Statistics in Society. 2011 Apr 1;174(2):369–86.

24. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR. Transportability of Trial Results Using Inverse Odds of Sampling Weights. Am J Epidemiol. 2017 Oct 15;186(8):1010–4.

25. Jang HY, Kim IW, Oh JM. Using real-world data for supporting regulatory decision making: Comparison of cardiovascular and safety outcomes of an empagliflozin randomized clinical trial versus real-world data. Frontiers in Pharmacology [Internet]. 2022 [cited 2023 Jul 25];13. Available from: https://www.frontiersin.org/articles/10.3389/fphar.2022.928121

26. Abrahami D., Pradhan R., Yin H., Honig P., Andre E.B., Azoulay L. Use of real-world data to emulate a clinical trial and support regulatory decision-making: assessing the impact of temporality, comparator choice and methods of adjustment. Clin Pharmacol Ther. 2020;((Abrahami, Pradhan, Yin, Azoulay) Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Canada).

27. Sciannameo V, Berchialla P, Avogaro A, Fadini GP, Consoli A, Formoso G, et al. Transposition of cardiovascular outcome trial effects to the real-world population of patients with type 2 diabetes. Cardiovascular Diabetology. 2021 May 10;20(1):103.

28. Heyard R, Held L, Schneeweiss S, Wang SV. Design differences and variation in results between randomised trials and non-randomised emulations: meta-analysis of RCT-DUPLICATE data. BMJ Medicine [Internet]. 2024 Feb 1 [cited 2024 Apr 7];3(1). Available from: https://bmjmedicine.bmj.com/content/3/1/e000709

29. Li X, Shen C. Doubly Robust Estimation of Causal Effect. Circulation: Cardiovascular Quality and Outcomes. 2020 Jan;13(1):e006065.

Table 4.1 Baseline characteristics of CPRD, DEVOTE eligible and DEVOTE ineligible populations of patients with type 2 diabetes who initiated insulin degludec or insulin glargine before inverse probability of treatment weighting and multiple imputation.

	CPRD			DEVOTE eligible			DEVOTE ineligible		
	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD
Number of patients	9,618	812		4,904	376		4,714	436	
Age (years) Mean (SD)	63.9 (14.1)	60.5 (12.9)	0.25	71.8 (10.3)	68.1 (10.0)	0.38	55.6 (12.9)	54.1 (11.6)	0.13
Males, n (%)	5,381 (55.9)	443 (54.6)	0.03	2,724 (55.5)	210 (55.9)	0.01	2,657 (56.4)	233 (53.4)	0.06
Ethnicity, n (%)									
Caucasian	7,440 (77.4)	664 (81.8)	0.19	4,103 (83.7)	332 (88.3)	0.18	3,337 (70.8)	332 (76.1)	0.21
Non-Caucasian	1,403 (14.6)	70 (8.6)		628 (12.8)	28 (7.4)		775 (16.4)	42 (9.6)	
Year of Cohort Entry, n (%)									
2013	1,024 (10.6)	19 (2.3)	0.34	532 (10.6)	10 (2.7)	0.33	492 (10.4)	9 (2.1)	0.35
2014	1,298 (13.5)	32 (3.9)	0.34	658 (13.4)	12 (3.2)	0.38	640 (13.6)	20 (4.6)	0.32
2015	1,561 (16.2)	78 (9.6)	0.20	749 (15.3)	37 (9.8)	0.16	812 (17.2)	41 (9.4)	0.24
2016	1,783 (18.5)	149 (18.3)	0.01	907 (18.5)	74 (19.7)	0.03	876 (18.6)	75 (17.2)	0.04
2017	2,008 (20.9)	256 (31.5)	0.23	1,029 (21.0)	120 (31.9)	0.25	979 (20.8)	136 (31.2)	0.24
2018	1,944 (20.2)	278 (7.6)	0.32	1,029 (21.0)	123 (32.7)	0.27	915 (19.4)	155 (35.6)	0.37
Diabetes duration (years)									
Mean (SD)[a]	10.8 (7.0)	11.6 (6.1)	0.12	13.2 (6.9)	13.6 (6.1)	0.06	8.4 (6.2)	9.9 (5.5)	0.25
Lifestyle and other covariates									
BMI, Mean (SD)[b]	31.0 (7.0)	33.5 (7.5)	0.35	30.8 (6.7)	33.3 (7.1)	0.37	31.2 (7.3)	33.6 (7.8)	0.32
Smoking, n (%)									
Current	3,588 (37.3)	292 (36.0)	0.03	1,830 (37.3)	128 (34.0)	0.07	1,758 (37.3)	164 (37.8)	0.00
Former	3,174 (33.0)	300 (36.9)	0.08	1,735 (35.4)	151 (40.2)	0.10	1,439 (30.5)	149 (34.2)	0.07
Never	2,699 (28.1)	213 (26.2)	0.05	1,282 (26.1)	94 (25.0)	0.03	1,417 (30.1)	119 (27.2)	0.07

Systolic Blood Pressure Level, Mean (SD)[b]	131.1 (15.8)	131.1 (14.2)	0.00	131.7 (16.4)	131.0 (14.9)	0.04	130.5 (15.1)	131.2 (13.5)	0.04
Diastolic Blood Pressure Level, Mean (SD)[b]	75.2 (10.0)	75.4 (9.4)	0.02	73.1 (9.9)	73.2 (9.5)	0.01	77.5 (9.5)	77.3 (8.9)	0.02
A1C, Mean (SD)	9.8 (2.2)	9.9 (2.0)	0.01	9.7 (2.2)	9.9 (2.0)	0.08	10.0 (2.2)	9.9 (2.0)	0.06
eGFR, Mean (SD)	66.5 (21.3)	70.9 (19.0)	0.22	57.2 (21.6)	62.5 (19.9)	0.26	76.6 (15.7)	78.2 (14.7)	0.11
Comorbidities, n (%)									
Acute kidney injury	1,612 (16.8)	93 (11.5)	0.15	1,263 (25.8)	79 (21.0)	0.11	349 (7.4)	14 (3.2)	0.19
Alcohol related disorders	1,937 (20.1)	140 (17.2)	0.07	988 (20.1)	63 (16.8)	0.09	949 (20.1)	77 (17.7)	0.06
Atrial fibrillation	782 (8.1)	59 (7.3)	0.03	687 (14.0)	50 (13.3)	0.02	95 (2.0)	9 (2.1)	0.00
Cancer[c]	2,340 (24.3)	147 (18.1)	0.15	1,444 (29.4)	91 (24.2)	0.12	896 (19.0)	56 (12.8)	0.17
Cerebrovascular disease	1,303 (13.5)	79 (9.7)	0.12	1,146 (23.4)	72 (19.1)	0.10	157 (3.3)	7 (1.6)	0.11
Chronic kidney disease	2,683 (27.9)	155 (19.1)	0.21	2,480 (50.6)	141 (37.5)	0.27	203 (4.3)	14 (3.2)	0.06
COPD	1,102 (11.5)	72 (8.9)	0.09	811 (16.5)	54 (14.4)	0.06	291 (6.2)	18 (4.1)	0.09
Coronary artery disease	2,853 (29.7)	216 (26.6)	0.07	2,606 (53.1)	204 (54.3)	0.02	247 (5.2)	12 (2.8)	0.13
Coronary revascularization	781 (8.1)	55 (6.8)	0.05	716 (14.6)	53 (14.1)	0.01	65 (1.4)	S*	-
Dialysis	152 (1.6)	15 (1.8)	0.02	96 (2.0)	12 (3.2)	0.08	56 (1.2)	S*	-
Hyperlipidemia	4,184 (43.5)	347 (42.7)	0.02	2,757 (56.2)	214 (56.9)	0.01	1,427 (30.3)	133 (30.5)	0.01
Hypertension	7,069 (73.5)	607 (74.8)	0.03	4,370 (89.1)	342 (91.0)	0.06	2,699 (57.3)	265 (60.8)	0.07
Hypoglycemia	1,011 (10.5)	89 (11.0)	0.01	624 (12.7)	51 (13.6)	0.02	387 (8.2)	38 (8.7)	0.02
Myocardial Infarction	2,228 (23.2)	165 (20.3)	0.07	2,031 (41.4)	157 (41.8)	0.01	197 (4.2)	8 (1.8)	0.14
Neuropathy	2,588 (26.9)	206 (25.4)	0.04	1,585 (32.3)	112 (29.8)	0.05	1,003 (21.3)	94 (21.6)	0.01
Peripheral vascular disease	1,174 (12.2)	92 (11.3)	0.03	1,020 (20.8)	82 (21.8)	0.02	154 (3.3)	10 (2.3)	0.06
Retinopathy	4,454 (46.3)	409 (50.4)	0.08	2,676 (54.6)	224 (59.6)	0.10	1,778 (37.7)	185 (42.4)	0.10
Stroke	988 (10.3)	71 (8.7)	0.05	896 (18.3)	70 (18.6)	0.01	92 (2.0)	S*	-
Cirrhosis	168 (1.7)	7 (0.9)	0.08	90 (1.8)	6 (1.6)	0.02	78 (1.7)	S*	-
Antidiabetic medications, n (%)									
Metformin	7,427 (77.2)	646 (79.6)	0.06	3,617 (73.8)	279 (74.2)	0.01	3,810 (80.8)	367 (84.2)	0.09
Sulfonylureas	5,889 (61.2)	388 (47.8)	0.27	3,199 (65.2)	185 (49.2)	0.33	2,690 (57.1)	203 (46.6)	0.21
Thiazolidinediones	626 (6.5)	45 (5.5)	0.04	338 (6.9)	18 (4.8)	0.09	288 (6.1)	27 (6.2)	0.00
DPP-4 inhibitors	3,996 (41.5)	259 (31.9)	0.20	2,291 (46.7)	127 (33.8)	0.27	1,705 (36.2)	132 (30.3)	0.13

GLP-1 receptor agonists	1,571 (16.3)	414 (51.0)	0.79	653 (13.3)	169 (44.9)	0.74	918 (19.5)	245 (56.2)	0.82
Alpha-glucosidase inhibitors	37 (0.4)	S*	-	23 (0.5)	S*	-	14 (0.3)	S*	-
Meglitinides	102 (1.1)	7 (0.9)	0.02	54 (1.1)	6 (1.6)	0.04	48 (1.0)	S*	-
SGLT-2 inhibitors	1,212 (12.6)	217 (26.7)	0.36	438 (8.9)	83 (22.1)	0.37	774 (16.4)	134 (30.7)	0.34
Insulin (bolus or long-acting not including insulin glargine and degludec)	3,451 (35.9)	429 (52.8)	0.35	1,668 (34.0)	219 (58.2)	0.50	1,783 (37.8)	210 (48.2)	0.21

[a] Duration of type 2 diabetes, defined as time since first diagnostic A1C, diagnostic code, or initiation of antihyperglycemic medication.

[b] Percentage of missing data presented in supplementary material, along with categorical presentation of the data.

[c] Not including non-melanoma skin cancer

S* supressed small cells with N <5

Abbreviations: A1C: glycated hemoglobin, BMI: body mass index, COPD: chronic obstructive pulmonary disorder, CPRD: Clinical Practice Research Datalink, DPP-4 inhibitors: dipeptidyl peptidase-4 inhibitors, eGFR: estimated glomerular filtration rate, GLP-1 receptor agonists: glucagon-like peptide-1 receptor agonists, SD: standard deviation, SGLT-2 inhibitors: sodium-glucose cotransporter-2 inhibitors, SMD: standardized mean difference

Table 4.2 Baseline characteristics of CPRD, DEVOTE eligible and DEVOTE ineligible populations of patients with type 2 diabetes who initiated insulin degludec or insulin glargine after inverse probability of treatment weighting and multiple imputation.

	CPRD			DEVOTE eligible			DEVOTE ineligible		
	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD
Number of patients	9,637	769		4,916	320		4,722	414	
Age (years) Mean (SD)	63.6 (14.1)	63.2 (12.7)	0.03	71.5 (10.2)	70.6 (9.4)	0.09	55.5 (12.8)	56.4 (12.2)	0.03
Males, n (%)	5,380 (55.8)	453 (58.9)	0.06	2,732 (55.6)	183 (57.0)	0.03	2,650 (56.1)	247 (59.6)	0.07
Ethnicity [b], n (%)									
Caucasian	8,139 (84.5)	665 (86.4)	0.06	4,287 (87.2)	294 (91.7)	0.15	3,851 (81.6)	346 (83.5)	0.05
Non-Caucasian	1,498 (15.5)	104 (13.6)		629 (12.8)	27 (8.3)		871 (18.4)	68 (16.5)	
Year of Cohort Entry, n (%)									
2013	962 (10.0)	112 (14.6)	0.14	504 (10.2)	35 (11.1)	0.03	459 (9.7)	61 (14.6)	0.15
2014	1,227 (12.7)	103 (13.4)	0.02	622 (12.7)	33 (10.2)	0.08	604 (12.8)	61 (14.8)	0.06
2015	1,509 (15.7)	83 (10.8)	0.14	729 (14.8)	34 (10.5)	0.13	779 (16.5)	40 (9.6)	0.20
2016	1,780 (18.5)	126 (16.4)	0.05	913 (18.6)	65 (20.2)	0.04	868 (18.4)	57 (13.8)	0.12
2017	2,090 (21.7)	172 (22.3)	0.01	1,069 (21.7)	75 (23.4)	0.04	1,022 (21.6)	97 (23.4)	0.04
2018	2,069 (21.5)	173 (22.5)	0.03	1,079 (22.0)	79 (24.6)	0.06	989 (21.0)	98 (23.7)	0.07
Diabetes duration (years)									
Mean (SD)[a]	10.9 (7.0)	10.8 (6.0)	0.01	13.2 (6.9)	13.5 (6.2)	0.05	8.5 (6.2)	9.1 (5.5)	0.09
Lifestyle and other covariates									
BMI, Mean (SD)[b]	31.2 (7.2)	31.6 (6.7)	0.07	31.0 (6.8)	31.8 (5.9)	0.14	31.3 (7.3)	31.4 (7.0)	0.01
Smoking [b], n (%)									
Current	3,640 (37.8)	290 (37.6)	0.00	1,843 (37.5)	121 (37.9)	0.01	1,801 (38.1)	152 (36.8)	0.03
Former	3,249 (33.7)	290 (37.6)	0.00	1,767 (35.9)	111 (34.6)	0.03	1,478 (31.3)	130 (31.3)	0.00
Never	2,748 (28.5)	254 (33.1)	0.01	1,306 (26.6)	88 (27.5)	0.02	1,444 (30.6)	132 (31.9)	0.03

Systolic Blood Pressure Level, Mean (SD)[b]	131.1 (15.8)	132.8 (14.1)	0.11	131.7 (16.3)	132.1 (14.1)	0.03	130.6 (15.1)	132.6 (13.6)	0.14
Diastolic Blood Pressure Level, Mean (SD)[b]	75.2 (9.9)	75.7 (9.4)	0.05	73.1 (9.9)	72.7 (9.0)	0.04	77.5 (9.5)	78.0 (9.4)	0.06
A1C, Mean (SD) [b]	9.9 (2.2)	10.0 (2.0)	0.07	9.7 (2.2)	9.8 (1.9)	0.02	10.0 (2.2)	10.2 (2.0)	0.10
eGFR, Mean (SD) [b]	67.0 (21.1)	67.2 (19.6)	0.01	57.7 (21.6)	58.8 (19.0)	0.05	76.7 (15.6)	76.6 (15.3)	0.01
Comorbidities, n (%)									
Acute kidney injury	1,573 (16.3)	124 (16.2)	0.00	1,248 (25.4)	85 (26.6)	0.03	333 (7.0)	32 (7.8)	0.03
Alcohol related disorders	1,915 (19.9)	164 (21.3)	0.04	976 (19.8)	66 (20.6)	0.02	939 (19.9)	92 (22.1)	0.05
Atrial fibrillation	776 (8.1)	58 (7.5)	0.02	685 (13.9)	44 (13.6)	0.01	95 (2.0)	8 (2.0)	0.00
Cancer[c]	2,299 (23.9)	206 (26.7)	0.07	1,428 (29.0)	105 (32.6)	0.08	875 (18.5)	86 (20.7)	0.05
Cerebrovascular disease	1,276 (13.2)	112 (14.6)	0.04	1,134 (23.1)	86 (26.9)	0.09	150 (3.2)	7 (1.6)	0.10
Chronic kidney disease	2,619 (27.2)	185 (24.0)	0.07	2,437 (49.6)	150 (46.8)	0.05	200 (4.2)	13 (3.1)	0.06
COPD	1,085 (11.3)	95 (12.4)	0.03	806 (16.4)	54 (16.8)	0.01	283 (6.0)	24 (5.7)	0.01
Coronary artery disease	2,836 (29.4)	222 (28.9)	0.01	2,616 (53.2)	178 (55.6)	0.05	237 (5.0)	13 (3.2)	0.09
Coronary revascularization	770 (8.0)	51 (6.6)	0.05	713 (14.5)	43 (13.5)	0.03	61 (1.3)	S*	-
Dialysis	154 (1.6)	11 (1.5)	0.01	101 (2.0)	6 (1.7)	0.02	54 (1.1)	S*	-
Hyperlipidemia	4,189 (43.5)	321 (41.7)	0.04	2,765 (56.3)	182 (56.8)	0.01	1,434 (30.4)	130 (31.4)	0.02
Hypertension	7,092 (73.6)	545 (70.8)	0.06	4,388 (89.3)	292 (91.0)	0.06	2,718 (57.6)	221 (53.4)	0.08
Hypoglycemia	1,019 (10.6)	82 (10.6)	0.00	626 (12.7)	42 (13.2)	0.02	396 (8.4)	42 (10.1)	0.06
Myocardial Infarction	2,210 (22.9)	169 (22.0)	0.02	2,036 (41.4)	134 (41.9)	0.01	187 (4.0)	7 (1.7)	0.14
Neuropathy	2,583 (26.8)	234 (30.4)	0.08	1,579 (32.1)	105 (32.7)	0.01	1,011 (21.4)	107 (25.9)	0.11
Peripheral vascular disease	1,167 (12.1)	107 (13.9)	0.05	1,023 (20.8)	75 (23.4)	0.06	151 (3.2)	24 (5.8)	0.13
Retinopathy	4,499 (46.7)	356 (46.3)	0.01	2,704 (55.0)	197 (61.3)	0.13	1,805 (38.2)	142 (34.2)	0.08
Stroke	979 (10.2)	100 (13.1)	0.09	901 (18.3)	83 (25.8)	0.18	85 (1.8)	S*	-
Cirrhosis	162 (1.7)	29 (3.8)	0.13	90 (1.8)	16 (5.1)	0.18	72 (1.5)	S*	-
Antidiabetic medications, n (%)									
Metformin	7,461 (77.4)	600 (78.0)	0.01	3,627 (73.8)	233 (72.8)	0.02	3,833 (81.2)	349 (84.2)	0.08
Sulfonylureas	5,786 (60.0)	435 (56.6)	0.07	3,141 (63.9)	175 (54.6)	0.19	2,645 (56.0)	236 (57.0)	0.02
Thiazolidinediones	620 (6.4)	45 (5.8)	0.03	330 (6.7)	15 (4.6)	0.09	289 (6.1)	26 (6.4)	0.01
DPP-4 inhibitors	3,920 (40.7)	296 (38.5)	0.04	2,243 (45.6)	124 (38.7)	0.14	1,680 (35.6)	152 (36.7)	0.02

GLP-1 receptor agonists	1,851 (19.2)	149 (19.4)	0.00	775 (15.8)	55 (17.3)	0.04	1,074 (22.7)	94 (22.8)	0.00
Alpha-glucosidase inhibitors	35 (0.4)	S*	-	22 (0.5)	S*	-	13 (0.3)	S*	-
Meglitinides	100 (1.0)	S*	-	55 (1.1)	S*	-	45 (0.9)	S*	-
SGLT-2 inhibitors	1,332 (13.8)	116 (15.0)	0.03	492 (10.0)	35 (11.0)	0.03	838 (17.8)	82 (19.8)	0.05
Insulin	3,609 (37.5)	377 (49.0)	0.24	1,771 (36.0)	174 (54.4)	0.38	1,840 (39.0)	173 (41.7)	0.06

[a] Duration of type 2 diabetes, defined as time since first diagnostic A1C, diagnostic code, or initiation of antihyperglycemic medication.

[b] Percentage of missing data presented in supplementary material, along with categorical presentation of the data.

[c] Not including non-melanoma skin cancer

S* supressed small cells with N <5

Abbreviations: A1C: glycated hemoglobin, BMI: body mass index, COPD: chronic obstructive pulmonary disorder, CPRD: Clinical Practice Research Datalink, DPP-4 inhibitors: dipeptidyl peptidase-4 inhibitors, eGFR: estimated glomerular filtration rate, GLP-1 receptor agonists: glucagon-like peptide-1 receptor agonists, SD: standard deviation, SGLT-2 inhibitors: sodium-glucose cotransporter-2 inhibitors, SMD: standardized mean difference

Table 4.3 The association between use of insulin degludec compared to insulin glargine and the risk of MACE among patients with type 2 diabetes in the CPRD, DEVOTE eligible and DEVOTE ineligible populations

	No. of patients	No. of events	Person-years	Incidence rate* (95% CI)	Unadjusted	Adjusted**	Observed Agreement****		
					HR (95% CI)	HR (95% CI***)	SA	EA	SD
CPRD population									
Insulin degludec	812	47	1,194	39.4 (29.6, 52.4)	0.82 (0.61, 1.10)	1.36 (0.83, 1.86)	N	N	N
Insulin glargine	9,618	910	20,001	45.5 (42.6, 48.6)	1.00 (Reference)	1.00 (Reference)			
DEVOTE eligible									
Insulin degludec	376	39	546	71.4 (52.2, 97.8)	0.86 (0.62, 1.19)	1.07 (0.63, 1.58)	N	N	Y
Insulin glargine	4,904	748	9,389	79.7 (74.2, 85.6)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	8	648	12.3 (6.2, 24.7)	0.77 (0.38, 1.57)	2.19 (0.30, 3.83)			
Insulin glargine	4,714	162	10,612	15.3 (13.1, 17.8)	1.00 (Reference)	1.00 (Reference)	N	N	N
DEVOTE Trial									
Insulin degludec	3,818	325	N/A	42.9	0.91 (0.78, 1.06)	N/A			
Insulin glargine	3,819	356	N/A	51.0	1.00 (Reference)	N/A			

*Incidence rates are expressed as events per 1,000 person-year. Confidence interval estimated using poisson model.

** The following baseline characteristics were included in the propensity score model used for inverse probability of treatment weighting: age, sex, ethnic origin, year of cohort entry, duration of diabetes, BMI, smoking status, A1C level, blood pressure level, eGFR category, comorbidities, antidiabetic drugs, comedications. Multiple imputation was applied for race, Index of Multiple Deprivation decile, smoking status, A1C, eGFR, body mass index, systolic blood pressure, diastolic blood pressure.

*** 95% CI were bootstrapped using 1000 simulations

S* supressed small cells with N <5

Abbreviations: CI: confidence interval, CPRD: Clinical Practice Research Datalink, HR: hazard ratio, MACE: major adverse cardiovascular event-composite of myocardial infarction, stroke or cardiovascular death,

****SA: Full statistical significance agreement (Y/N/P) - is said to occur if the RWD and RCT have estimates and CIs on the same side of the null; P: partial significance agreement- met the prespecified noninferiority criteria even though the database study may have indicated superiority, EA: Estimate agreement (Y/N) - is said to occur if the effect estimate of the RWD falls within the 95% confidence interval of the RCT effect estimate. SD: Standardized difference agreement is defined by standardized differences $|Z| < 1.96$ (Y = yes, N = no)

Table 4.4 The association between use of insulin degludec compared to insulin glargine for the risk of secondary outcomes including individual components of MACE, heart failure, all-cause mortality and hospitalization due to hypoglycemia, among patients with type 2 diabetes

	No. of patients	No. of events	Person-years	Incidence rate* (95% CI)	Unadjusted	Adjusted**	Observed Agreement****		
					HR (95% CI)	HR (95% CI***)	SA	EA	SD
MI									
CPRD population									
Insulin degludec	812	17	1,203	14.1 (8.8, 22.7)	1.04 (0.63, 1.69)	0.84 (0.41, 1.41)	Y	Y	Y
Insulin glargine	9,618	266	20,157	13.2 (11.7, 14.9)	1.00 (Reference)	1.00 (Reference)			
DEVOTE eligible									
Insulin degludec	376	15	551	27.2 (16.4, 45.1)	1.27 (0.75, 2.15)	0.90 (0.40, 1.56)	Y	Y	Y
Insulin glargine	4,904	198	9,517	20.8 (18.1, 23.9)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	S*	S*	2.0 (0.5, 7.9)	0.38 (0.00, 1.06)	0.29 (0.00, 0.83)	P	N	N
Insulin glargine	4,714	68	10,640	4.9 (3.9, 6.3)	1.00 (Reference)	1.00 (Reference)			
DEVOTE Trial	N/A	N/A	N/A	N/A	0.85 (0.68, 1.06)	N/A			
Stroke									
CPRD population									
Insulin degludec	812	11	1,203	9.1 (5.1, 16.5)	0.80 (0.43, 1.46)	1.55 (0.42, 3.22)	N	N	Y
Insulin glargine	9,618	218	20,264	10.8 (9.4, 12.3)	1.00 (Reference)	1.00 (Reference)			
DEVOTE eligible									
Insulin degludec	376	6	555	10.8 (4.9, 24.1)	0.52 (0.23, 1.19)	0.53 (0.12, 1.23)	Y	N	Y
Insulin glargine	4,904	188	9,573	19.6 (17.0, 22.7)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	5	648	7.7 (3.2, 18.5)	2.57 (0.99, 6.69)	7.34 (0.66, 20.43)	N	N	N
Insulin glargine	4,714	30	10,691	2.8 (2.0, 4.0)	1.00 (Reference)	1.00 (Reference)			
DEVOTE Trial	N/A	N/A	N/A	N/A	0.90 (0.65, 1.23)	N/A			
CV Death									
CPRD population									
Insulin degludec	812	29	1,212	23.9 (16.6, 34.4)	0.74 (0.51, 1.07)	1.33 (0.67, 2.13)	N	N	Y
Insulin glargine	9,618	631	20,426	30.9 (28.6, 33.4)	1.00 (Reference)	1.00 (Reference)			

DEVOTE eligible									
Insulin degludec	376	26	560	46.4 (31.6, 68.2)	0.81 (0.55, 1.20)	1.21 (0.60, 1.84)	N	Y	Y
Insulin glargine	4,904	543	9,706	55.9 (51.4, 60.9)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	S*	S*	4.6 (1.9, 14.3)	0.51 (0.16, 1.61)	1.77 (0.00, 4.97)	N	N	N
Insulin glargine	4,714	88	10,720	8.2 (6.7, 10.1)	1.00 (Reference)	1.00 (Reference)			
DEVOTE Trial	N/A	N/A	N/A	N/A	0.96 (0.76, 1.21)	N/A			
Heart Failure[‡]									
CPRD population									
Insulin degludec	812	44	1,183	37.2 (27,7, 50.0)	0.63 (0.47, 0.85)	0.79 (0.47, 1.18)	-	-	-
Insulin glargine	9,618	1,020	19,353	52.7 (49.6, 56.0)	1.00 (Reference)	1.00 (Reference)			
DEVOTE eligible									
Insulin degludec	376	41	532	77.0 (56.7, 104.6)	0.70 (0.51, 0.96)	0.84 (0.51, 1.26)	-	-	-
Insulin glargine	4,904	886	8,783	100.9 (94.5, 107.7)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	S*	S*	4.6 (1.5, 14.3)	0.35 (0.11, 1.09)	0.44 (0.00, 1.21)	-	-	-
Insulin glargine	4,714	134	10,571	12.7 (10.7, 15.0)	1.00 (Reference)	1.00 (Reference)			
All-cause Mortality									
CPRD population									
Insulin degludec	812	42	1,212	34.7 (25.6, 46.9)	0.54 (0.39, 0.73)	1.09 (0.66, 1.59)	N	Y	Y
Insulin glargine	9,618	1,217	20,426	59.6 (56.3, 63.0)	1.00 (Reference)	1.00 (Reference)			
DEVOTE eligible									
Insulin degludec	376	34	560	60.7 (43.4, 84.9)	0.61 (0.43, 0.86)	1.16 (0.70, 1.63)	N	N	Y
Insulin glargine	4,904	915	9,706	94.3 (88.6, 100.6)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	8	652	12.3 (6.1, 24.6)	0.38 (0.19, 0.77)	0.72 (0.13, 1.69)	Y	N	Y
Insulin glargine	4,714	302	10,720	28.2 (25.2, 31.5)	1.00 (Reference)	1.00 (Reference)			
DEVOTE Trial	N/A	N/A	N/A	N/A	0.91 (0.76, 1.11)	N/A			
Hospitalization due to Hypoglycemia[‡]									
CPRD population									
Insulin degludec	812	34	1,184	28.7 (20.5, 40.2)	0.96 (0.68, 1.35)	1.26 (0.72, 1.89)	-	-	-
Insulin glargine	9,618	573	19,860	28.9 (26.6, 31.3)	1.00 (Reference)	1.00 (Reference)			

DEVOTE eligible									
Insulin degludec	376	27	536	50.4 (34.5, 73.4)	1.08 (0.73, 1.59)	1.67 (0.85, 2.65)	-	-	-
Insulin glargine	4,904	424	9,309	45.5 (41.4, 50.1)	1.00 (Reference)	1.00 (Reference)			
DEVOTE ineligible									
Insulin degludec	436	7	648	10.8 (5.2, 22.7)	0.74 (0.35, 1.58)	0.52 (0.10, 1.19)	-	-	-
Insulin glargine	4,714	149	10,551	14.1 (12.0, 16.6)	1.00 (Reference)	1.00 (Reference)			

*Incidence rates are expressed as events per 1,000 person-year. Confidence interval estimated using poisson model.

** The following baseline characteristics were included in the propensity score model used for inverse probability of treatment weighting: age, sex, ethnic origin, year of cohort entry, duration of diabetes, BMI, smoking status, A1C level, blood pressure level, eGFR category, comorbidities, antidiabetic drugs, comedications. Multiple imputation was applied for race, Index of Multiple Deprivation decile, smoking status, A1C, eGFR, body mass index, systolic blood pressure, diastolic blood pressure. Hospitalization due to hypoglycemia had additional covariates of history of hypoglycemia in the year prior to cohort entry (yes/no), thyroid disease, liver cirrhosis, previous medication usages of acetaminophen, opioids, and glucagon.

*** 95% CIs were bootstrapped using 1000 simulations

S* supressed small cells with N <5

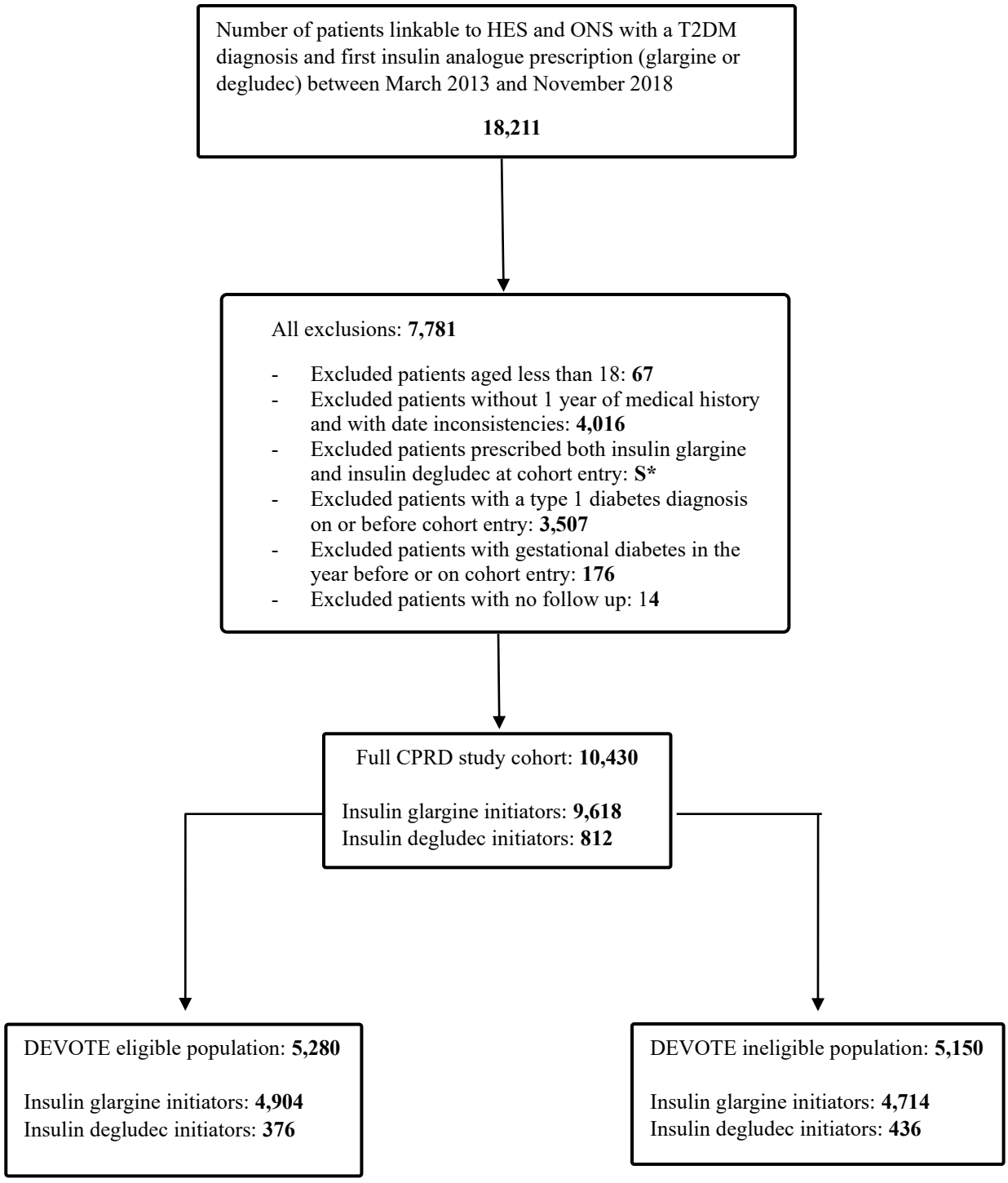
ǂ Heart failure and hospitalization due to hypoglycemia were not evaluated in the original DEVOTE trial and agreement statistics could not be calculated.

CI: confidence interval, CPRD: Clinical Practice Research Datalink, CV Death: cardiovascular death, HR: hazard ratio, MI: myocardial infarction, ****SA: Full statistical significance agreement (Y/N/P) - Is said to occur if the RWD and RCT have estimates and CIs on the same side of the null; P: partial significance agreement- met the prespecified noninferiority criteria even though the database study may have indicated superiority, EA: Estimate agreement (Y/N) - is said to occur if the effect estimate of the RWD falls within the 95% confidence interval of the RCT effect estimate. SD: Standardized difference agreement is defined by standardized differences $|Z| < 1.96$ (Y = yes, N = no)

4.11 Figure Legend

Figure 4.1 Flowchart of cohort to define sub-populations and exposures group of insulin glargine and insulin degludec initiators to examine risk of MACE in patients with type 2 diabetes.

Figure 4.1



S* supressed small cells with N <5

4.12 Supplementary Methods

4.12.1 Definition of agreement statistics:

Full statistical significance agreement is said to have occurred if the RWD and RCT have estimates and CIs on the same side of the null. *Partial significance agreement* is defined as RCT meeting the prespecified noninferiority criteria even though the CPRD study may have indicated superiority[1]. *Estimate agreement* is said to occur if the effect estimate of the RWD falls within the 95% CI of the RCT effect estimate *Standardized difference agreement* is defined by standardized differences $|Z| < 1.96$.

$$Z = \frac{\hat{\theta}_{RWE} - \hat{\theta}_{RCT}}{\sqrt{\hat{\sigma}^2_{RWE} + \hat{\sigma}^2_{RCT}}}$$ $\hat{\theta}$ are effect estimates and $\hat{\sigma}^2$ are variances.

As per the FDA, all the major CVOTs were designed for non-inferiority with an upper CI limit of 1.3[12]. For our statistical significance agreement statistic, we first assessed whether the trial was able to demonstrate non-inferiority and then evaluated whether the RWD study was able to replicate the non-inferiority finding within the same margin. If the trial achieved non-inferiority but the RWD study achieved superiority, partial statistical significance agreement was established. In addition to non-inferiority, if superiority was established in the trial, we assessed if the RWD study also found superiority to achieve full statistical significance agreement.

4.12.2 Sensitivity analyses:

We conducted 6 different sensitivity analyses. First, to compare our subpopulations to the trial ones, we combined inverse odds weights (IOW) targeting the distribution of prior MI in DEVOTE trial with our stabilized IPTW[2]. In this analysis, we reweighed our three study populations to mirror the distribution of MI in the DEVOTE trial (34%)[3]. Second, we explored the use of applying IOW to the DEVOTE eligible population to target the DEVOTE ineligible population rather than the DEVOTE trial. These IOW included all the covariates initially used in our IPTW model rather than only MI. Third, as trials usually are shorter in duration compared to observational studies, we repeated our primary analysis with follow-up time restricted to a maximum of 2 years to replicate the follow-up duration of the DEVOTE trial. Fourth, to examine the impact of exposure discontinuation and switching, we described the rate and timing of exposure discontinuation and switching. Fifth, we repeated our primary analysis using an 'on treatment' exposure definition in which we censored patients who discontinued or switched treatments. Sixth,

we matched our exposure groups based on age, year of cohort entry, and the logit of the PS with a caliper at 0.05 and re-ran our primary analysis.

4.12.3 Supplementary References:

1. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376–85.
2. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR. Transportability of Trial Results Using Inverse Odds of Sampling Weights. Am J Epidemiol. 2017 Oct 15;186(8):1010–4.
3. Marso SP, McGuire DK, Zinman B, Poulter NR, Emerson SS, Pieber TR, et al. Design of DEVOTE (Trial Comparing Cardiovascular Safety of Insulin Degludec vs Insulin Glargine in Patients With Type 2 Diabetes at High Risk of Cardiovascular Events) – DEVOTE 1. American Heart Journal. 2016 Sep 1;179:175–83.

Supplementary Table 4.1 Operationalization of DEVOTE trial inclusion criteria for observational study and number of patients for each corresponding criteria

RCT inclusion Criteria	Observational Study Operationalization of Inclusion Criteria	Number of patients
Criteria 1: Age ≥ 50 years old at cohort entry and at least 1 of the following conditions:		4,826
1. Prior myocardial infarction	1. Code for MI in CPRD or HES before cohort entry	2,280
2. Prior stroke or prior Transient Ischemic Attack (TIA)	2. Codes for stroke or TIA in CPRD or HES before cohort entry	1,188
3. Prior coronary, carotid, or peripheral arterial revascularization	3. Coronary, carotid or peripheral arterial revascularization code in CPRD or HES before cohort entry	938
4. >50% stenosis on angiography or other imaging of coronary, carotid or lower-extremity artery	4. Stenosis of coronary, carotid or lower-extremity artery codes in CPRD or HES before cohort entry	118
5. History of symptomatic coronary heart disease documented by positive exercise stress test of any cardiac imaging, or unstable angina pectoris with ECG changes	5. Coronary artery disease codes in CPRD or HES before cohort entry	2,917
6. Chronic heart failure NYHA class II-III	6. Heart failure codes in CPRD or HES before cohort entry	1,523
7. Chronic kidney disease corresponding to estimated glomerular filtration rate (eGFR) of 30-59 mL/min per 1.73m^2 per CKD-EPI	7. Codes for eGFR 30-59 mL/min per 1.73m^2 or codes for stage 3 CKD in CPRD or HES before cohort entry	2,441
Criteria 2: Age ≥ 60 years old at cohort entry and at least 1 of the following risk factors:		2,922
1. Microalbuminuria or proteinuria	1. Codes for microalbuminuria or proteinuria in CPRD or HES before cohort entry	2,102
2. Hypertension and left ventricular hypertrophy by ECG or imaging	2. Codes for hypertension or left ventricular hypertrophy in CPRD or HES before cohort entry	489
3. Left ventricular systolic and diastolic dysfunction by imaging	3. Codes for left ventricular systolic and diastolic dysfunction in CPRD or HES before cohort entry	19

4. Ankle/brachial index <0.9	4. Codes for peripheral vascular disease in CPRD or HES before cohort entry	1,023
Total participants that match Criteria 1 or 2		**5,470**
Include patients with A1C ≥ 7.0% or current insulin treatment corresponding to ≥ 20 U/d of basal insulin	A1C ≥ 7.0% or prescription of basal insulin (NPH insulin or insulin detemir) before cohort entry	5,338 (132 excluded from above)
One or more oral or injectable antidiabetic agent(s)	Prescription of metformin, sulfonylureas, thiazolidinediones, DPP-4 inhibitor, GLP-1 RA, alpha-glucosidase inhibitors, meglitinides, SGLT-2 inhibitors, insulin (except insulin glargine or insulin degludec) before cohort entry	5,280 (58 excluded from above)
Total DEVOTE eligible population		**5,280**

Abbreviations: A1C: glycated hemoglobin, CKD-EPI: Chronic Kidney Disease Epidemiology Collaboration, CPRD: Clinical Practice Research Datalink, DPP-4 inhibitors: dipeptidyl peptidase-4 inhibitors, ECG: electrocardiogram, eGFR: estimated glomerular filtration rate, GLP-1 RA: Glucagon-like peptide-1 receptor agonists, HES: Hospital Episode Statistics, U/d: units per day, MI: myocardial infarction, NYHA: New York Heart Association, SGLT-2 inhibitors: sodium-glucose cotransporter-2 inhibitors, TIA: transient ischemic stroke

Supplementary Table 4.2 Baseline characteristics of CPRD, DEVOTE eligible and DEVOTE ineligible populations of patients with type 2 diabetes who initiated insulin degludec or insulin glargine before inverse probability of treatment weighting and multiple imputation.

	CPRD			DEVOTE eligible			DEVOTE ineligible		
	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD
Number of patients	9,618	812		4,904	376		4,714	436	
Ethnicity- Missing	*775 (8.1)*	*78 (9.6)*	*0.05*	*173 (3.5)*	*16 (4.3)*	*0.04*	*602 (12.8)*	*62 (14.2)*	*0.04*
Age group, n (%)									
<49.9	1,562 (16.2)	145 (17.9)	0.04	0 (0.0)	0 (0.0)	---	1,562 (33.1)	145 (33.3)	0.00
50-59.9	2,166 (22.5)	247 (30.4)	0.18	632 (12.9)	78 (20.7)	0.21	1,534 (32.5)	169 (38.8)	0.13
60-69.9	2,299 (23.9)	222 (27.3)	0.08	1,414 (28.8)	144 (38.3)	0.20	885 (18.8)	78 (17.9)	0.02
70-79.9	2,126 (22.1)	135 (16.6)	0.14	1,593 (32.5)	99 (26.3)	0.14	533 (11.3)	36 (8.3)	0.10
80+	1,465 (15.2)	63 (7.8)	0.24	1,265 (25.8)	55 (14.6)	0.28	200 (4.2)	8 (1.8)	0.14
Index of Multiple Deprivation, n(%)									
1	1,612 (16.8)	178 (21.9)	0.13	869 (17.7)	90 (23.9)	0.15	743 (15.8)	88 (20.2)	0.12
2	1,797 (18.7)	159 (19.6)	0.02	947 (19.3)	68 (18.1)	0.03	850 (18.0)	91 (20.9)	0.07
3	1,867 (19.4)	149 (18.3)	0.03	958 (19.5)	75 (19.9)	0.01	909 (19.3)	74 (17.0)	0.06
4	2,065 (21.5)	163 (20.1)	0.03	1,014 (20.7)	72 (19.1)	0.04	1,051 (22.3)	91 (20.9)	0.03
5	2,271 (23.6)	163 (20.1)	0.09	1,112 (22.7)	71 (18.9)	0.09	1,159 (24.6)	92 (21.1)	0.08
Missing	*6 (0.1)*	*S**	-	*4 (0.1)*	*S**	-	*S**	*S**	-
Lifestyle and other covariates									
BMI									
< 25	1,756 (18.3)	92 (11.3)	0.20	890 (18.1)	48 (12.8)	0.15	866 (18.4)	44 (10.1)	0.25
25-29.9	2,843 (29.6)	183 (22.5)	0.17	1,526 (31.1)	76 (20.2)	0.26	1,317 (27.9)	107 (24.5)	0.09
30.0-34.9	2,500 (26.0)	217 (26.7)	0.01	1,286 (26.2)	97 (25.8)	0.01	1,214 (25.8)	120 (27.5)	0.03
35-39.9	1,338 (13.9)	169 (20.8)	0.18	689 (14.0)	86 (22.9)	0.23	649 (13.8)	83 (19.0)	0.13
40+	935 (9.7)	143 (17.6)	0.23	427 (8.7)	65 (17.3)	0.26	508 (10.8)	78 (17.9)	0.20
Missing	*246 (2.6)*	*8 (1.0)*	*0.12*	*86 (1.8)*	*S**	-	*160 (3.4)*	*S**	-
Smoking- Missing	*157 (1.6)*	7 (0.9)	*0.07*	*57 (1.2)*	*3 (0.8)*	*0.04*	*100 (2.1)*	*S**	-

SBP- Missing	*43 (0.4)*	*S**	-	*S**	*S**	-	*42 (0.9)*	*S**	-
DBP- Missing	*43 (0.4)*	*S**	-	*S**	*S**	-	*42 (0.9)*	*S**	-
eGFR, ml/min/1.73m2									
<60	2,897 (30.1)	166 (20.4)	0.24	2,510 (51.2)	140 (37.2)	0.28	387 (8.2)	26 (6.0)	0.10
60+	6,491 (67.5)	638 (78.6)		2,366 (48.2)	233 (62.0)		4,125 (87.5)	405 (92.9)	
Missing	*230 (2.4)*	*8 (1.0)*	*0.11*	*28 (0.6)*	*S**	-	*202 (4.3)*	*5 (1.1)*	*0.19*
A1C									
≤ 7 %	765 (8.0)	43 (5.3)	0.11	421 (8.6)	15 (4.0)	0.19	344 (7.3)	28 (6.4)	0.04
7.1-8.0%	1,186 (12.3)	96 (11.8)	0.02	706 (14.4)	46 (12.2)	0.06	480 (10.2)	50 (11.5)	0.03
> 8.0%	7,477 (77.7)	670 (82.5)	0.09	3,774 (77.0)	315 (83.8)	0.17	3,703 (78.6)	355 (81.4)	0.00
Missing	*190 (2.0)*	*S**	-	*S**	*S**	-	*187 (4.0)*	*S**	-
Comedications, n (%)									
ACE inhibitors	4,375 (45.5)	400 (49.3)	0.08	2,594 (52.9)	205 (54.5)	0.03	1,781 (37.8)	195 (44.7)	0.14
Acetylsalicylic acid	2,982 (31.0)	225 (27.7)	0.07	2,196 (44.8)	154 (41.0)	0.08	786 (16.7)	71 (16.3)	0.01
Antiplatelets	3,522 (36.6)	269 (33.1)	0.07	2,666 (54.4)	194 (51.6)	0.06	856 (18.2)	75 (17.2)	0.03
Angiotensin II receptor blockers	1,773 (18.4)	159 (19.6)	0.03	1,190 (24.3)	98 (26.1)	0.04	583 (12.4)	61 (14.0)	0.05
Beta-blockers	3,029 (31.5)	227 (28.0)	0.08	2,348 (47.9)	168 (44.7)	0.06	681 (14.4)	59 (13.5)	0.03
Calcium-channel blockers	3,005 (31.2)	245 (30.2)	0.02	1,953 (39.8)	147 (39.1)	0.01	1,052 (22.3)	98 (22.5)	0.00
Diuretics	3,260 (33.9)	247 (30.4)	0.07	2,389 (48.7)	178 (47.3)	0.03	871 (18.5)	69 (15.8)	0.07
Fibrates	286 (3.0)	23 (2.8)	0.01	173 (3.5)	7 (1.9)	0.10	113 (2.4)	16 (3.7)	0.07
Oral anticoagulants	1,029 (10.7)	72 (8.9)	0.06	885 (18.0)	60 (16.0)	0.06	144 (3.1)	12 (2.8)	0.02
Nonsteroidal anti-inflammatory drugs	2,421 (25.2)	222 (27.3)	0.05	1,238 (25.2)	111 (29.5)	0.10	1,183 (25.1)	111 (25.5)	0.01
Statins	7,309 (76.0)	642 (79.1)	0.07	4,109 (83.8)	326 (86.7)	0.08	3,200 (67.9)	316 (72.5)	0.10
Hypoglycemia specific covariates									
Thyroid disease	1,357 (14.1)	115 (14.2)	0.00	836 (17.0)	66 (17.6)	0.01	521 (11.1)	49 (11.2)	0.01
Acetaminophen	4,177 (43.4)	328 (40.4)	0.06	2,536 (51.7)	189 (50.3)	0.03	1,641 (34.8)	139 (31.9)	0.06
Opioids	3,962 (41.2)	344 (42.4)	0.02	2,228 (45.4)	181 (48.1)	0.05	1,734 (36.8)	163 (37.4)	0.01
Glucagon	78 (0.8)	7 (0.9)	0.01	32 (0.7)	5 (1.3)	0.07	46 (1.0)	S*	-

S* supressed small cells with N <5 Abbreviations: A1C: glycated hemoglobin, ACE inhibitors: angiotensin-converting enzyme inhibitors, BMI: body mass index, DBP: diastolic blood pressure, CPRD: Clinical Practice Research Datalink, eGFR: estimated glomerular filtration rate, SBP: systolic blood pressure, SD: standard deviation, SMD: standardized mean difference

Supplementary Table 4.3 Baseline characteristics of CPRD, DEVOTE eligible and DEVOTE ineligible populations of patients with type 2 diabetes who initiated insulin degludec or insulin glargine after inverse probability of treatment weighting and multiple imputation.

	CPRD			DEVOTE eligible			DEVOTE ineligible		
	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD	Insulin glargine users N/mean (%/SD)	Insulin degludec users N/mean (%/SD)	SMD
Number of patients	9,637	769		4,916	320		4,722	414	
Ethnicity- Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
Age group, n (%)									
<49.9	1,576 (16.4)	110 (14.3)	0.06	0 (0.0)	0 (0.0)	---	1,565 (33.1)	118 (28.5)	0.10
50-59.9	2,234 (23.2)	216 (28.1)	0.11	662 (13.5)	48 (14.9)	0.04	1,562 (33.1)	160 (38.6)	0.12
60-69.9	2,333 (24.2)	173 (22.5)	0.04	1,458 (29.7)	95 (29.5)	0.00	883 (18.7)	69 (16.7)	0.05
70-79.9	2,085 (21.6)	188 (24.4)	0.07	1,571 (31.9)	120 (37.5)	0.12	521 (11.0)	46 (11.1)	0.00
80+	1,408 (14.6)	82 (10.6)	0.12	1,225 (24.9)	58 (18.1)	0.17	191 (4.0)	21 (5.1)	0.05
Index of Multiple Deprivation, n(%)									
1	1,661 (17.2)	155 (20.2)	0.08	898 (18.3)	72 (22.6)	0.11	764 (16.2)	81 (19.5)	0.09
2	1,808 (18.8)	143 (18.6)	0.00	945 (19.2)	55 (17.1)	0.06	865 (18.3)	87 (20.9)	0.07
3	1,863 (19.3)	147 (19.1)	0.01	960 (19.5)	69 (21.6)	0.05	901 (19.1)	65 (15.7)	0.09
4	2,060 (21.4)	176 (22.9)	0.04	1,013 (20.6)	62 (19.5)	0.03	1,047 (22.2)	100 (24.1)	0.04
5	2,244 (23.3)	147 (19.1)	0.10	1,099 (22.4)	62 (19.3)	0.08	1,146 (24.3)	82 (19.8)	0.11
Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
Lifestyle and other covariates									
BMI									
< 25	1,753 (18.2)	113 (14.8)	0.09	887 (18.0)	45 (14.0)	0.11	858 (18.2)	63 (15.3)	0.08
25-29.9	2,855 (29.6)	235 (30.6)	0.02	1,515 (30.8)	88 (27.4)	0.08	1,366 (28.9)	138 (33.2)	0.09
30.0-34.9	2,580 (26.8)	222 (28.8)	0.05	1,314 (26.7)	102 (32.0)	0.12	1,260 (26.7)	111 (26.8)	0.00
35-39.9	1,431 (14.8)	113 (14.6)	0.01	740 (15.1)	52 (16.2)	0.03	692 (14.7)	52 (12.5)	0.06
40+	1,018 (10.6)	86 (11.2)	0.02	460 (9.4)	34 (10.5)	0.04	547 (11.6)	51 (12.2)	0.02
Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
Smoking- Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---

SBP- Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
DBP- Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
eGFR, ml/min/1.73m2									
<60	2,859 (29.7)	217 (28.2)	0.03	2,468 (50.2)	158 (49.3)	0.02	395 (8.4)	29 (7.0)	0.05
60+	6,778 (70.3)	552 (71.8)		2,448 (49.8)	163 (50.7)		4,327 (91.6)	385 (93.0)	0.00
Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
HbA1c									
≤ 7 %	758 (7.9)	40 (5.2)	0.11	404 (8.2)	17 (5.4)	0.11	361 (7.6)	22 (5.3)	0.10
7.1-8.0%	1,197 (12.4)	82 (10.7)	0.05	700 (14.2)	52 (16.3)	0.06	500 (10.6)	29 (7.1)	0.12
> 8.0%	7,681 (79.7)	647 (84.1)	0.12	3,811 (77.5)	251 (78.2)	0.02	3,861 (81.8)	363 (87.7)	0.16
Missing	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---	*0 (0.0)*	*0 (0.0)*	---
Comedications, n (%)									
ACE inhibitors	4,411 (45.8)	364 (47.4)	0.03	2,608 (53.1)	172 (53.6)	0.01	1,809 (38.3)	168 (40.4)	0.04
Acetylsalicylic acid	2,961 (30.7)	257 (33.3)	0.06	2,186 (44.5)	149 (46.3)	0.04	785 (16.6)	80 (19.3)	0.07
Antiplatelets	3,500 (36.3)	305 (39.7)	0.07	2,661 (54.1)	183 (57.2)	0.06	854 (18.1)	95 (22.8)	0.12
Angiotensin II receptor blockers	1,788 (18.6)	133 (17.3)	0.03	1,197 (24.4)	69 (21.6)	0.07	595 (12.6)	53 (12.8)	0.01
Beta-blockers	3,007 (31.2)	227 (29.5)	0.04	2,342 (47.6)	147 (46.0)	0.03	677 (14.3)	47 (11.4)	0.09
Calcium-channel blockers	2,999 (31.1)	216 (28.0)	0.07	1,953 (39.7)	123 (38.4)	0.03	1,054 (22.3)	76 (18.5)	0.10
Diuretics	3,239 (33.6)	254 (33.0)	0.01	2,388 (48.6)	161 (50.2)	0.03	864 (18.3)	69 (16.8)	0.04
Fibrates	285 (3.0)	26 (3.3)	0.02	168 (3.4)	17 (5.4)	0.10	117 (2.5)	11 (2.6)	0.01
Oral anticoagulants	1,016 (10.5)	84 (10.9)	0.01	878 (17.9)	57 (17.8)	0.00	143 (3.0)	16 (4.0)	0.05
Nonsteroidal anti-inflammatory drugs	2,442 (25.3)	186 (24.2)	0.03	1,254 (25.5)	78 (24.4)	0.03	1,189 (25.2)	100 (24.2)	0.02
Statins	7,348 (76.2)	583 (75.8)	0.01	4,129 (84.0)	267 (83.3)	0.02	3,226 (68.3)	279 (67.3)	0.02
Hypoglycemia specific covariates									
Thyroid disease	1,359 (14.1)	123 (16.0)	0.05	839 (17.1)	70 (22.0)	0.12	524 (11.1)	49 (11.9)	0.03
Acetaminophen	4,167 (43.2)	315 (41.0)	0.05	2,536 (51.6)	172 (53.6)	0.04	1,638 (34.7)	121 (29.1)	0.12
Opioids	3,977 (41.3)	336 (43.7)	0.05	2,242 (45.6)	166 (51.7)	0.12	1,739 (36.8)	152 (36.8)	0.00
Glucagon	76 (0.8)	15 (2.0)	0.10	32 (0.6)	14 (4.2)	0.23	45 (0.9)	*S**	-

Abbreviations: A1C: glycated hemoglobin, ACE inhibitors: angiotensin-converting enzyme inhibitors, BMI: body mass index, DBP: diastolic blood pressure, CPRD: Clinical Practice Research Datalink, eGFR: estimated glomerular filtration rate, SBP: systolic blood pressure, SD: standard deviation, SMD: standardized mean difference

Supplementary Table 4.4 Sensitivity analyses for the association of insulin degludec compared to insulin glargine and risk of MACE among patients with type 2 diabetes

	No. of patients	No. of events	Person-years	Incidence rate* (95% CI)	Unadjusted HR (95% CI)	Adjusted** HR (95% CI***)
2 Year Follow-up						
CPRD population						
Insulin degludec	812	38	965	39.4 (28.6, 54.1)	0.77 (0.56, 1.07)	1.36 (0.82, 2.27)
Insulin glargine	9,618	672	13,336	50.4 (46.7, 54.3)	1.00 (Reference)	1.00 (Reference)
DEVOTE eligible						
Insulin degludec	376	32	446	71.7 (50.7, 101.4)	0.83 (0.58, 1.18)	0.91 (0.57, 1.47)
Insulin glargine	4,904	557	6,468	86.1 (79.3, 93.6)	1.00 (Reference)	1.00 (Reference)
DEVOTE ineligible						
Insulin degludec	436	6	519	11.6 (5.2, 25.7)	0.68 (0.30, 1.54)	2.67 (0.90, 7.91)
Insulin glargine	4,714	115	6,868	16.7 (13.9, 20.1)	1.00 (Reference)	1.00 (Reference)
'On-treatment'[a]						
CPRD population						
Insulin degludec	812	14	462	30.3 (18.0, 51.2)	0.62 (0.36, 1.07)	1.41 (0.60, 3.29)
Insulin glargine	9,618	229	4,490	51.0 (44.8, 58.1)	1.00 (Reference)	1.00 (Reference)
DEVOTE eligible						
Insulin degludec	376	12	227	52.9 (30.0, 93.2)	0.68 (0.38, 1.22)	0.87 (0.43, 1.74)
Insulin glargine	4,904	190	2,237	84.9 (73.7, 97.9)	1.00 (Reference)	1.00 (Reference)
DEVOTE ineligible						
Insulin degludec	436	*S**	*S**	8.5 (2.1, 34.1)	0.48 (0.12, 1.99)	3.23 (1.65, 6.31)
Insulin glargine	4,714	39	2,253	17.3 (12.6, 23.7)	1.00 (Reference)	1.00 (Reference)
Propensity Score Matching[b]						
CPRD population						
Insulin degludec	769	35	1,102	31.8 (22.8, 44.2)	1.46 (0.90, 2.38)	1.18 (0.63, 2.20)
Insulin glargine	769	47	1,110	42.3 (31.8, 56.3)	1.00 (Reference)	1.00 (Reference)
DEVOTE eligible						
Insulin degludec	332	36	454	79.3 (57.2, 109.9)	1.14 (0.63, 2.05)	1.10 (0.50, 2.38)
Insulin glargine	332	37	456	81.1 (58.8, 112.0)	1.00 (Reference)	1.00 (Reference)

DEVOTE ineligible						
Insulin degludec	402	6	589	10.2 (4.6, 22.7)	1.05 (0.30, 3.61)	---****
Insulin glargine	402	6	612	9.8 (4.4, 22.7)	1.00 (Reference)	1.00 (Reference)
IOW with MI[c]						
Overall CPRD population						
Insulin degludec	812	47	1,194	39.4 (29.6, 52.4)	0.82 (0.61, 1.10)	1.35 (0.87, 2.08)
Insulin glargine	9,618	910	20,001	45.5 (42.6, 48.6)	1.00 (Reference)	1.00 (Reference)
DEVOTE eligible						
Insulin degludec	376	39	546	71.4 (52.2, 97.8)	0.86 (0.62, 1.19)	1.07 (0.69, 1.67)
Insulin glargine	4,904	748	9,389	79.7 (74.2, 85.6)	1.00 (Reference)	1.00 (Reference)
DEVOTE ineligible						
Insulin degludec	436	8	648	12.3 (6.2, 24.7)	0.77 (0.38, 1.57)	0.78 (0.28, 2.20)
Insulin glargine	4,714	162	10,612	15.3 (13.1, 17.8)	1.00 (Reference)	1.00 (Reference)
IOW with DEVOTE ineligible[d]						
DEVOTE eligible						
Insulin degludec	376	39	546	71.4 (52.2, 97.8)	0.86 (0.62, 1.19)	1.07 (0.57, 2.01)
Insulin glargine	4,904	748	9,389	79.7 (74.2, 85.6)	1.00 (Reference)	1.00 (Reference)

*Incidence rates are expressed as events per 1,000 person-year. Confidence interval estimated using Poisson model.

** The following baseline characteristics were included in the propensity score model used for inverse probability of treatment weighting: age, sex, ethnic origin, year of cohort entry, duration of diabetes, BMI, smoking status, A1C level, blood pressure level, eGFR category, comorbidities, antidiabetic drugs, comedications. Multiple imputation was applied for race, Index of Multiple Deprivation decile, smoking status, A1C, eGFR, body mass index, systolic blood pressure, diastolic blood pressure.

*** 95% CIs were estimated using robust sandwich variance estimator.

**** Estimate was unable to converge

S* supressed small cells with N <5

[a]On-treatment analysis where patients were censored for treatment discontinuation or switching

[b]Matching analysis where we matched our exposure groups based on age, year of cohort entry and the logit of the PS with a caliper at 0.05

[c]Inverse odds weighting analysis where we inverse odds weights targeting the distribution of prior MI in DEVOTE trial participants (34%) with our stabilized and truncated IPTW

[d]Inverse odds weighting to the DEVOTE eligible population to target the DEVOTE ineligible. These IOW included al the covariates initially used in our IPTW model

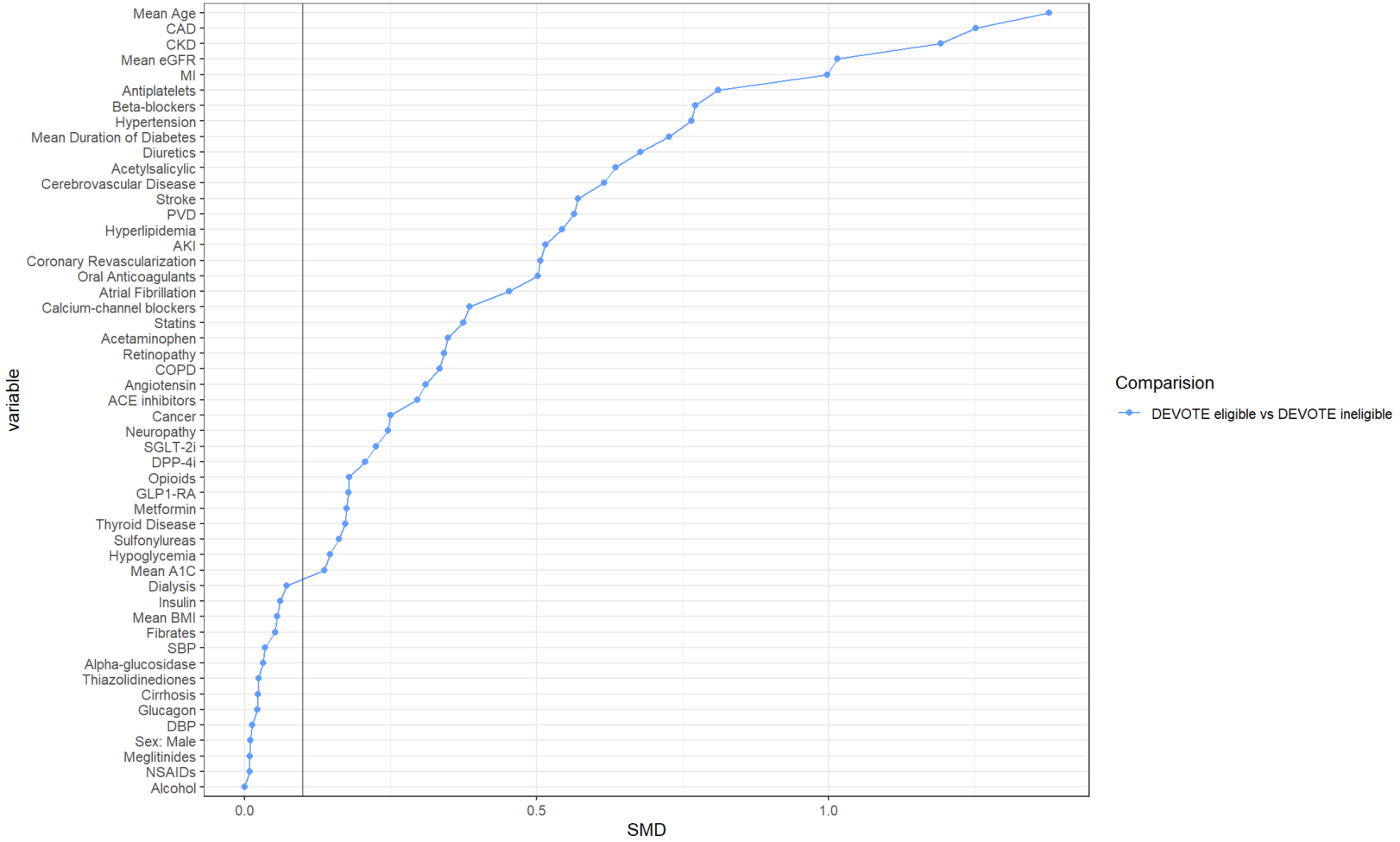
Abbreviations: CI: confidence interval, CPRD: Clinical Practice Research Datalink, IOW: inverse odds weights, IPTW: inverse probability of treatment weighting, MI: myocardial infarction

Supplementary Table 4.5 Censoring reasons among CPRD, DEVOTE eligible and DEVOTE ineligible populations

Events	No. events (N%)	Person-years	Incidence rate* (95% CI)
CPRD population			
Outcome[a]	243 (2.3%)	123	1.98 (1.74-2.24)
Switched	56 (0.5%)	41	1.37 (1.05-1.78)
Discontinuation	8281 (79.4%)	3,211	2.58 (2.52-2.63)
End of study period	1445 (13.8%)	1,385	1.04 (0.9-1.10)
Death	166 (1.6%)	43	3.82 (3.28-4.44)
End of practice	196 (1.9%)	116	1.69 (1.47-1.94)
Last data collection	43 (0.4%)	32	1.34 (0.99-1.80)
DEVOTE eligible			
Outcome[a]	202 (3.8%)	103	1.96 (1.71-2.25)
Switched	21 (0.4%)	11	1.87 (1.22-2.87)
Discontinuation	4,101 (77.7%)	1,591	2.58 (2.50-2.66)
End of study period	738 (14.0%)	662	1.11 (1.04-1.20)
Death	103 (2.0%)	31	3.28 (2.70-3.98)
End of practice	95 (1.8%)	48	1.99 (1.62-2.43)
Last data collection	20 (0.4%)	16	1.21 (0.78-1.88)
DEVOTE ineligible			
Outcome[a]	41 (0.8%)	20	2.07 (1.53-2.82)
Switched	35 (0.7%)	30	1.18 (0.84-1.64)
Discontinuation	4,180 (81.2%)	1,620	2.58 (2.50-2.66)
End of study period	707 (13.7%)	722	0.98 (0.91-1.05)
Death	63 (2.0%)	12	5.21 (4.07-6.67)
End of practice	101 (2.0%)	68	1.48 (1.22-1.80)
Last data collection	23 (0.5%)	16	1.47 (0.98-2.21)

*Incidence rates are expressed as events per 1 person-year. Interval of confidence estimated using Poisson model.
[a]Patient had a major adverse cardiovascular event of myocardial infarction, stroke or cardiovascular death
CPRD: Clinical Practice Research Datalink

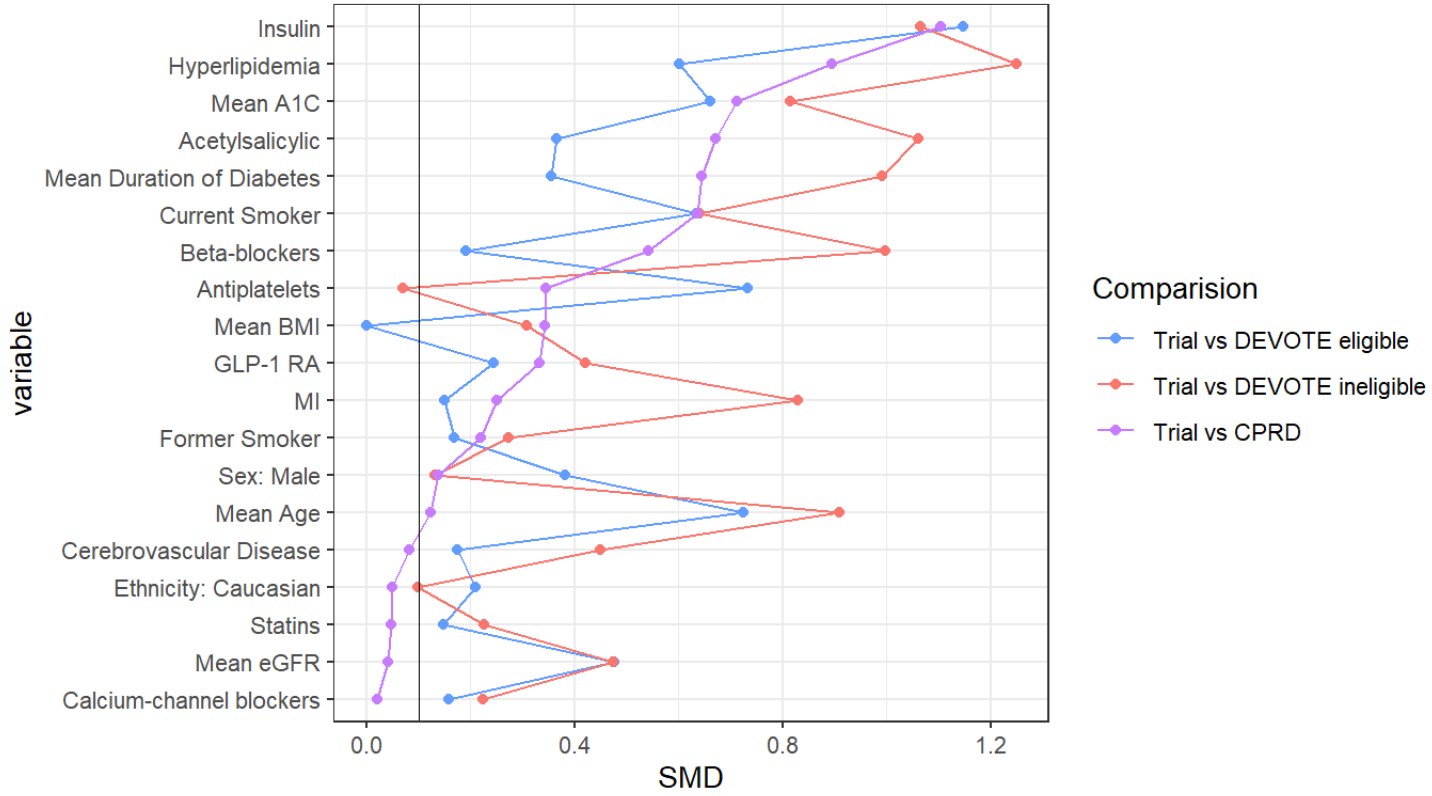
Supplementary Figure 4.1



Footnote: Graph depicts absolute standardized mean differences between DEVOTE eligible and DEVOTE ineligible total populations before weighting or imputation. Vertical line at 0.1 delineates clinically important differences

Abbreviations: A1C: glycated hemoglobin, ACE inhibitors: angiotensin-converting enzyme inhibitors, AKI: acute kidney injury, BMI: body mass index, CAD: coronary artery disease, CKD: chronic kidney disease, COPD: chronic obstructive pulmonary disease, DBP: diastolic blood pressure, DPP-4i: dipeptidyl peptidase-4 inhibitors, eGFR: estimated glomerular filtration rate, GLP-1 RA: Glucagon-like peptide-1 receptor agonists, IPTW: inverse probability of treatment weighting, MI: myocardial infarction, NSAIDs: Non-steroidal anti-inflammatory drugs, PVD: peripheral vascular disease, SBP: systolic blood pressure, SGLT-2i: sodium-glucose cotransporter-2 inhibitors, SMD: standardized mean difference

Supplementary Figure 4.2



Footnote: Graph depicts absolute standardized mean differences between DEVOTE trial and CPRD, DEVOTE eligible and DEVOTE ineligible total populations before weighting or imputation in purple, blue and red respectively. Vertical line at 0.1 delineates clinically important differences.
Abbreviations: A1C: glycated hemoglobin, BMI: body mass index, CPRD: Clinical Practice Research Datalink, eGFR: estimated glomerular filtration rate, GLP-1 RA: Glucagon-like peptide-1 receptor agonists, IPTW: inverse probability of treatment weighting, MI: myocardial infarction, SMD: standardized mean differences

CHAPTER 5: DISCUSSION

5.1 Summary

My thesis explored the emulation of CVOTs using RWD among patients with T2DM. First, we conducted a systematic review of 19 observational studies examining the generalizability of newer antidiabetic medications (GLP-1 RAs, SGLT-2 inhibitors, and DPP-4 inhibitors). We reviewed 15 cross-sectional studies which evaluated the proportion of the real-world population eligible for seminal RCTs in the area and found that eligibility percentage ranged from 7% to 59% for GLP-1 RAs and 6% to 48% for SGLT-2 inhibitor use. In addition, four cohort studies that replicated RCTs using RWD showed high levels of agreement for all drug classes we examined. However, none of these studies examined long-acting insulin analogues, which are often given later in diabetes management, especially among patients with longer duration of disease and increased disease severity.

Therefore, for my pharmacoepidemiologic database study, I emulated the DEVOTE trial by examining the risk of MACE among patients with T2DM prescribed insulin degludec compared to insulin glargine. We found that half of our CPRD population would have been eligible for the DEVOTE trial with notable differences in age, comorbidity, and comedications between the eligible and ineligible populations. In addition, HRs for MACE in our DEVOTE eligible population achieved standardized difference agreement with the original DEVOTE trial, both not indicating a clinically significant difference. The overall CPRD population and DEVOTE ineligible populations suggested consistent results for risk of MACE but were not able to achieve any level of agreement with the DEVOTE trial. For the secondary outcomes of individual components of MACE, hospitalization for heart failure, all-cause mortality, and hospitalization due to hypoglycemia, there were some variabilities in outcome rates across our overall CPRD, DEVOTE eligible and DEVOTE ineligible populations. Agreement statistics between our observational study and the original DEVOTE trial varied with the outcome and population. Sensitivity analyses addressing potential sources of bias and examining external validity for our study demonstrated consistent results. Overall, several of the studies in our systematic review found that, despite differences in population, results were generalizable. Our database study adds to the growing body of evidence that RWD and RCT can provide complementary results and emphasizes the lack of generalizability of RCTs in this area.

5.2 Emulation differences between RCTs and RWD

My thesis examines potential areas for emulation differences for RCTs using RWD, which may lead to differences in results not attributable to bias or chance. From our systematic review on the emulation of CVOTs in patients with T2DM, we found that previous studies used data sources from healthcare claims, previous cohort studies and electronic medical records. For our database study, we used CPRD data, which is a primary care database from the UK. As expected from the use of secondary data, the RWD studies' ability to replicate variable definitions for treatment, covariates, and outcomes varied from RCTs. For example, many of the studies from our systematic review did not have access to cause-of death records, and thus used all-cause mortality instead of cardiovascular death in the MACE outcome definition[143,144]. In our database study, to operationalize the inclusion criteria from the DEVOTE trial, we substituted clinical markers such as ECG measurements with diagnostic codes for diseases. These substitutions may lead to differences in the sensitivity of outcomes and inclusion criteria, leading to differences in the population, and treatment effects.

Moreover, we found that studies in this area utilized two main types of exposure definitions: ITT and on-treatment. We chose to use an ITT approach to replicate an RCT more closely and previous literature supported its use in replicating RCTs using observational data[145]. However, adherence was low in our study, with 80% of our study population discontinuing their treatment before the end of follow-up. When we conducted an 'on-treatment' analysis and censored patients who switched or discontinued their treatment, we found that our point estimate more closely resembled that of the DEVOTE trial (DEVOTE trial HR 0.91, 95% CI: 0.78, 1.06; DEVOTE eligible HR: 1.07, 95% CI: 0.63, 1.58; DEVOTE eligible censored HR: 0.87, 95% CI 0.43, 1.74). While the purpose of an ITT approach is to estimate the effect of treatment assignment, low levels of adherence can cause the effect of treatment assignment to vary substantially between trial and real-world populations.

Many of the seminal cardiovascular outcome trials used a placebo control group. To replicate these trials, studies identified in our systematic review used an active comparator of either DPP-4 inhibitors, sulfonylureas, or other second or third line antidiabetic drugs. Since diabetes medications recommends the prescription of second or third line diabetic drug classes based on individual patient goals (i.e. GLP-1 RA prescribed for those with weight loss goals), these comparisons across drug classes may be prone to confounding by indication[42]. Abrahami et al.

emulated the LEADER trial, which examined the risk of MACE for liraglutide and reported discrepant point estimates when using sulfonylureas as a control compared to second-to-third line antidiabetic drugs more broadly[146]. To mitigate the potential biases, we chose to replicate the DEVOTE trial, which used an active comparator of insulin glargine[147]. As the trial compared two different types of long-acting insulin analogues, the risk of confounding by indication and differences in patient populations between the trial and real-world control populations were reduced.

RCTs are considered the gold-standard for clinical evidence due to the randomization process, which eliminates the potential for confounding[148]. However, confounding remains an underlying bias in observational studies. Studies in our systematic review adjusted for potential confounders by using PS matching[143,144]. Abrahami et al. compared different types of PS adjustment and found heterogeneity in point estimates[146]. For our database study, we used stabilized and truncated IPTW weighting due to our imbalance in sample populations between treatment groups. Previous research has shown that in healthcare databases with small sample sizes and rare binary outcomes, some propensity score methods are more prone to bias than others[149]. When we conducted a sensitivity analysis with matching based on age, cohort entry and PS, results were consistent between IPTW and matching analyses for the CPRD overall and DEVOTE eligible populations as point estimates were on the same side of the null but varied in magnitude. However, due to the low number of events for the DEVOTE ineligible population, our results for the matching analysis were unable to converge. Given our sample population, the IPTW method allowed us to preserve more of our sample size and account for confounding in our analysis. However, it may be a source of differences in effect estimates.

Agreement statistics of statistical significance agreement, estimate agreement, and standardized difference agreement were coined by the DUPLICATE team[150]. Of the studies we evaluated in the systematic review, across the three drug classes of SGLT-2 inhibitors, GLP-1 RA and DPP-4 inhibitors, all emulation studies were able to replicate either non-inferiority findings or superiority findings from the original RCT for the risk of MACE. Seven of the nine studies were able to achieve estimate agreement where the RWD estimate was within the CI of the RCT, and all achieved standardized difference agreement. Our database study examining long-acting insulin analogues, failed to reject the null similar to the DEVOTE trial but did not achieve significance agreement with a non-inferiority upper limit threshold of 1.3 per the FDA guidelines for CVOTs[123].

Estimate agreement could not be met in any of our populations with the DEVOTE trial for estimates of MACE. Only the DEVOTE eligible population achieved standardized difference agreement with the DEVOTE trial. These discrepancies may be due to our low number of events, leading to wide CIs and low statistical power to detect differences. The emulation differences and areas of bias outlined above may have also influenced our results.

5.3 Implication of Findings

My thesis highlights the potential differences between RWD and RCTs in generating evidence on the safety and effectiveness of long-acting insulin analogues for patients with T2DM. We found that the study population is often different between the two types of evidence, with RCTs are composed of a selective population that meets specific inclusion criteria whereas RWD is more representative of the general population. Results from our database study showed that there were variations in point estimates and the width of the CI based on DEVOTE eligibility. The DEVOTE eligible population had a more precise CI with higher incidence rates for both exposure groups due to capturing older individuals with elevated CV risk. As there has been a recent push to utilize RWD in decision-making from regulatory agencies, understanding reasons for potential heterogeneity between RCT and RWD and disentangling differences in study population, emulation, and potential for biases are crucial for critically judging the quality of evidence. By using rigorous methods to analyze RWD, we can complement existing RCTs.

5.4 Limitations

The thesis has several limitations. For our systematic review, due to the limited number of studies emulating RCTs using RWD among patients with T2DM, there was heterogeneity in study design. The small number of studies with high variability in methodology and differences in exposure drug class prevented us from conducting a meta-analysis. Second, as an emerging area of research, there are currently no established guidelines on how to assess the quality of these studies for external validity. Third, the agreement statistics are only assessing the effect estimate; to our knowledge, there are no guidelines or standards on reporting or assessing methods of RCT replication studies for external validity.

For our database study, due to the small sample size, our study had a low number of events, leading to a lack of precision in our estimates. Second, even after IPTW weighting, stabilization and truncation of the weights, there were differences in covariates between exposure groups. This

may be cause for potential residual confounding. We chose to bootstrap our estimates for robustness but did not further adjust due to current literature in the area recommending against doing so in this setting[151]. Third, treatment discontinuation largely impacted our trial, and using an ITT approach may have biased our results towards the null. Our sensitivity analysis saw a shift of our effect estimate towards the DEVOTE trial.

5.5 Future Directions

Based on our literature review, there are few studies that have replicated RCTs using RWD for cardiovascular outcomes in patients with T2DM taking anti-diabetic medications. The thesis highlights how there is considerable heterogeneity in study design among current studies in the area and how these differences can influence the effect estimate. While we used the DUPLICATE team's agreement statistics[170] to assess the estimates achieved from the emulation in comparison to the RCT that was being emulated, to our knowledge, there are no guidelines or standards on reporting or assessing methodology of RCT replication studies for external validity. More research is needed to understand the best practices and reporting guideline standard for full transparency. Additional areas of future research include repeating our emulation of the DEVOTE trial using RWD with a larger sample population using different sources of data and different methodologies to examine the reproducibility and robustness of our findings. In addition, performing additional transportability analyses using patient-level trial data to better compare the impact of patient selection and effect modification on effect estimates in different study populations.

CHAPTER 6: CONCLUSION

Decisions on the safety and effectiveness of diabetic medications in patients with T2DM have been informed by cardiovascular outcome trials. However, these trials have their own shortcomings such as strict inclusion criteria that restrict the study population. Due to the lack of generalizability, there has been a recent push to assess the complementary nature of RWD and RCTs for clinical decision-making. To better understand the evidence generated by RCTs and RWD for antidiabetic medication in patients with T2DM, I first synthesized existing knowledge in the area. Our systematic review suggests that seminal cardiovascular outcome trials for the drug classes of SGLT-2 inhibitors, GLP-1 RAs, and DPP-4 inhibitors were not representative of the general population. When replicating seminal cardiovascular outcome trials using RWD, we found that the agreement between these two types of evidence were high. However, there has been no studies replicating previous cardiovascular outcome trials for long-acting insulin analogues using RWD in patients with T2DM. We conducted a cohort study emulating the DEVOTE trial to evaluate the risk of MACE in patients with T2DM taking insulin degludec compared to insulin glargine. We established a cohort of patients from the CPRD, and two subpopulations based on the eligibility criteria of the DEVOTE trial. In this study, there was no clinically relevant difference in risk across the populations for the risk of MACE, and our DEVOTE eligible subpopulation risks were compatible with the original DEVOTE trial. My thesis contributes to the growing body of evidence comparing RCTs to RWD in CVOTs among patients with T2DM. Future research should aim to replicate these findings in a highly powered population and alternative study designs. Understanding how emulation differences may correspond to differences in effect estimates allows us to better understand the complementary nature of RCTs and RWD in decision-making.

CHAPTER 7: REFERENCES

1. Diabetes Canada Clinical Practice Guidelines Expert Committee, Punthakee Z, Goldenberg R, Katz P. Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. Can J Diabetes. 2018 Apr;42 Suppl 1:S10–5.

2. Petersmann A, Müller-Wieland D, Müller UA, Landgraf R, Nauck M, Freckmann G, et al. Definition, Classification and Diagnosis of Diabetes Mellitus. Exp Clin Endocrinol Diabetes. 2019 Dec;127(S 1):S1–7.

3. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. Lancet. 2005 Apr 9;365(9467):1333–46.

4. Bartolomé A. The Pancreatic Beta Cell: Editorial. Biomolecules. 2023 Mar 8;13(3):495.

5. Galicia-Garcia U, Benito-Vicente A, Jebari S, Larrea-Sebal A, Siddiqi H, Uribe KB, et al. Pathophysiology of Type 2 Diabetes Mellitus. Int J Mol Sci. 2020 Aug 30;21(17):6275.

6. Wilcox G. Insulin and Insulin Resistance. Clin Biochem Rev. 2005 May;26(2):19–39.

7. Christensen AA, Gannon M. The Beta Cell in Type 2 Diabetes. Curr Diab Rep. 2019 Aug 9;19(9):81.

8. Chen L, Magliano DJ, Zimmet PZ. The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives. Nat Rev Endocrinol. 2012 Apr;8(4):228–36.

9. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. The Lancet. 2011 Jul 2;378(9785):31–40.

10. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends. J Epidemiol Glob Health. 2020 Mar;10(1):107–11.

11. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Research and Clinical Practice. 2010 Jan 1;87(1):4–14.

12. Hosseini Z, Whiting SJ, Vatanparast H. Type 2 diabetes prevalence among Canadian adults — dietary habits and sociodemographic risk factors. Appl Physiol Nutr Metab. 2019 Oct;44(10):1099–104.

13. Zghebi SS, Steinke DT, Carr MJ, Rutter MK, Emsley RA, Ashcroft DM. Examining trends in type 2 diabetes incidence, prevalence and mortality in the UK between 2004 and 2014. Diabetes, Obesity and Metabolism. 2017;19(11):1537–45.

14. Manson JE, Skerrett PJ, Greenland P, VanItallie TB. The Escalating Pandemics of Obesity and Sedentary Lifestyle: A Call to Action for Clinicians. Archives of Internal Medicine. 2004 Feb 9;164(3):249–58.

15. Bellou V, Belbasis L, Tzoulaki I, Evangelou E. Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses. PLoS One. 2018 Mar 20;13(3):e0194127.

16. Narayan KMV, Boyle JP, Thompson TJ, Gregg EW, Williamson DF. Effect of BMI on lifetime risk for diabetes in the U.S. Diabetes Care. 2007 Jun;30(6):1562–6.

17. Schellenberg ES, Dryden DM, Vandermeer B, Ha C, Korownyk C. Lifestyle interventions for patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. Ann Intern Med. 2013 Oct 15;159(8):543–51.

18. Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. N Engl J Med. 2001 Sep 13;345(11):790–7.

19. Steyn NP, Mann J, Bennett PH, Temple N, Zimmet P, Tuomilehto J, et al. Diet, nutrition and the prevention of type 2 diabetes. Public Health Nutrition. 2004 Feb;7(1a):147–65.

20. Helmrich SP, Ragland DR, Leung RW, Paffenbarger RS. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. N Engl J Med. 1991 Jul 18;325(3):147–52.

21. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, et al. The genetic architecture of type 2 diabetes. Nature. 2016 Aug 4;536(7614):41–7.

22. Haines L, Wan KC, Lynn R, Barrett TG, Shield JPH. Rising incidence of type 2 diabetes in children in the U.K. Diabetes Care. 2007 May;30(5):1097–101.

23. Karter AJ, Schillinger D, Adams AS, Moffet HH, Liu J, Adler NE, et al. Elevated Rates of Diabetes in Pacific Islanders and Asian Subgroups. Diabetes Care. 2013 Mar;36(3):574–9.

24. Rosella LC, Lebenbaum M, Fitzpatrick T, Zuk A, Booth GL. Prevalence of Prediabetes and Undiagnosed Diabetes in Canada (2007-2011) According to Fasting Plasma Glucose and HbA1c Screening Criteria. Diabetes Care. 2015 Jul;38(7):1299–305.

25. Santaguida PL, Balion C, Hunt D, Morrison K, Gerstein H, Raina P, et al. Diagnosis, prognosis, and treatment of impaired glucose tolerance and impaired fasting glucose. Evid Rep Technol Assess (Summ). 2005 Aug;(128):1–11.

26. Viigimaa M, Sachinidis A, Toumpourleka M, Koutsampasopoulos K, Alliksoo S, Titma T. Macrovascular Complications of Type 2 Diabetes Mellitus. CVP. 2020 Jan 27;18(2):110–6.

27. Litwak L, Goh SY, Hussein Z, Malek R, Prusty V, Khamseh ME. Prevalence of diabetes complications in people with type 2 diabetes mellitus and its association with baseline

characteristics in the multinational A1chieve study. Diabetol Metab Syndr. 2013 Oct 24;5(1):57.

28. Fowler MJ. Microvascular and Macrovascular Complications of Diabetes. Clinical Diabetes. 2008 Apr 1;26(2):77–82.

29. Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, et al. Diabetic Retinopathy. Diabetes Care. 2003 Jan 1;26(suppl_1):s99–102.

30. Gheith O, Farouk N, Nampoory N, Halim MA, Al-Otaibi T. Diabetic kidney disease: world wide difference of prevalence and risk factors. J Nephropharmacol. 2015 Oct 9;5(1):49–56.

31. Hicks CW, Selvarajah S, Mathioudakis N, Sherman RL, Hines KF, Black JH, et al. Burden of Infected Diabetic Foot Ulcers on Hospital Admissions and Costs. Ann Vasc Surg. 2016 May;33:149–58.

32. Laing SP, Swerdlow AJ, Slater SD, Burden AC, Morris A, Waugh NR, et al. Mortality from heart disease in a cohort of 23,000 patients with insulin-treated diabetes. Diabetologia. 2003 Jun 1;46(6):760–5.

33. Hogan P, Dall T, Nikolov P, American Diabetes Association. Economic costs of diabetes in the US in 2002. Diabetes Care. 2003 Mar;26(3):917–32.

34. Nwaneri C, Cooper H, Bowen-Jones D. Mortality in type 2 diabetes mellitus: magnitude of the evidence from a systematic review and meta-analysis. Diabetes & Vascular Disease. 2013 Jul 1;13(4):192–207.

35. Seshasai SRK, Kaptoge S, Thompson A, Angelantonio ED, Gao P, Sarwar N, et al. Diabetes Mellitus, Fasting Glucose, and Risk of Cause-Specific Death. The New England journal of medicine. 2011 Mar 3;364(9):829.

36. Afkarian M, Sachs MC, Kestenbaum B, Hirsch IB, Tuttle KR, Himmelfarb J, et al. Kidney Disease and Increased Mortality Risk in Type 2 Diabetes. J Am Soc Nephrol. 2013 Jan 31;24(2):302–8.

37. Zoppini G, Fedeli U, Gennaro N, Saugo M, Targher G, Bonora E. Mortality From Chronic Liver Diseases in Diabetes. Official journal of the American College of Gastroenterology | ACG. 2014 Jul;109(7):1020.

38. Zoppini G, Fedeli U, Schievano E, Dauriz M, Targher G, Bonora E, et al. Mortality from infectious diseases in diabetes. Nutrition, Metabolism and Cardiovascular Diseases. 2018 May;28(5):444–50.

39. Lin EHB, Von Korff M. Mental disorders among persons with diabetes—Results from the World Mental Health Surveys. J Psychosom Res. 2008 Dec;65(6):571–80.

40. American Diabetes Association. Standards of Care in Diabetes—2023 Abridged for Primary Care Providers. Clinical Diabetes. 2022 Dec 12;41(1):4–31.

41. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 5. Facilitating Positive Health Behaviors and Well-being to Improve Health Outcomes: Standards of Care in Diabetes—2023. Diabetes Care. 2022 Dec 12;46(Supplement_1):S68–96.

42. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Care in Diabetes—2023. Diabetes Care. 2022 Dec 12;46(Supplement_1):S140–57.

43. Lipscombe L, Booth G, Butalia S, Dasgupta K, Eurich DT, Goldenberg R, et al. Pharmacologic Glycemic Management of Type 2 Diabetes in Adults. Canadian Journal of Diabetes. 2018 Apr;42:S88–103.

44. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HAW. 10-Year Follow-up of Intensive Glucose Control in Type 2 Diabetes. New England Journal of Medicine. 2008 Oct 9;359(15):1577–89.

45. Maruthur NM, Tseng E, Hutfless S, Wilson LM, Suarez-Cuervo C, Berger Z, et al. Diabetes Medications as Monotherapy or Metformin-Based Combination Therapy for Type 2 Diabetes. Ann Intern Med. 2016 Jun 7;164(11):740–51.

46. Davies MJ, Aroda VR, Collins BS, Gabbay RA, Green J, Maruthur NM, et al. Management of Hyperglycemia in Type 2 Diabetes, 2022. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care. 2022 Nov 1;45(11):2753–86.

47. Swinnen SG, Hoekstra JB, DeVries JH. Insulin Therapy for Type 2 Diabetes. Diabetes Care. 2009 Nov;32(Suppl 2):S253–9.

48. Rick Turner J. FDA's New Guidance for Industry Addressing Evaluation of the Safety of New Drugs for Improving Glycemic Control: A Case Study in Regulatory Science. Ther Innov Regul Sci. 2021 Jan 1;55(1):1–5.

49. Nissen SE, Wolski K. Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes. New England Journal of Medicine. 2007 Jun 14;356(24):2457–71.

50. Hirshberg B, Raz I. Impact of the U.S. Food and Drug Administration Cardiovascular Assessment Requirements on the Development of Novel Antidiabetes Drugs. Diabetes Care. 2011 Apr 22;34(Supplement_2):S101–6.

51. Center for Drug Evaluation and Research. Type 2 Diabetes Mellitus: Evaluating the Safety of New Drugs for Improving Glycemic Control Guidance for Industry [Internet]. FDA; 2020 [cited 2023 Sep 8]. Available from: https://www.fda.gov/regulatory-

information/search-fda-guidance-documents/type-2-diabetes-mellitus-evaluating-safety-new-drugs-improving-glycemic-control-guidance-industry

52. European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment or prevention of diabetes mellitus [Internet]. United Kingdom; 2012 May p. 28. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-prevention-diabetes-mellitus-revision_en.pdf

53. Eckel RH, Farooki A, Henry RR, Koch GG, Leiter LA. Cardiovascular Outcome Trials in Type 2 Diabetes: What Do They Mean for Clinical Practice? Clin Diabetes. 2019 Oct;37(4):316–37.

54. Deacon CF. Dipeptidyl peptidase 4 inhibitors in the treatment of type 2 diabetes mellitus. Nat Rev Endocrinol. 2020 Nov;16(11):642–53.

55. Baggio LL, Drucker DJ. Biology of Incretins: GLP-1 and GIP. Gastroenterology. 2007 May 1;132(6):2131–57.

56. Christensen MB, Calanna S, Holst JJ, Vilsbøll T, Knop FK. Glucose-dependent Insulinotropic Polypeptide: Blood Glucose Stabilizing Effects in Patients With Type 2 Diabetes. The Journal of Clinical Endocrinology & Metabolism. 2014 Mar 1;99(3):E418–26.

57. Richter B, Bandeira-Echtler E, Bergerhoff K, Lerch C. Dipeptidyl peptidase-4 (DPP-4) inhibitors for type 2 diabetes mellitus. Cochrane Database Syst Rev. 2008 Apr 23;2008(2):CD006739.

58. Karagiannis T, Paschos P, Paletas K, Matthews DR, Tsapas A. Dipeptidyl peptidase-4 inhibitors for treatment of type 2 diabetes mellitus in the clinical setting: systematic review and meta-analysis. BMJ. 2012 Mar 12;344:e1369.

59. Wang T, McNeill AM, Chen Y, O'Neill EA, Engel SS. Characteristics of Elderly Patients Initiating Sitagliptin or Non-DPP-4-Inhibitor Oral Antihyperglycemic Agents: Analysis of a Cross-Sectional US Claims Database. Diabetes Ther. 2018 Feb;9(1):309–15.

60. White WB, Cannon CP, Heller SR, Nissen SE, Bergenstal RM, Bakris GL, et al. Alogliptin after Acute Coronary Syndrome in Patients with Type 2 Diabetes. New England Journal of Medicine. 2013 Oct 3;369(14):1327–35.

61. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B, et al. Saxagliptin and Cardiovascular Outcomes in Patients with Type 2 Diabetes Mellitus. New England Journal of Medicine. 2013 Oct 3;369(14):1317–26.

62. Green JB, Bethel MA, Armstrong PW, Buse JB, Engel SS, Garg J, et al. Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes. New England Journal of Medicine. 2015 Jul 16;373(3):232–42.

63. Rosenstock J, Kahn SE, Johansen OE, Zinman B, Espeland MA, Woerle HJ, et al. Effect of Linagliptin vs Glimepiride on Major Adverse Cardiovascular Outcomes in Patients With Type 2 Diabetes. JAMA. 2019 Sep 24;322(12):1–12.

64. Rosenstock J, Perkovic V, Johansen OE, Cooper ME, Kahn SE, Marx N, et al. Effect of Linagliptin vs Placebo on Major Cardiovascular Events in Adults With Type 2 Diabetes and High Cardiovascular and Renal Risk. JAMA. 2019 Jan 1;321(1):69–79.

65. Filion KB, Azoulay L, Platt RW, Dahl M, Dormuth CR, Clemens KK, et al. A Multicenter Observational Study of Incretin-based Drugs and Heart Failure. New England Journal of Medicine. 2016 Mar 24;374(12):1145–54.

66. Sinha B, Ghosal S. Meta-analyses of the effects of DPP-4 inhibitors, SGLT2 inhibitors and GLP1 receptor analogues on cardiovascular death, myocardial infarction, stroke and hospitalization for heart failure. Diabetes Research and Clinical Practice. 2019 Apr 1;150:8–16.

67. Cornell S. A review of GLP-1 receptor agonists in type 2 diabetes: A focus on the mechanism of action of once-weekly agents. J Clin Pharm Ther. 2020 Sep;45(S1):17–27.

68. Meier JJ. GLP-1 receptor agonists for individualized treatment of type 2 diabetes mellitus. Nat Rev Endocrinol. 2012 Dec;8(12):728–42.

69. Nauck MA, Heimesaat MM, Behle K, Holst JJ, Nauck MS, Ritzel R, et al. Effects of Glucagon-Like Peptide 1 on Counterregulatory Hormone Responses, Cognitive Functions, and Insulin Secretion during Hyperinsulinemic, Stepped Hypoglycemic Clamp Experiments in Healthy Volunteers. The Journal of Clinical Endocrinology & Metabolism. 2002 Mar 1;87(3):1239–46.

70. Nauck MA, Kleine N, Ørskov C, Holst JJ, Willms B, Creutzfeldt W. Normalization of fasting hyperglycaemia by exogenous glucagon-like peptide 1 (7-36 amide) in Type 2 (non-insulin-dependent) diabetic patients. Diabetologia. 1993 Aug 1;36(8):741–4.

71. Hare KJ, Vilsbøll T, Asmar M, Deacon CF, Knop FK, Holst JJ. The Glucagonostatic and Insulinotropic Effects of Glucagon-Like Peptide 1 Contribute Equally to Its Glucose-Lowering Action. Diabetes. 2010 Jul 1;59(7):1765–70.

72. Meier JJ, Gallwitz B, Salmen S, Goetze O, Holst JJ, Schmidt WE, et al. Normalization of Glucose Concentrations and Deceleration of Gastric Emptying after Solid Meals during Intravenous Glucagon-Like Peptide 1 in Patients with Type 2 Diabetes. The Journal of Clinical Endocrinology & Metabolism. 2003 Jun 1;88(6):2719–25.

73. Prasad-Reddy L, Isaacs D. A clinical review of GLP-1 receptor agonists: efficacy and safety in diabetes and beyond. Drugs Context. 2015 Jul 9;4:212283.

74. Kolterman OG, Buse JB, Fineman MS, Gaines E, Heintz S, Bicsak TA, et al. Synthetic Exendin-4 (Exenatide) Significantly Reduces Postprandial and Fasting Plasma Glucose in

Subjects with Type 2 Diabetes. The Journal of Clinical Endocrinology & Metabolism. 2003 Jul 1;88(7):3082–9.

75. Drucker DJ, Buse JB, Taylor K, Kendall DM, Trautmann M, Zhuang D, et al. Exenatide once weekly versus twice daily for the treatment of type 2 diabetes: a randomised, open-label, non-inferiority study. The Lancet. 2008 Oct 4;372(9645):1240–50.

76. Aroda VR, Henry RR, Han J, Huang W, DeYoung MB, Darsow T, et al. Efficacy of GLP-1 Receptor Agonists and DPP-4 Inhibitors: Meta-Analysis and Systematic Review. Clinical Therapeutics. 2012 Jun 1;34(6):1247-1258.e22.

77. Garber AJ. Long-Acting Glucagon-Like Peptide 1 Receptor Agonists. Diabetes Care. 2011 May;34(Suppl 2):S279–84.

78. Buse JB, Rosenstock J, Sesti G, Schmidt WE, Montanya E, Brett JH, et al. Liraglutide once a day versus exenatide twice a day for type 2 diabetes: a 26-week randomised, parallel-group, multinational, open-label trial (LEAD-6). Lancet. 2009 Jul 4;374(9683):39–47.

79. Pratley RE, Nauck M, Bailey T, Montanya E, Cuddihy R, Filetti S, et al. Liraglutide versus sitagliptin for patients with type 2 diabetes who did not have adequate glycaemic control with metformin: a 26-week, randomised, parallel-group, open-label trial. Lancet. 2010 Apr 24;375(9724):1447–56.

80. Nauck MA, Ratner RE, Kapitza C, Berria R, Boldrin M, Balena R. Treatment With the Human Once-Weekly Glucagon-Like Peptide-1 Analog Taspoglutide in Combination With Metformin Improves Glycemic Control and Lowers Body Weight in Patients With Type 2 Diabetes Inadequately Controlled With Metformin Alone. Diabetes Care. 2009 Jul;32(7):1237–43.

81. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. N Engl J Med. 2016 Jul 28;375(4):311–22.

82. Pfeffer MA, Claggett B, Diaz R, Dickstein K, Gerstein HC, Køber LV, et al. Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome. New England Journal of Medicine. 2015 Dec 3;373(23):2247–57.

83. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. N Engl J Med. 2016 Nov 10;375(19):1834–44.

84. Holman RR, Bethel MA, Mentz RJ, Thompson VP, Lokhnygina Y, Buse JB, et al. Effects of Once-Weekly Exenatide on Cardiovascular Outcomes in Type 2 Diabetes. N Engl J Med. 2017 Sep 28;377(13):1228–39.

85. Hernandez AF, Green JB, Janmohamed S, D'Agostino RB, Granger CB, Jones NP, et al. Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and

cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. The Lancet. 2018 Oct 27;392(10157):1519–29.

86. Gerstein HC, Colhoun HM, Dagenais GR, Diaz R, Lakshmanan M, Pais P, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. The Lancet. 2019 Jul 13;394(10193):121–30.

87. Husain M, Birkenfeld AL, Donsmark M, Dungan K, Eliaschewitz FG, Franco DR, et al. Oral Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. New England Journal of Medicine. 2019 Aug 29;381(9):841–51.

88. Kristensen SL, Rørth R, Jhund PS, Docherty KF, Sattar N, Preiss D, et al. Cardiovascular, mortality, and kidney outcomes with GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. The Lancet Diabetes & Endocrinology. 2019 Oct 1;7(10):776–85.

89. Brown GK. Glucose transporters: structure, function and consequences of deficiency. J Inherit Metab Dis. 2000 May;23(3):237–46.

90. Chao EC, Henry RR. SGLT2 inhibition — a novel strategy for diabetes treatment. Nat Rev Drug Discov. 2010 Jul;9(7):551–9.

91. Rahmoune H, Thompson PW, Ward JM, Smith CD, Hong G, Brown J. Glucose transporters in human renal proximal tubular cells isolated from the urine of patients with non-insulin-dependent diabetes. Diabetes. 2005 Dec;54(12):3427–34.

92. Nauck M. Update on developments with SGLT2 inhibitors in the management of type 2 diabetes. DDDT. 2014 Sep;1335.

93. Berhan A, Barker A. Sodium glucose co-transport 2 inhibitors in the treatment of type 2 diabetes mellitus: a meta-analysis of randomized double-blind controlled trials. BMC Endocrine Disorders. 2013 Dec 17;13(1):58.

94. Vasilakou D, Karagiannis T, Athanasiadou E, Mainou M, Liakos A, Bekiari E, et al. Sodium–Glucose Cotransporter 2 Inhibitors for Type 2 Diabetes. Ann Intern Med. 2013 Aug 20;159(4):262–74.

95. Clar C, Gill JA, Court R, Waugh N. Systematic review of SGLT2 receptor inhibitors in dual or triple therapy in type 2 diabetes. BMJ Open. 2012 Jan 1;2(5):e001007.

96. Research C for DE and. FDA revises labels of SGLT2 inhibitors for diabetes to include warnings about too much acid in the blood and serious urinary tract infections. FDA [Internet]. 2023 Aug 23 [cited 2024 Jan 12]; Available from: https://www.fda.gov/drugs/drug-safety-and-availability/fda-revises-labels-sglt2-inhibitors-diabetes-include-warnings-about-too-much-acid-blood-and-serious

97. Forxiga 5 mg film-coated tablets - Summary of Product Characteristics (SmPC) - (emc) [Internet]. [cited 2023 Sep 15]. Available from: https://www.medicines.org.uk/emc/product/2865/smpc

98. EMA. Invokana [Internet]. European Medicines Agency. 2018 [cited 2023 Sep 15]. Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/invokana

99. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. N Engl J Med. 2015 Nov 26;373(22):2117–28.

100. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. N Engl J Med. 2017 Aug 17;377(7):644–57.

101. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. New England Journal of Medicine. 2019 Jan 24;380(4):347–57.

102. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJL, Charytan DM, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. N Engl J Med. 2019 Jun 13;380(24):2295–306.

103. Cannon CP, Pratley R, Dagogo-Jack S, Mancuso J, Huyck S, Masiukiewicz U, et al. Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. New England Journal of Medicine. 2020 Oct 8;383(15):1425–35.

104. Goldenberg RM, Berard LD, Cheng AYY, Gilbert JD, Verma S, Woo VC, et al. SGLT2 Inhibitor–associated Diabetic Ketoacidosis: Clinical Review and Recommendations for Prevention and Diagnosis. Clinical Therapeutics. 2016 Dec 1;38(12):2654-2664.e1.

105. McGuire DK, Shih WJ, Cosentino F, Charbonnel B, Cherney DZI, Dagogo-Jack S, et al. Association of SGLT2 Inhibitors With Cardiovascular and Kidney Outcomes in Patients With Type 2 Diabetes: A Meta-analysis. JAMA Cardiology. 2021 Feb 1;6(2):148–58.

106. Wallia A, Molitch ME. Insulin Therapy for Type 2 Diabetes Mellitus. JAMA. 2014 Jun 11;311(22):2315.

107. Tibaldi JM. Evolution of Insulin: From Human to Analog. The American Journal of Medicine. 2014 Oct 1;127(10):S25–38.

108. Gualandi-Signorini AM, Giorgi G. Insulin formulations--a review. Eur Rev Med Pharmacol Sci. 2001;5(3):73–83.

109. DeWitt DE, Hirsch IB. Outpatient Insulin Therapy in Type 1 and Type 2 Diabetes MellitusScientific Review. JAMA. 2003 May 7;289(17):2254–64.

110. Rosenstock J, Schwartz SL, Clark CM, Park GD, Donley DW, Edwards MB. Basal Insulin Therapy in Type 2 Diabetes. DIABETES CARE. 2001;24(4).

111. Rosenstock J, Dailey G, Massi-Benedetti M, Fritsche A, Lin Z, Salzman A. Reduced Hypoglycemia Risk With Insulin Glargine: A meta-analysis comparing insulin glargine with human NPH insulin in type 2 diabetes. Diabetes Care. 2005 Apr 1;28(4):950–5.

112. Semlitsch T, Engler J, Siebenhofer A, Jeitler K, Berghold A, Horvath K. (Ultra-)long-acting insulin analogues versus NPH insulin (human isophane insulin) for adults with type 2 diabetes mellitus. Cochrane Database of Systematic Reviews [Internet]. 2020 [cited 2023 Nov 3];(11). Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD005613.pub4/full

113. Monami M, Marchionni N, Mannucci E. Long-acting insulin analogues versus NPH human insulin in type 2 diabetes: A meta-analysis. Diabetes Research and Clinical Practice. 2008 Aug 1;81(2):184–9.

114. Kalra S, Mukherjee JJ, Venkataraman S, Bantwal G, Shaikh S, Saboo B, et al. Hypoglycemia: The neglected complication. Indian J Endocrinol Metab. 2013;17(5):819–34.

115. Rys P, Wojciechowski P, Rogoz-Sitek A, Niesyczyński G, Lis J, Syta A, et al. Systematic review and meta-analysis of randomized clinical trials comparing efficacy and safety outcomes of insulin glargine with NPH insulin, premixed insulin preparations or with insulin detemir in type 2 diabetes mellitus. Acta Diabetol. 2015 Aug 1;52(4):649–62.

116. Horvath K, Jeitler K, Berghold A, Ebrahim SH, Gratzer TW, Plank J, et al. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. Cochrane Metabolic and Endocrine Disorders Group, editor. Cochrane Database of Systematic Reviews [Internet]. 2007 Apr 18 [cited 2023 Nov 3]; Available from: https://doi.wiley.com/10.1002/14651858.CD005613.pub3

117. Zhang XW, Zhang XL, Xu B, Kang LN. Comparative safety and efficacy of insulin degludec with insulin glargine in type 2 and type 1 diabetes: a meta-analysis of randomized controlled trials. Acta Diabetol. 2018 May 1;55(5):429–41.

118. Rodbard HW, Gough S, Lane W, Korsholm L, Bretler DM, Handelsman Y. Reduced Risk of Hypoglycemia with Insulin Degludec Versus Insulin Glargine in Patients with Type 2 Diabetes Requiring High Doses of Basal Insulin: A Meta-Analysis of 5 Randomized Begin Trials. Endocrine Practice. 2014 Apr 1;20(4):285–92.

119. Erpeldinger S, Rehman MB, Berkhout C, Pigache C, Zerbib Y, Regnault F, et al. Efficacy and safety of insulin in type 2 diabetes: meta-analysis of randomised controlled trials. BMC Endocrine Disorders. 2016 Jul 8;16(1):39.

120. Ivers NM, Jiang M, Alloo J, Singer A, Ngui D, Casey CG, et al. Diabetes Canada 2018 clinical practice guidelines. Can Fam Physician. 2019 Jan;65(1):14–24.

121. The ORIGIN Trial Investigators. Basal Insulin and Cardiovascular and Other Outcomes in Dysglycemia. N Engl J Med. 2012 Jul 26;367(4):319–28.

122. Marso SP, McGuire DK, Zinman B, Poulter NR, Emerson SS, Pieber TR, et al. Efficacy and Safety of Degludec versus Glargine in Type 2 Diabetes. N Engl J Med. 2017 Aug 24;377(8):723–32.

123. Regier EE, Venkat MV, Close KL. More Than 7 Years of Hindsight: Revisiting the FDA's 2008 Guidance on Cardiovascular Outcomes Trials for Type 2 Diabetes Medications. Clin Diabetes. 2016 Oct;34(4):173–80.

124. Cefalu WT, Kaul S, Gerstein HC, Holman RR, Zinman B, Skyler JS, et al. Cardiovascular Outcomes Trials in Type 2 Diabetes: Where Do We Go From Here? Reflections From a Diabetes Care Editors' Expert Forum. Diabetes Care. 2018 Jan;41(1):14–31.

125. Novo Nordisk receives Complete Response Letter in the US for Tresiba® and Ryzodeg® [Internet]. European Pharmaceutical Review. [cited 2024 Jan 18]. Available from: https://www.europeanpharmaceuticalreview.com/news/16863/novo-nordisk-receives-complete-response-letter-in-the-us-for-tresiba-and-ryzodeg/

126. Van Spall HGC, Toren A, Kiss A, Fowler RA. Eligibility Criteria of Randomized Controlled Trials Published in High-Impact General Medical JournalsA Systematic Sampling Review. JAMA. 2007 Mar 21;297(11):1233–40.

127. Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. Trials. 2015 Nov 3;16(1):495.

128. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" Lancet. 2005 Jan 1;365(9453):82–93.

129. Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues Clin Neurosci. 2011 Jun;13(2):217–24.

130. CADTH. Guidance for Reporting Real-World Evidence [Internet]. Canada: Canada's Drug and Health Technology Agency; 2023 May p. 66. (CADTH Methods and Guidelines). Available from: https://www.cadth.ca/sites/default/files/RWE/MG0020/MG0020-RWE-Guidance-Report.pdf

131. Tadrous M, Ahuja T, Ghosh B, Kropp R. Developing a Canadian Real-World Evidence Action Plan across the Drug Life Cycle. Healthc Policy. 2020 May;15(4):42–7.

132. Health Canada's position on the CADTH Guidance for Reporting RWE to Support Decision-making [Internet]. 2023 [cited 2024 Jan 22]. Available from: https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/announcements/health-canada-position-guidance-reporting-real-world-evidence-supporting-decision-making.html

133. Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products.

134. U.S. Food & Drug Administration. Framework for FDA's Real-World Evidence Program [Internet]. 2018. Available from: https://www.fda.gov/media/120060/download

135. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-World Evidence — What Is It and What Can It Tell Us? N Engl J Med. 2016 Dec 8;375(23):2293–7.

136. Eberly LA, Yang L, Eneanya ND, Essien U, Julien H, Nathan AS, et al. Association of Race/Ethnicity, Gender, and Socioeconomic Status With Sodium-Glucose Cotransporter 2 Inhibitor Use Among Patients With Diabetes in the US. JAMA Network Open. 2021 Apr 15;4(4):e216139.

137. Dennis JM, Henley WE, McGovern AP, Farmer AJ, Sattar N, Holman RR, et al. Time trends in prescribing of type 2 diabetes drugs, glycaemic response and risk factors: A retrospective analysis of primary care data, 2010–2017. Diabetes, Obesity & Metabolism. 2019 Jul;21(7):1576.

138. Falkentoft AC, Andersen J, Malik ME, Selmer C, Gæde PH, Staehr PB, et al. Impact of socioeconomic position on initiation of SGLT-2 inhibitors or GLP-1 receptor agonists in patients with type 2 diabetes – a Danish nationwide observational study. The Lancet Regional Health – Europe [Internet]. 2022 Mar 1 [cited 2024 Jan 23];14. Available from: https://www.thelancet.com/journals/lanepe/article/PIIS2666-7762(22)00001-1/fulltext

139. Overview | Type 2 diabetes in adults: management | Guidance | NICE [Internet]. NICE; 2015 [cited 2024 Jan 23]. Available from: https://www.nice.org.uk/guidance/ng28

140. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ. 2008 Jun 26;336(7659):1475–82.

141. Young KG, Hopkins R, Shields BM, Thomas NJ, McGovern AP, Dennis JM, et al. Recent UK type 2 diabetes treatment guidance represents a near whole population indication for SGLT2-inhibitor therapy. Cardiovascular Diabetology [Internet]. 2023 [cited 2024 Jan 23];22. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10623735/

142. Evidence | Type 2 diabetes in adults: management | Guidance | NICE [Internet]. NICE; 2015 [cited 2024 Jan 23]. Available from: https://www.nice.org.uk/Guidance/NG28/evidence

143. Jang HY, Kim IW, Oh JM. Using real-world data for supporting regulatory decision making: Comparison of cardiovascular and safety outcomes of an empagliflozin randomized clinical trial versus real-world data. Frontiers in Pharmacology [Internet]. 2022 [cited 2023 Jul 25];13. Available from: https://www.frontiersin.org/articles/10.3389/fphar.2022.928121

144. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies. Circulation. 2021 Mar 9;143(10):1002–13.

145. Toh S, Manson JE. An Analytic Framework for Aligning Observational and Randomized Trial Data: Application to Postmenopausal Hormone Therapy and Coronary Heart Disease. Stat Biosci. 2013 Nov;5(2):344–60.

146. Abrahami D., Pradhan R., Yin H., Honig P., Andre E.B., Azoulay L. Use of real-world data to emulate a clinical trial and support regulatory decision-making: assessing the impact of temporality, comparator choice and methods of adjustment. Clin Pharmacol Ther. 2020;((Abrahami, Pradhan, Yin, Azoulay) Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Canada).

147. Marso SP, McGuire DK, Zinman B, Poulter NR, Emerson SS, Pieber TR, et al. Design of DEVOTE (Trial Comparing Cardiovascular Safety of Insulin Degludec vs Insulin Glargine in Patients With Type 2 Diabetes at High Risk of Cardiovascular Events) – DEVOTE 1. American Heart Journal. 2016 Sep 1;179:175–83.

148. Hariton E, Locascio JJ. Randomised controlled trials—the gold standard for effectiveness research. BJOG. 2018 Dec;125(13):1716.

149. Franklin JM, Eddings W, Austin PC, Stuart EA, Schneeweiss S. Comparing the performance of propensity score methods in healthcare database studies with rare outcomes. Statistics in Medicine. 2017;36(12):1946–63.

150. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376–85.

151. Li X, Shen C. Doubly Robust Estimation of Causal Effect. Circulation: Cardiovascular Quality and Outcomes. 2020 Jan;13(1):e006065.