CARDIOVASCULAR EFFECTS OF INCRETIN-BASED ANTIHYPERGLYCEMIC DRUGS RELATIVE TO TREATMENT ALTERNATIVES IN OLDER ADULTS

Mugdha Gokhale

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2016

Approved by:

Til Stürmer

John Buse

Michele Jonsson Funk

Ross Simpson

Jennifer Lund

ABSTRACT

Mugdha Gokhale: Cardiovascular Effects Of Incretin-Based Antihyperglycemic Drugs Relative To Treatment Alternatives In Older Adults
(Under the direction of Til Stürmer)

Randomized placebo-controlled trials have examined the cardiovascular effects of dipeptidyl peptidase-4 inhibitors (DPP-4i), but limited data exist on the comparative incidence relative to therapeutic alternatives, including sulfonylureas (SU) and thiazolidinediones (TZD). In this study we therefore examined the comparative incidence of cardiovascular events with DPP-4i compared with relevant comparators using a new-user design. In recent years the use of SU was constant but DPP-4i use increased with a corresponding decrease in TZD use. Using hospitalization for heart failure (HF) as a positive control outcome we explored the use of calendar time as an instrumental variable (IV) and compared this approach to an active comparator new-user study comparing DPP-4i versus TZD.

Using 2007-2013 US Medicare claims data, we identified two new user cohort pairs – DPP-4i versus SU and DPP-4i versus TZD. Since TZDs are contraindicated in patients with HF, we further excluded patients with diagnoses of HF or related conditions for the DPP-4i versus TZD analyses. Using propensity score-weighted survival analysis methods accounting for competing risk by death, we estimated hazard ratios (HR), risk differences (RD) and 95% confidence intervals (CI) for myocardial infarction (MI), stroke, HF hospitalization, and a composite outcome (MI, stroke, or all-cause mortality). For the IV analyses, we examined the IV strength and estimated RD for HF using Kaplan-Meier curves.

The magnitude of RD per 100 patients for MI, stroke and HF hospitalization was <1 at one year after initiation with DPP-4i versus SU or TZD. The IV analysis compared patients initiating treatment during October 2010 to December 2013 versus January 2008 to May 2010 resulting in IV strength 40%. The 1- and 2-year RD of HF using the IV approach (scaled by IV-strength) and propensity score weighting were between 0 and -1 per 100 patients.

Our well-controlled population based study suggests no increased short-term CV risk with DPP-4i relative to comparators. Both IV and propensity score-weighted approaches indicate lesser risk of HF hospitalizations among DPP-4i vs TZD initiators. The use of calendar time as an IV in settings where real-world market dynamics lead to profound changes in treatments is worth consideration.

To my husband, Nikhil, for your love and support.
and
To my parents, Nitin and Shruti, for your support and encouragement.

ACKNOWLEDGEMENTS

I would first like to thank my adviser, Til Stürmer, for being a great mentor throughout my years at UNC and for providing an excellent atmosphere for learning Pharmacoepidemiology. Your guidance, motivation and example continue to immensely help me grow as a researcher. I am grateful for the opportunity of working with you for the last few years. A special thanks to John Buse for providing encouragement, your infectious enthusiasm and the many hours you invested in providing your guidance and expertise.

I would like to thank Michele Jonsson Funk, Jennifer Lund and Ross Simpson for always being available to answer questions and provide their expertise and for my research in the last few years. Thanks to Virginia Pate for her help with statistical programming and to Alison Marquis for coordinating research meetings. I am also very grateful for the support from friends and colleagues in the Pharmacoepidmiology program.

Lastly, I would like to express my deep appreciation and love for my husband, Nikhil and daughter, Myra for their love, support and patience. I would also like to express my deep gratitude to my parents, brother and grandfather for their love, encouragement during my years at UNC and beyond.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACEI	Angiotensin-Converting-Enzyme Inhibitor
AT	As-Treated Analysis
CI	Confidence Interval
CAROLINA	Cardiovascular Outcome Study of Linagliptin Versus Glimepiride in patients with Type 2 Diabetes
DPP-4i	Dipeptidyl peptidase 4 inhibitors
EXAMINE	Cardiovascular Outcomes Study of Alogliptin in Patients With Type 2 Diabetes and Acute Coronary Syndrome
FDA	Food and Drug Administration
HR	Hazard Ratio
HF	Heart failure
ITT	Intention-To-Treated Analysis
IV	Instrumental variable
IPTW	Inverse probability of treatment weight
MI	Myocardial infarction
PS	Propensity Score
RD	Risk Difference
SU	Sulfonylurea
SAVOR-TIMI	Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus (SAVOR) thrombolysis in myocardial infarction (TIMI)
SMR	Standardized Mortalityy Ratio
TZD	Thiazolidinediones
TECOS	Sitagliptin cardiovascular outcome study

CHAPTER 1. STATEMENT OF SPECIFIC AIMS

Cardiovascular disease (CVD) is a leading cause of morbidity and mortality in people with diabetes and therefore managing cardiovascular (CV) risk is important in diabetes care. [1] Three randomized trials (SAVOR-TIMI 53 for saxagliptin, EXAMINE for alogliptin and TECOS for sitagliptin) examined the CV effects of adding dipeptidyl peptidase 4 inhibitor (DPP-4i) drugs versus placebo to existing therapy but found no difference in the risk of myocardial infarction (MI), stroke and CV death with DPP-4i compared to placebo.[2-4] SAVOR-TIMI however, reported increased rate of heart failure with saxagliptin compared to placebo and there is increased interest in evaluating this further.[2, 5, 6] The evidence from observational studies examining the risk of heart failure with DPP-4i is mixed.[7-12] To date, there are no epidemiologic studies exploring the comparative cardiovascular effects of DPP-4i in an older U.S. population with a high prevalence of comorbidities and long duration of diabetes, both of which could affect the effects of DPP-4i on CVD risk. Moreover the randomized trials recruited high risk populations, largely patients with preexisting CVD. Therefore it would be of interest to study the comparative incidence of cardiovascular events among initiators of DPP-4i compared to other second-line antihyperglycemic drugs (sulfonylureas and thiazolidinediones) in a real world population.

A recent study[13] examining the market trends of antihyperglycemic drugs reported that newly approved drugs, particularly the DPP-4i class quickly gained significant market share and was the most commonly prescribed new drug class by 2012. On the other hand, the use of TZD decreased by 2012 potentially due to the safety concerns associated with this class in the last

decade. DPP-4i and TZD are generally used as second-line antihyperglycemic drugs, with intermediate (DPP-4i) to high (TZD) efficacy, low risk of hypoglycemia[14] and comparable costs before the generic form of pioglitazone was introduced in August 2012.[15] The DPP-4i and TZD drugs are therefore expected to be in therapeutic equipoise in a population excluding anyone with clear contraindications. The increase in the use of DPP-4i with a simultaneous decrease in TZD use over a short period of time also provides an opportunity to use calendar time as an instrumental variable (IV) as suggested previously in other settings.[16]

The specific aims of this project are:

Aim 1: To compare the incidence of non-fatal MI, non-fatal stroke, all-cause mortality and hospitalization for heart failure among initiators of DPP-4i versus SU and TZD using Medicare claims data 2007 – 2013.

Hypothesis: We hypothesized that DPP-4i would not be associated with an increased risk of cardiovascular outcomes relative to comparators.

Rationale: Clinical studies have signaled potential benefit on cardiovascular risk factors with DPP-4i drugs. Recently completed randomized placebo controlled trials have also not found an increased risk of MI, stroke with DPP-4i compared to placebo. Two out of the three trials have reported no increased risk of heart failure with DPP-4i.

Aim 2: Using hospitalization for heart failure (HF) as a positive control outcome, explore the use of calendar time as an instrumental variable and compare this approach to an active comparator new-user study comparing DPP-4i versus TZD (actual treatment received) in sample of Medicare beneficiaries without a history of HF or related conditions.

Hypothesis: Both the instrumental variable and treatment received approach will yield similar results of no increased risk of HF with DPP-4i versus TZD.

Rationale: DPP-4i and TZD are both second line antihyperglycemic drugs expected to be in therapeutic equipoise in a population excluding anyone with clear contraindications. The increase in the use of DPP-4i with a simultaneous decrease in TZD use over a short period of time also provides an opportunity to use calendar time as an instrumental variable in this setting. This is therefore a unique setting where both analytic approaches are expected to perform equally well.

CHAPTER 2. BACKGROUND AND REVIEW OF LITERATURE

I. Background

Type 2 diabetes is a complex metabolic disorder characterized by hyperglycemia because of insufficient insulin secretion and resistance to insulin action.[17] Cardiovascular safety of antidiabetic drugs is a question of great interest invigorated by the ongoing discussions about cardiovascular safety of thiazolidinediones. Cardiovascular (CV) disease is a leading cause of morbidity and mortality in diabetic patients and management of CV risk factors is critical in diabetes care.[1] The Food and Drug Administration (FDA) therefore requires an assessment of cardiovascular events in antihyperglycemic drug development programs.[1]

Dipeptidyl peptidase 4 inhibitors (DPP-4i) were first introduced in the U.S. in 2006. Since then, a number of new drugs in each class have been developed. GLP-1 (glucagon-like peptide 1) is one of the two incretin hormones that stimulate the release of insulin from the pancreas after food intake in a glucose-dependent manner thereby reducing the average blood glucose levels. GLP-1 is degraded within 2-3 minutes by the enzyme dipeptidyl peptidase-4 and the DPP-4i act by raising the endogenous levels of GLP-1 by inhibiting the enzyme DPP-4.[1]

II. Review of the literature:

The mechanism of DPP-4 inhibition has been shown to exert cardioprotective effects in preclinical models of cardiovascular dysfunction, and short-term studies in human subjects demonstrate modest beneficial actions on cardiac function (example, reduction of ischemic area at risk with GLP-1 infusion) in subjects with myocardial infarction.[18] DPP-4i have demonstrated improved glycemic control with beneficial effect of cardiovascular risk factors like decrease in

blood pressure, low risk of hypoglycemia and neutral (DPP-4i) weight profile suggesting potential CV benefit.[19]

Recently completed randomized controlled trials have examined the CV effects of adding DPP-4i versus placebo to existing therapy and found no evidence of increased risk of myocardial infarction (MI), stroke and CV death with DPP-4i.[2, 3, 4] All these trials met the FDA criteria for non-inferiority of these agents over placebo but found no evidence of CV risk reduction[1-3] This is unlike the observations in previous meta-analyses of about 70 clinical trials on DPP-4i agents that found reduced CV risk with DPP-4i compared to placebo or active comparators (SU, TZD or metformin).[20, 21] The saxagliptin trial (SAVOR-TIMI53) reported a 27% increased risk of heart failure hospitalizations with the addition of saxagliptin versus placebo to existing therapy, while the other two trials on alogliptin (EXAMINE) and sitagliptin (TECOS) did not find this effect.[2, 3, 4] Pre-clinical data in rats suggest that inhibition of the DPP-4 enzyme may reduce the rate of heart failure. Moreover, the effect of DPP-4i may be different in subgroups with and without heart failure. Brain natriuretic peptides are substrates for DPP-4i inhibitors and are much higher in patients with heart failure compared to patients without heart failure. Given the observations of the SAVOR-TIMI trial, the FDA has requested the clinical trial data from manufacturers of saxagliptin to explore this further.[22] Heart failure is a concern with older glucose-lowering drugs (example thiazolidinediones) and evaluating heart failure in clinical trials of antidiabetic drugs is important as recently suggested.[23] Therefore it is of interest to compare the incidence of heart failure with DPP-4i versus other agents.[24]

There have been a few observational studies examining the risk of cardiovascular outcomes with DPP-4i.[7-12, 25] While all studies report no increased relative risk of MI and stroke with DPP-4i, the evidence on heart failure is mixed with all but two studies reporting no

increased risk. All the studies mainly reported summary relative risk measures but not the absolute risk measures which may be important to put the issue in context. Further, several studies used a combined pool of non-DPP-4i drugs as the comparator making interpretation of the results difficult and less useful for physicians for making treatment decisions. To date there has not been any epidemiologic study comparing the incidence of CV events with DPP-4i versus clinically relevant active comparators in a US population of older adults with a high prevalence of comorbidity and long duration of diabetes, both of which could affect the effects of DPP-4i on CV risk. While randomized trials have had an important role in assessing the CV safety of DPP-4i, all of the trials to date have compared the addition of a DPP-4i versus adding no drug to existing therapy which may not necessarily represent real world treatment patterns which involve a lot of switching or stopping treatments. Moreover, other than the ongoing trial comparing linagliptin to glimepiride (CAROLINA trial),[26] all of the completed trials were placebo-controlled making it difficult to assess the comparative incidence of cardiovascular events relative to other second line antihyperglycemic drugs. Finally, the trials have recruited high risk populations, largely patients with a prior history of CV events.

These results may therefore be supplemented by well-designed observational studies based on real-world treatment practices that assess comparative incidence of cardiovascular events with DPP-4i. We therefore proposed to examine the comparative incidence of CV events with DPP-4i relative to other oral therapeutic alternatives SU and TZD.

In recent years, the use of DPP-4i increased with a corresponding decrease in the use of TZDs while SU use was more or less constant. According to a recent analysis [13] of the IMS health data examining the market trends of antihyperglycemic drugs, the newly approved drugs, particularly the DPP-4i class quickly gained significant market share and was the most

commonly prescribed new drug class by 2012. On the other hand, the use of thiazolidinediones decreased by 2012 with the use of pioglitazone reaching only half its peak use in 2008 and negligible use of rosiglitazone. DPP-4i and TZD are both second-line antihyperglycemic drugs, with intermediate (DPP-4i) to high (TZD) efficacy, low risk of hypoglycemia [14] and comparable costs before the generic form of pioglitazone was introduced in August 2012.[15]

The DPP-4i and TZD drugs are therefore expected to be in therapeutic equipoise in a population excluding anyone with clear contraindications. The increase in the use of DPP-4i with a simultaneous decrease in TZD use over a short period of time also provides an opportunity to use calendar time as an instrumental variable (IV) as suggested previously in other settings.[16] According to our knowledge, this is a first unique setting where there exists both an active comparator and also a potential IV driven by the antihyperglycemic drug market which enables comparison of the two methods. Using hospitalization for heart failure (HF) as a positive control outcome we explored the use of calendar time as an IV and compared this approach to an active comparator new-user study comparing DPP-4i versus TZD.

CHAPTER 3. METHODS

I. Study Designs, Exposure and Comparisons

For aim 1, we used an active comparator, new-user cohort study design. New-user design helps to avoid the biases inherent in the prevalent user design. This design is analogous to a head-to-head clinical trial (except that drugs are not assigned by randomization)[27] and mimics the most relevant clinical decision ('which treatment to initiate' rather than 'treatment or not'). It allows synchronizing follow-up in all cohorts which is the basis for sensitivity analyses of induction periods.

We compared DPP-4i with SU and TZD in separate analyses. Metformin is the first-line pharmacotherapy for type 2 diabetes. In case of insufficient control of blood glucose levels, one of the second-line agents in the table below is added to the treatment regimen.[14] Comparing DPP-4i to other antidiabetic drugs that are likely to be indicated at a similar stage of diabetes progression and severity ensures balance in most confounders at baseline.

For aim 2, we compared two analytic approaches – 'actual treatment received' i.e., DPP-4i versus TZD and analyses using calendar time as an instrumental variable. Exposure was for the former approach was initiation of either DPP-4i or TZD. For the later, an 'optimal' IV was defined using drug initiation curves and substantive information that included FDA drug safety communications about the TZDs.

II. Data Source

For both aim 1 and aim 2, we will use a 20% random sample of Medicare beneficiaries >65 years with fee-for-service Part A (hospital coverage), B (outpatient medical care) and D

(dispensed prescription drugs) enrollment in at least one month during a calendar year from January 1, 2007 (2006 for Part A and B) to December 31, 2013. This data contains information about demographics characteristics, plan enrollment, diagnoses, procedures and prescription drugs for each enrollee. Data is recorded by the *International Classification of Diseases, Ninth Revision* (ICD-9), Current Procedural Terminology-4 (CPT-4) codes, and the Healthcare Common Procedure Coding System (HCPCS). Dispensed prescriptions are coded using National Drug Codes (NDCs).

III. Study cohort, eligibility criteria

For aim 1, we identified two new-user cohort pairs – DPP-4i versus SU (not exposed to either DPP-4i or SU in the previous 6 months) and DPP-4i versus TZD (not exposed to either DPP-4i or SU in the previous 6 months). Patients could be on any antihyperglycemic drugs other than the drugs compared in the 6 months pre-initiation. For example, in the DPP-4i versus SU comparison, patients could be on any drugs except DPP-4i and SU during the washout period. A DPP-4i initiator with no previous prescription of SU and TZD would be eligible for inclusion in both comparisons. Initiation was defined as the first prescription of the drug with the index date defined as the date of dispensing. Patients were required to have at least 6 months of continuous Part D enrollment and at least 12 months parts A and B enrollment pre-index. To reduce the potential for bias toward the null due to secondary nonadherence, we restricted our cohorts to patients with a second prescription for the same drug class dispensed within 6 months after the index date and follow-up started from the second fill date. Since TZDs are contraindicated in patients with existing heart failure (which can lead to intractable confounding by contraindication if not properly accounted for), for all analyses comparing DPP-4i versus TZD we further excluded patients with diagnoses of heart failure, hypertensive disease with heart

failure and closely related conditions of cardiomyopathy, arrhythmias, chronic kidney disease, edema and use of loop diuretics.

For aim 2, for the 'treatment initiated' approach, we identified new-users of DPP-4i or TZD as described above from a population without previous diagnoses of heart failure and related conditions. Next we defined a cohort of new-users of DPP-4i and TZD for our IV analysis. An IV is an observed variable associated with the variation in exposure like randomized assignment in a clinical trial.[28] To be valid, an IV 1) should be associated with the treatment, 2) should be unrelated to patient characteristics and 3) should be related to the outcome only through its association with the treatment (exclusion restriction). Additionally, in order to estimate the average treatment effect in the 'compliers' an additional assumption of 'monotonicity' (no 'defiers') should hold. Monotonicity requires that the instrument affects the treatment deterministically in one direction.[28]

We defined the IV as a binary variable (post versus pre-period) anchored around the time of the cross-over of the drug initiation curves (June – September 2010) which was also the time when the FDA issued communications about restriction of rosiglitazone and concerns about the safety of pioglitazone. In our study, 'compliers' are patients whose initial treatment is determined by the calendar time in which they initiated treatment – i.e, TZD if initiating before the crossover of the drug initiation curves and DPP-4i if initiating after the crossover. We identified an 'optimal' IV by evaluating the strength of the instrument's effect on the received treatment.[16] We examined the instrument in relation to IV assumptions using measured covariates, falsification tests and expert knowledge. The binary measure of calendar time was decided before examining the effect estimates.

IV. Outcomes

Aim 1 –

The outcomes for this aim 1 were non-fatal MI, stroke, hospitalization for heart failure (HF) and all-cause mortality and a composite CV outcome of non-fatal MI, stroke and all-cause mortality based on the outcome definition in the RCTs. Medicare claims data do not include information on causes of death and we could not directly identify cardiovascular death. However since cardiovascular deaths are responsible for more than half of the deaths in patients with diabetes, we used all-cause mortality a proxy in this study.[29] MI was defined using International Classification of Diseases, Ninth Revision (ICD-9) code 410 in the first or second position of the inpatient claims, a definition with a positive predictive value of 94% in a Medicare population.[30] Stroke was defined using ICD-9 codes 430, 431, 433.x1, 434.x1, and 436, located in the first position only (specificity 95–97%, sensitivity 74–90%).[31] HF hospitalization was defined using ICD-9 code 428.xx in the primary position only which has a specificity >98% but a very low sensitivity of 21% in a Medicare population.[32]

Aim 2 –

For aim 2, the outcome was hospitalization for HF[33] defined using ICD-9 code 428.xx in the primary position in inpatient claims as for aim 1. TZDs are known to be associated with an increased risk of heart failure.[34] While recently reported placebo-controlled SAVOR-TIMI trial reported a 27% increased risk of heart failure with saxagliptin, the trials of other DPP-4i (EXAMINE and TECOS) indicate no increased incidence of HF with DPP-4i drugs sitagliptin and alogliptin (which is >80% of our cohort).[2, 3, 35] HF hospitalization is therefore a positive control outcome in our study enabling the comparison of performance of the propensity score weighting method comparing the 'actual treatments initiated' and IV analysis comparing levels

of the calendar time instrument. Patients were followed up from the second fill date till the earliest of the following – outcome (HF hospitalization), death, end of enrollment or 2 years after the index date.

V. Follow-up

The primary analysis approach used for aim 1 was 'as-treated'. Patients were followed from the second prescription until the earliest of: the outcome of interest, discontinuation (no new prescription for the initiated drug, within days-supply plus a 180 days grace period to allow for dose adjustment/irregular use), switching to or augmentation with the comparator drug, non-end point event (example, stroke would be a non-end point event in the analysis of MI), end of enrollment, or December 31, 2013.

For aim 2, we used an intent-to-treat approach. Patients were followed up from the second prescription date till the earliest of the following – outcome (HF hospitalization), death, end of enrollment or 2 years after the index date.

Confounding control and analysis:

For aim 1, we used propensity scores (PS) to control for measured confounding. Using baseline variables (comorbidities, medication use, health care utilization) measured before initiation, we predicted the probability for initiating DPP-4i versus SU and DPP-4i versus TZD for each patient (the PS) using two separate logistic regression models.[36] We then assigned a weight of 1 to the treated (DPP-4i) and a weight of (PS/(1-PS)) to SU and TZD. Such weighting creates pseudo-populations of SU and TZD initiators with similar covariate distribution as in DPP-4i initiators. This covariate balance across groups allows us to estimate the unconfounded treatment effect in a population of patients similar to those actually initiating DPP-4i.[37, 38] Our

weighted analysis thus answers the question "what would have happened to patients who initiated DPP-4i if they had initiated SU or TZD, instead".

Competing risks arise when the occurrence of one event precludes the occurrence of other events. In our study of older adults, mortality is a competing event and censoring patients who die and analyzing the data using standard Cox proportional hazards models yields biased estimates because this type of censoring may be 'informative'.[39] We therefore used weighted cumulative incidence curves accounting for competing risk by death to estimate the risk, risk differences (RD) and risk ratios (RR) for non-fatal MI, stroke and hospitalization for HF among initiators of DPP-4i versus comparators.[39] We obtained confidence intervals by bootstrapping 1000 replicates. We analyzed the composite outcome of non-fatal MI, stroke or all-cause mortality using traditional weighted Cox models.

Analyses were repeated in pre-specified subgroups based on history of CV disease (diagnoses of MI, stroke, angina, other ischemic disease, atherosclerosis for the DPP-4i versus TZD comparison and the above diagnoses plus heart failure related diagnoses for the DPP-4i versus SU comparison).

Under the 'treatment received' approach, we used inverse probability of treatment (IPTW) weighting using PS to control for confounding. PS for DPP-4i versus TZD was estimated using baseline covariates without including the year of initiation in the model as it can increase the variability of effect estimates and can increase bias in case of unmeasured confounding. [36, 40, 41] We then assigned a weight of 1/PS to DPP-4i and a weight of (1/(1-PS)) to TZD and stabilized both groups by marginal prevalence of the treatment actually received.[42] Such inverse probability of treatment (IPTW) weighting creates a pseudo-population in which the association between covariates and treatment is removed by the weighting described above. Under the assumption of

no unmeasured confounding this method allows us to contrast two scenarios: "what would have happened if the entire population initiated DPP-4i' versus '"what would have happened if the entire population initiated TZD'.[42] Using IPTW weighted Kaplan Meier curves we then estimated the RDs for HF hospitalization using treatment as the strata variable.

The effect of the IV levels on HF hospitalization was observed by the IV estimates of RD scaled by the IV strength to estimate the average treatment effect among the 'compliers'. Next we generated covariate-adjusted IV estimates of the RD using weighted Kaplan-Meier methods analogous to the IPTW weighted methods described above and scaled this by the strength of the instrument. Under this approach, we first predicted the probability (PP) of the instrument as a function of baseline covariates and assigned a weight of 1/PP to the 'post' period and 1/(1-PP) to the 'pre' period and stabilized the weights by the marginal prevalence of the instrument.

VI. Sensitivity Analysis

Aim 1 –

The following sensitivity analyses were performed. First, to increase the probability of cardiovascular cause of death, we excluded the deaths of patients with codes for metastatic cancer anytime during follow-up. Second, since SU are older drugs and to ensure that patients were indeed treated with SU as second line drugs (rather than first line therapy), we repeated the DPP-4i versus SU analyses requiring patients to have at least one metformin prescription in the 6 months before the index date. Finally we repeated all analyses using an intent-to-treat approach where patients were not censored for treatment changes but followed from the second prescription to the earliest of the outcome, non-event end point, end of enrollment, or December 31, 2013.

Aim 2 –

The following sensitivity analyses were performed for aim 2. First, propensity score models used the full study population and we performed sensitivity analyses in the reduced instrumental variable cohort to evaluate selection differences that may have been introduced based on instrumental variable exclusions. Next we performed a sensitivity analysis restricting the study cohort to patients who initiated DPP-4i or TZD between July 2007 and December 2011, with follow-up allowed through December 2013. This approach excluded anyone who initiated therapy during or after January 2012, who would not have the potential for a 2-year follow-up before the end of the study (and therefore potentially violate the assumption of 'non-informative' censoring i.e., the assumption that patients who drop out of the study should do so for reasons unrelated to the study).[43] Finally we performed additional analyses with different definitions of the instrument (IV defined using a cut-point and by comparing periods with maximum separation; described in supplemental figures 2 and 3).

All analyses were performed with the SAS software, 9.3 version.

CHAPTER 4. COMPARITIVE INCIDENCE OF CARDIOVASCULAR EVENTS IN OLDER ADULTS INITIATING DIPEPTIDYL PEPTIDASE 4 INHIBITORS VERSUS THERAPEUTIC ALTERNATIVES

I. Introduction

In the United States over 25% of the population 65 years or older has diabetes.[44] Cardiovascular (CV) disease is a leading cause of morbidity and mortality in diabetes patients, with the risk increasing with age.[45] While improved glycemic control by antihyperglycemic drugs reduces microvascular complications, uncertainty remains regarding risk reduction for CV events. International agencies now require a thorough assessment of CV risk in antihyperglycemic drug development programs.[46,47]

The dipeptidyl peptidase-4 inhibitors (DPP-4i) are relatively new antihyperglycemic drugs that were incorporated into diabetes treatment algorithms as second line therapy since 2011. These drugs have good tolerability, low risk of hypoglycemia and are weight neutral compared to other second line drugs.[14] Three randomized placebo-controlled trials (RCT) have recently evaluated the CV safety of DPP-4i (saxagliptin, alogliptin and sitagliptin) in high-risk patients with type 2 diabetes.[2-4] All RCTs found no increase in the risk of non-fatal myocardial infarction, stroke, CV death with adding a DPP-4i agent versus placebo to existing therapy. However, the saxaglipin trial found an increased risk of hospitalization for heart failure, whereas the other two trials did not find any association between DPP-4i treatment and heart failure.[2]

While randomized trials have had an important role in assessing the CV safety of DPP-4i, all the trials to date have compared the addition of a DPP-4i versus adding no drug to existing therapy which may not represent real world treatment patterns which involve a lot of switching

or stopping treatments. Moreover, other than the ongoing trial comparing linagliptin to glimepiride (CAROLINA),[26] all the completed trials were placebo-controlled making it difficult to assess the comparative incidence of cardiovascular events relative to therapeutic alternatives. Finally, the trials have recruited high risk populations, largely patients with a prior history of CV events.

Observational studies examining CV risk with DPP-4i report no increased relative risk of myocardial infarction and stroke with DPP-4i, but the evidence on heart failure is mixed.[7-12, 25] These studies mainly reported summary relative risk measures but not the absolute risk measures which may be important to put the issue in context. Further, several studies used a combined pool of non-DPP-4i drugs as the comparator making the results less useful for physicians for making treatment choices. To date there has not been any epidemiologic study comparing the incidence of CV events with DPP-4i versus clinically relevant comparators in a US population of older adults with a high prevalence of comorbidity and long duration of diabetes, both of which could affect the effects of DPP-4i on CV risk.

We therefore compared the relative and absolute risk of CV outcomes among initiators of DPP-4i versus relevant oral drug alternatives sulfonylureas (SU) and thiazolidinediones (TZD) using a 20% sample of the Medicare fee-for-service beneficiaries. Specifically, we examined the risk of non-fatal myocardial infarction (MI), stroke, hospitalization for heart failure (HF) and a composite outcome including MI, stroke, and all-cause mortality.

II. Methods

Study Population

We conducted an active-comparator new-user cohort study using a 20% random sample of Medicare beneficiaries >65 years with fee-for-service Part A, B and D enrollment in at least

one month during a calendar year from January 1, 2007 to December 31, 2013. This dataset contains information about demographics, enrollment, diagnoses, procedures and prescription drugs for each enrollee and has been previously used to study antihyperglycemic drugs.[49-51]

From this population, we identified two new-user active-comparator cohort pairs (TABLE 4.19, TABLE 4.20) mimicking a clinical treatment decision:[52] 1. DPP-4i versus SU (not exposed to either DPP-4i or SU in the previous 6 months) and 2. DPP-4i versus TZD (not exposed to either DPP-4i or TZD in the previous 6 months). Initiation was defined as the first prescription of the drug after a 6 month washout. Patients were allowed to be on any drugs except the drugs compared. A DPP-4i initiator with no previous prescription of SU and TZD would be eligible for inclusion in both comparisons. Patients were required to have at least 6 months of continuous Part D enrollment and at least 12 months parts A and B enrollment pre-initiation. To reduce the potential for bias toward the null due to secondary nonadherence, we restricted our cohorts to patients with a second prescription for the same drug class dispensed within 6 months after initiation and follow-up started from the second fill date. Since TZDs are contraindicated in patients with HF (which can lead to intractable confounding by contraindication), for DPP-4i versus TZD analyses we further excluded patients with diagnoses of HF and related conditions (cardiomyopathy, arrhythmias, chronic kidney disease, edema and loop diuretics use).

Outcomes

The outcomes assessed were non-fatal MI, stroke, HF hospitalization and all-cause mortality and a composite outcome of non-fatal MI, stroke and all-cause mortality based on the outcome definition in the RCTs. Medicare claims do not include information on causes of death and we could not identify cardiovascular death. However since cardiovascular deaths account for

>50% deaths in diabetes patients, we used all-cause mortality as proxy.[29,53] MI was defined using *International Classification of Diseases, Ninth Revision* (*ICD-9*) code 410 in the first or second position of the inpatient claims (definition with a positive predictive value of 94% in a Medicare population).[30] Stroke was defined using ICD-9 codes 430, 431, 433.x1, 434.x1, and 436, located in the first position (specificity 95–97%, sensitivity 74–90%).[31] HF hospitalization was defined using ICD-9 code 428.xx in the primary position which has a specificity >98% but a very low sensitivity of 21% in a Medicare population.[32]

Patients were followed from the second prescription until the earliest of: the outcome of interest, discontinuation, switching to or augmentation with the comparator drug, non-end point event (example, stroke is a non-end point event in the analysis of MI), end of enrollment, or December 31, 2013.

Confounding control and analysis

We used propensity scores (PS) to control for measured confounding. Using variables in tables 1 and 2 measured before initiation, we predicted the probability for initiating DPP-4i versus SU and DPP-4i versus TZD for each patient (PS) using two separate logistic regression models.[36] We then assigned a weight of 1 to DPP-4i and a weight of (PS/(1-PS)) to SU and TZD. Such weighting creates pseudo-populations of SU and TZD initiators with similar covariate distribution as in DPP-4i.[37, 38] Our weighted analysis thus answers the question "what would have happened to patients who initiated DPP-4i if they had initiated SU or TZD, instead".[42]

Competing Risk

Competing risks arise when the occurrence of one event precludes the occurrence of other events. In our study of older adults, mortality is a competing event and standard Cox models censoring patients who die yield biased estimates because this type of censoring may be

'informative'.[39] We therefore used weighted cumulative incidence curves accounting for competing risk by death to estimate the risk, risk differences (RD) and risk ratios (RR) for non-fatal MI, stroke and HF hospitalizations among initiators of DPP-4i versus comparators.[39] We obtained confidence intervals by bootstrapping 1000 replicates. We analyzed the composite outcome of non-fatal MI, stroke or all-cause mortality using traditional weighted Cox models.

Subgroup and Sensitivity analyses:

Analyses were repeated in pre-specified subgroups based on CVD history. Several sensitivity analyses were performed. To increase the probability of CV death, we excluded the deaths of patients with codes for metastatic cancer anytime during follow-up. Since SU are older drugs and to ensure that patients were indeed treated with SU as second line drugs (rather than first-line), we repeated the DPP-4i versus SU analyses requiring patients to have at least one metformin prescription in the 6 months before initiation. Finally repeated all analyses using an intent-to-treat approach where patients were not censored for treatment changes but followed from the second prescription to the earliest of the outcome, non-event end point, end of enrollment, or December 31, 2013.

III. Results

For the DPP-4i versus SU comparison (TABLE 4.1), there were 44,771 DPP-4i and 119,436 SU initiators. Compared with the DPP-4i, SU initiators were slightly older, less likely to have hyperlipidemia, diabetes complications, less likely to be on metformin or other antihyperglycemic medications, statins, and less likely to have had influenza vaccinations and lipid panels at baseline. For the DPP-4i versus TZD comparison (TABLE 4.2), there were 26,198 DPP-4i and 18,842 TZD initiators without previous HF/related diagnoses (excluded because TZD are contraindicated in those with pre-existing HF).[54] TZD initiators were generally more

comparable with DDP4-i initiators than SU initiators. Compared with the DPP-4i initiators, the TZD initiators were nevertheless more likely to be male and non-white. TZD initiators were less likely to have hyperlipidemia at baseline and less likely to get influenza vaccinations and lipid panels compared to DPP-4i initiators. After weighting all covariates in the weighted TZD and SU pseudo-populations were identical to the distribution of the DPP-4i initiators.

For the DPP-4i versus SU comparison, based on 4,720 events (725 MI, 593 stroke and 3,770 deaths) among DPP-4i initiators and 20,274 events (2,765 MI, 2,209 stroke and 17,046 deaths) among SU initiators, the adjusted hazard ratio was 0.83 (0.80, 0.86) (TABLE 4.3, FIGURE 4.1). This was mainly driven by death (FIGURE 4.2) rather than MI and stroke for which risks were approximately 1% at the median of ~1 year of treatment (FIGURE 4.3). The adjusted RD per 100 patients for MI comparing DPP-4i versus SU ranged from -0.27 (-0.46, -0.06) in year 1 to -0.78 (-1.62, 0.12) in year 5 and ranged from -0.11 (-0.23, 0.03) to -0.81 (-1.30, -0.22) for stroke (TABLE 4.4,TABLE 4.5) indicating no meaningful difference in the risk of MI or stroke between DPP-4i and SU. In the subgroup without prior CVD, the adjusted risks of MI and stroke for both DPP-4i and SU groups were <1% at 1 year after initiation and the magnitudes of RDs were <1 per 100 patients (TABLE 4.6, TABLE 4.7, FIGURE 4.4). In the subgroup with prior CVD, the risks for MI and stroke were slightly higher (~1.5% at 1 year), but the magnitude of RD per 100 patients was <1 (FIGURE 4.4, TABLE 4.8, TABLE 4.9). No increased risk of HF hospitalization was observed with DPP-4i versus SU (FIGURE 4.5).

For DPP-4i versus TZD, based on 1,760 events among DPP-4i initiators and 1,466 events among TZD initiators the adjusted HR for the composite outcome was 0.94 (0.88, 1.01) (FIGURE 4.1, TABLE 4.3). The MI risk for both DPP-4i and TZD groups was ~1% at 1 year after initiation with magnitude of RDs <1 per 100 patients (TABLE 4.10, TABLE 4.11). In the

subgroup without prior CVD, the 1-year risks of MI and stroke were <1% for DPP-4i and TZD (TABLE 4.12, TABLE 4.13, FIGURE 4.3). In the subgroup with prior CVD, the 1-year risks were >1% for MI and stroke for both DPP-4i and TZD (FIGURE 4.6,TABLE 4.14, TABLE 4.15). The RD per 100 patients for HF hospitalization comparing DPP-4i versus TZD were between 0 and -1 during the study period (TABLE 4.16, FIGURE 4.5).

Sensitivity analyses after excluding metastatic cancer deaths from all-cause mortality (which accounted for 14-16% of deaths in all treatment groups) did not change the results (TABLE 4.17). When we restricted the study population to patients who had a metformin script in the 6 months before initiation of DPP-4i or SU, the adjusted HR of the composite outcome was 0.77 (0.74, 0.81) and again no difference in individual risks of MI or stroke between DPP-4i or SU as evinced by 1-year RDs <1 per 100 patients. Additional analyses using an intent-to-treat approach did not change the results (TABLE 4.18).

IV. Discussion and Conclusions

We found no evidence of an increased short-term risk of CV events with DPP-4i versus SU or TZD in our new-user active-comparator cohort study based on a 20% random sample of all US Medicare fee-for-service beneficiaries 2007-2013. The results were consistent across subgroups based on prior CVD and several sensitivity analyses. The individual risks of MI and stroke at 1 year after initiation in the subgroup with prior CVD (just above 1%) were slightly higher than the subgroup without prior CVD (about 0.7%), but the risk differences in either subgroups for DPP-4i versus comparators were small (<1 per 100 patients). The apparent decreased risk of the composite CV outcome with DPP-4i versus SU is mainly driven by all-cause mortality. While there is debate about increased risk of all-cause mortality with SU, the theoretical risk differs based on which SU agent is used.

More information on the comparative risk with DPP-4i versus SU will be added once the CAROLINA trial comparing linagliptin with glimepiride is completed.[26, 55, 56] Our results of no increased risk of MI and stroke with DPP-4i are consistent with the RCTs on saxagliptin (SAVOR-TIMI), alogliptin (EXAMINE) and sitagliptin (TECOS) which found no increased risk of MI, stroke with DPP-4i versus placebo in high risk populations.[2-4]

We did not observe an increased risk of HF hospitalization with DPP-4i versus SU. While the EXAMINE and TECOS trials did not find an increased HF risk with alogliptin and sitagliptin respectively, the SAVOR-TIMI 53 trial reported a 27% increased risk of HF hospitalization with saxagliptin versus placebo. Several factors could explain the discrepancy. First, the SAVOR-TIMI 53 examined saxagliptin alone while our DPP-4i cohort mainly consisted of sitagliptin initiators (~75%) and sample size was not sufficient to study saxagliptin alone. Second, the treatment duration in our study was shorter than in the trial. Observational studies examining HF risk with DPP-4i using different designs, populations and comparators report mixed results. Two studies reported a reduced rate of HF compared to other anti-hyperglycemic drugs[9, 12], three studies suggest no difference in effect[8, 10, 57], while two studies reported increased HF risk with DPP-4i compared to other antihyperglycemic drugs.[7, 25] Most of these studies used a heterogeneous comparator of 'all other antihyperglycemic drugs' makes interpretation of results hard particularly in cases where the risks differed greatly depending on the comparator.[7, 11, 12, 25, 57] One study that reported increased risk of HF compared sitagliptin initiators to matched controls who were prevalent users of antihyperglycemic therapy which could bias the results due to prevalent users possibly being tolerant to other antihyperglycemic therapy.[25] Some clinical studies on the other hand, have suggested a protective role of DPP-4i in the pathogenesis of CHF.[58, 59]

Since TZDs are known to be associated with an increased HF risk, we used HF hospitalization as a positive control outcome in the DPP-4i versus TZD analysis expecting no increased risk with DPP-4i relative to TZD and that is what we observed. Taken together the evidence from our study and existing literature suggests that there is no concern of increased risk of HF hospitalizations with DPP-4i.

A strength of our study is the use of a new-user active-comparator cohort design which is analogous to a head-to-head clinical trial and answers the more relevant question of 'which second-line treatment to initiate' rather than 'treatment or not'.[27] Specifically, we identified initiators of DPP-4i or comparators after 6 months without use of DPP4i or comparator and covariates were measured before initiation thereby avoiding the problem of controlling for covariates potentially affected by treatment.[27] The good balance of measured covariates achieved by the study design implies that unmeasured covariates could also be balanced, although this cannot be proven. The balance of measured covariates was further improved by PS weighting and reassures us about absence of confounding by these covariates. We also used specific comparator groups and reported results separately for each comparison unlike a few other observational studies where the comparator group consisted of 'all other antihyperglycemic drugs' making interpretation difficult. Our study also raises an important consideration about which is the best comparator for a second line antihyperglycemic drug. As mentioned above, since TZDs are contraindicated in patients with existing HF, for this comparison we had to exclude patients with previous diagnoses of HF or related conditions in order to identify patients with treatment equipoise. On the other hand, more initiators of DPP-4i compared to SU had baseline use of metformin and long acting insulin possibly suggesting that the SU group consisted of newer diabetes patients compared to DPP-4i. To address this, we repeated the

analyses requiring a prescription of metformin (first line drug) at baseline to ensure that patients with second line therapy were being compared. Researchers should be cognizant of such differences and identify ways to increase treatment equipoise, i.e., patients in whom the choice of treatments not driven by predictors of the risk for the outcome.

Following caveats should be considered. First, the time-on-treatment in our primary as-treated analysis was short (median ~1 year) but that is mostly a function of the real-world treatment dynamics. Second, since Medicare claims do not contain information on causes of death, we could not identify cardiovascular death. Sensitivity analyses excluding deaths in patients with metastatic cancer to increase the contribution of cardiac death to all-cause mortality did not change results. Third, given the absence of clinical measures it is hard to identify HF using claims data and our definition of HF hospitalization had a near perfect specificity (which yields unbiased relative risks) but a low sensitivity which will lead to an underestimation of absolute risks. Fourth, since occurrence of one CV event during follow-up might affect the incidence of a subsequent CV event, (example, MI can affect the risk of stroke), we censored patients at non-end point CV events. This could theoretically lead to 'competing risk' from the non-end point event, but such censoring was extremely rare in our study. Finally we were not able to measure and adjust for lifestyle variables like smoking and body mass index directly (BMI). However we adjusted for codes for tobacco use and chronic obstructive pulmonary disease as proxies for smoking and smoking is unlikely to meaningfully affect the choice of second line treatments. We also previously found that BMI does not affect the choice of initiation of DPP-4i versus SU and TZD and therefore is unlikely to be a confounder in this setting.[60]

In summary, we did not observe an increased short-term risk of CV events with DPP-4i versus relevant oral second line diabetes drugs in a population of older adults. Along with the RCT results, our results based on real-world drug use and effects are relevant to physicians for making antihyperglycemic treatment choices.

V. Tables and Figures

TABLE 4.1. Characteristics of initiators[a] of dipeptidyl peptidase-4 inhibitors and sulfonylureas : Medicare claims data 2006 – 2013					
	DPP-4i (44,771)		**SU (N = 119,436)**		**weighted SU [b]**
	N	**%**	**N**	**%**	**%**
66 to 75 years old	24,839	55.5	63,885	53.5	55.7
76 to 85 years old	15,229	34.0	40,767	34.1	33.8
86 years and above	4,703	10.5	14,784	12.4	10.5
Male	17,520	39.1	50,230	42.1	39.0
White	33,076	73.9	92,933	77.8	73.9
Black	4,681	10.5	14,030	11.8	10.4
Other	7,014	15.7	12,473	10.4	15.7
Baseline cardiovascular comorbidities [c]					
Diagnosis of acute myocardial infarction	1,489	3.3	4,694	3.9	3.4
Diagnosis of old myocardial infarction	2,713	6.1	7,566	6.3	6.1
Angina	3,306	7.4	7,186	6.0	7.4
Cardiomyopathy	2,606	5.8	7,268	6.1	5.8
Arrhythmia, syncope or pacemaker	12,876	28.8	35,347	29.6	28.9
Heart failure	9,900	22.1	27,926	23.4	22.3
Hyperlipidemia/ lipid disorder	37,726	84.3	91,849	76.9	84.4
Hypertension	40,849	91.2	106,725	89.4	91.3
Hypertensive heart disease with heart failure	1,687	3.8	3,593	3.0	3.9
Diagnosis of Stroke or TIA	3,614	8.1	10,480	8.8	8.2
Other acute and sub-acute forms of ischemic heart disease	1,880	4.2	4,844	4.1	4.3
Other baseline comorbidities[c]					
Obesity diagnosis	658	1.5	1,599	1.3	1.5
Diabetic nephropathy	3,943	8.8	8,835	7.4	8.9
Diabetic neuropathy	9,256	20.7	19,045	16.0	20.9
Diabetic retinopathy	6,959	15.5	13,768	11.5	15.7
Bone fractures	3,171	7.1	8,496	7.1	7.2
Chronic obstructive pulmonary disease	9,637	21.5	26,968	22.6	21.4
Cancer (Non-Skin)	7,746	17.3	20,207	16.9	17.2
Chronic kidney disease (CKD)	8,706	19.5	23,177	19.4	19.6
CKD Stage1	518	1.2	1,131	1.0	1.2
CKD Stage2	1,155	2.6	2,693	2.3	2.6
CKD Stage3	4,683	10.5	11,240	9.4	10.6
CKD Stage4	1,355	3.0	3,415	2.9	3.1

CKD Stage5	283	0.6	776	0.7	0.6
End Stage renal disease	698	1.6	2,067	1.7	1.6
Chronic lung diseases	8,375	18.7	21,884	18.3	18.8
Connective tissue diseases	15,209	34.0	35,193	29.5	34.0
Dementia	3,833	8.6	12,189	10.2	8.6
Depression	7,608	17.0	19,477	16.3	17.0
Edema	7,912	17.7	20,059	16.8	17.7
Foot ulcers	5,072	11.3	12,054	10.1	11.4
Inflammatory GI diseases	412	0.9	1,051	0.9	0.9
Liver disease	2,891	6.5	6,707	5.6	6.5
Other chronic ischemic heart disease	17,683	39.5	45,214	37.9	39.6
Other neurological disorders	8,739	19.5	19,345	16.2	19.6
Other ocular disorders	18,662	41.7	43,968	36.8	41.7
Other renal disorders	5,355	12.0	15,730	13.2	12.1
Pancreatitis	557	1.2	1,634	1.4	1.3
Peripheral vascular disease (claudication)	7,795	17.4	18,526	15.5	17.4
Proteinuria	1,707	3.8	3,724	3.1	3.9
Tobacco use diagnosis	4,363	9.8	12,927	10.8	9.7
Urinary tract infection/cystitis	12,128	27.1	30,175	25.3	27.2
Upper Respiratory	13,018	29.1	31,451	26.3	29.2
Alcoholism	77	0.2	287	0.2	0.2
Aortic aneurysm	1,228	2.7	3,224	2.7	2.7
Dialysis	399	0.9	1,306	1.1	0.9
Baseline medication use [d]					
Metformin	28,526	63.7	65,528	54.9	64.3
Sulfonylureas	11,502	25.7	17,845	14.9	25.8
Glucagon-like peptide-1 agonists	805	1.8	1,444	1.2	1.9
Long acting insulin	8,529	19.1	16,046	13.4	19.5
Short acting insulin	4,311	9.6	9,244	7.7	9.7
Anticholinergic	1,222	2.7	3,760	3.2	2.7
Statins	28,295	63.2	67,347	56.4	63.4
Angiotensin converting enzyme inhibitors	17,146	38.3	50,479	42.3	38.3
Angiotensin receptor blockers	9,733	21.7	19,138	16.0	22.0
Beta blockers	21,990	49.1	58,999	49.4	49.3
Calcium channel blockers	14,587	32.6	39,654	33.2	32.7
Loop diuretics	12,531	28.0	35,898	30.1	28.1
Other diuretics	18,310	40.9	47,183	39.5	41.0
Healthcare utilization variables[c]					
Blood tests	4,348	9.7	10,178	8.5	9.8
Influenza vaccinations	25,334	56.6	63,769	53.4	56.6

Lipid panels	39,101	87.3	95,216	79.7	87.5
Mammogram	9,971	22.3	22,466	18.8	22.4

[a] Initiation defined as no dispensed prescriptions for DPP-4i or SU during the 6 months before initiation and filling a second prescription of the same drug/drug class within 6 months after the first prescription.
[b] Pseudo-population of SU initiators weighted to the distribution of covariates of the DPP-4i initiators using the propensity score to balance covariates (and therefore control for confounding).
[c] Measured in the 12 months before drug initiation.
[d] Measured in the 6 months before drug initiation

TABLE 4.2. Characteristics of initiators[a] of dipeptidyl peptidase-4 inhibitors and thiazolidinediones : Medicare claims data 2006 – 2013					
	DPP-4i (26,198)		**TZD (N = 18,842)**		**weighted TZD[b]**
	N	**%**	**N**	**%**	**%**
66 to 75 years old	16,975	64.80	12,387	65.74	64.60
76 to 85 years old	7,559	28.85	5,339	28.34	28.87
86 years and above	1,664	6.35	1,116	5.92	6.54
male	10,390	39.66	8,217	43.61	39.64
white	19,777	75.49	13,444	71.35	75.55
black	2,338	8.92	1,978	10.50	8.88
other	4,083	15.59	3,420	18.15	15.57
Baseline cardiovascular comorbidities[c]					
Diagnosis of acute myocardial infarction	221	0.84	98	0.52	0.86
Diagnosis of old myocardial infarction	581	2.22	304	1.61	2.27
Angina	921	3.52	490	2.60	3.53
Hyperlipidemia/ lipid disorder	21,843	83.38	14,370	76.27	83.44
Hypertension	22,895	87.39	15,615	82.87	87.36
Diagnosis of Stroke or TIA	1,222	4.66	753	4.00	4.66
Other acute and sub-acute forms of ischemic heart disease	405	1.55	195	1.03	1.55
Other baseline comorbidities[c]					
Obesity diagnosis	253	0.97	119	0.63	1.00
Diabetic nephropathy	776	2.96	532	2.82	2.95
Diabetic neuropathy	3,996	15.25	2,393	12.70	15.53
Diabetic retinopathy	3,625	13.84	2,436	12.93	13.91
Bone fractures	1,062	4.05	740	3.93	4.20
Chronic obstructive pulmonary disease	2,885	11.01	1,963	10.42	11.09
Cancer (Non-Skin)	3,962	15.12	2,342	12.43	15.16
Chronic lung diseases	2,936	11.21	1,824	9.68	11.19
Connective tissue diseases	6,946	26.51	4,139	21.97	26.74
Dementia	1,279	4.88	871	4.62	4.98
Depression	3,090	11.79	1,849	9.81	11.85
Foot ulcers	1,823	6.96	1,094	5.81	7.06
Inflammatory GI diseases	169	0.65	126	0.67	0.62
Liver disease	1,367	5.22	804	4.27	5.29
Other neurological disorders	3,615	13.80	2,162	11.47	13.85
Other ocular disorders	10,371	39.59	6,327	33.58	39.69
Other renal disorders	574	2.19	342	1.82	2.18
Pancreatitis	197	0.75	133	0.71	0.77

Peripheral vascular disease (claudication)	2,523	9.63	1,537	8.16	9.78
Proteinuria	598	2.28	355	1.88	2.27
Tobacco use diagnosis	1,653	6.31	915	4.86	6.32
Urinary tract infection/cystitis	4,920	18.78	2,920	15.50	18.93
Upper Respiratory	6,627	25.30	4,101	21.77	25.38
Alcoholism	23	0.09	21	0.11	0.09
Aortic aneurysm	421	1.61	233	1.24	1.64
Baseline medication use[d]					
Metformin	19,348	73.85	13,292	70.54	73.74
Sulfonylureas	13,628	52.02	10,193	54.10	52.01
Glucagon-like peptide-1 agonists	468	1.79	308	1.63	1.85
Long acting insulin	2,956	11.28	2,175	11.54	11.44
Short acting insulin	1,077	4.11	822	4.36	4.16
Anticholinergic	373	1.42	308	1.63	1.40
Statins	15,814	60.36	11,537	61.23	60.59
Angiotensin converting enzyme inhibitors	10,098	38.54	7,894	41.90	38.53
Angiotensin receptor blockers	4,521	17.26	2,713	14.40	17.34
Beta blockers	9,618	36.71	6,113	32.44	36.57
Calcium channel blockers	7,090	27.06	4,743	25.17	27.11
Other diuretics	10,609	40.50	7,423	39.40	40.41
Healthcare utilization variables[c]					
Blood tests	2,411	9.20	1,611	8.55	9.22
Influenza vaccinations	14,713	56.16	9,603	50.97	56.36
Lipid panels	23,379	89.24	16,076	85.32	89.32
Mammogram	6,517	24.88	3,974	21.09	24.98

[a] Initiation defined as no dispensed prescriptions for DPP-4i or TZD during the 6 months before initiation and filling a second prescription of the same drug/drug class within 6 months after the first prescription. For the DPP-4i versus TZD analysis, patients with prevalent diagnosis of heart failure, hypertensive disease with heart failure, chronic kidney disease, edema and use of loop diuretics were excluded.

[b] Pseudo-population of SU initiators weighted to the distribution of covariates of the DPP-4i initiators using the propensity score to balance covariates (and therefore control for confounding).

[c]Measured in the 12 months before drug initiation.

[d]Measured in the 12 months before drug initiation

TABLE 4.3. Number of initiators, events (composite of nonfatal myocardial infarction, stroke and death), treatment duration, person–years, event rates and crude and adjusted hazard ratios for DPP-4i versus comparators in the entire population as well as subgroups based on prior cardiovascular disease

Comparison	Treatment	Number of new-users[a]	events	Treatment duration [b] interquartile range (median)	Total person-years	Incidence (per 100 person years)	Unadjusted HR (95% CI) [c]	Adjusted HR (95%CI) [d]
Entire population								
DPP-4i vs SU	**DPP-4i[e]**	44,771	4,720	0.58 - 1.96 (1.04)	63,709	7.4	0.71 (0.69, 0.74)	0.83 (0.80, 0.86)
	SU	119,436	20,274	0.62 - 2.39 (1.21)	197,245	10.3	1.00 (reference)	1.00 (reference)
DPP-4i vs TZD	**DPP-4i[e]**	26,198	1,760	0.60 - 2.08 (1.09)	39,553	4.4	0.93 (0.87, 1.00)	0.94 (0.88, 1.01)
	TZD	18,842	1,466	0.68 - 2.27 (1.17)	30,635	4.8	1.00 (reference)	1.00 (reference)
Subgroup without CVD at baseline								
DPP-4i vs SU	**DPP-4i[e]**	15,602	792	0.59 - 2.05 (1.11)	23,517	3.4	0.67 (0.62, 0.73)	0.78 (0.72, 0.85)
	SU	41,934	3,847	0.67 - 2.64 (1.33)	75,520	5.0	1.00 (reference)	1.00 (reference)
DPP-4i vs TZD	**DPP-4i[e]**	17,742	994	0.61 - 2.10 (1.11)	27,198	3.6	0.88 (0.80, 0.96)	0.91 (0.83, 1.00)
	TZD	13,725	956	0.69 - 2.32 (1.21)	22,816	4.2	1.00 (reference)	1.00 (reference)
Subgroup with CVD at baseline								
DPP-4i vs SU	**DPP-4i[e]**	26,873	3,722	0.58 - 1.89 (0.99)	36,691	10.1	0.72 (0.70, 0.75)	0.83 (0.80, 0.86)
	SU	71,052	15,322	0.59 - 2.21 (1.12)	110,013	13.9	1.00 (reference)	1.00 (reference)
DPP-4i vs TZD	**DPP-4i[e]**	25,601	1,683	0.61 - 2.08 (1.09)	38,774	4.3	0.93 (0.87, 0.99)	0.94 (0.87, 1.01)
	TZD	18,445	1,406	0.68 - 2.27 (1.18)	30,017	4.7	1.00 (reference)	1.00 (reference)

DPP-4i - Dipeptidyl peptidase 4 inhibitors, TZD - thiazolidinediones, SU - sulfonylureas, CVD - cardiovascular disease

[a] Initiation defined as no dispensed prescriptions of the drugs being compared during the 6 months before initiation and filling a second prescription of the same drug/drug class within 6 months after the first prescription.

[b] Patients followed up from the 2nd prescription to the earliest of the following: outcome of interest,

[c]Hazard ratios and their 95 % confidence intervals from Cox proportional hazards models for the composite outcome with baseline treatment as the only independent covariate.

[d] Hazard ratios adjusted for variables in Table 1 and 2 using propensity score weighting (standardized to DPP-4i population).

[e]Number of DPP-4i initiators different in both cohorts because for the DPP-4i vs SU comparison, patients could be on any diabetes medication (including TZD) except for DPP-4i and SU during the washout period. Similarly for the DPP-4i versus TZD comparison, patients could be on any other drugs except DPP-4i and TZD. Further for the DPP-4i versus TZD analysis, patients with prevalent diagnosis of heart failure, hypertensive disease with heart failure, chronic kidney disease, edema and use of loop diuretics were excluded for all analyses.

TABLE 4.4. Risks, risk differences and relative risks for myocardial infarction accounting for competing risk by death: DPP-4i versus SU			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	1.22 (1.13, 1.32)	-0.27 (-0.46, -0.06)	0.81 (0.71, 0.95)
2 year	2.31 (2.15, 2.48)	-0.47 (-0.77, -0.15)	0.83 (0.74, 0.93)
3 year	3.11 (3.07, 3.54)	-0.63 (-1.10, -0.21)	0.84 (0.74, 0.94)
4 year	4.38 (4.04, 4.73)	-0.84 (-1.49, -0.22)	0.83 (0.73, 0.95)
5 year	5.45 (4.97, 6.00)	-0.78 (-1.62, 0.12)	0.88 (0.76, 1.02)
SU			
1 year	1.49 (1.32, 1.66)	Ref	Ref
2 year	2.79 (2.50, 3.04)	Ref	Ref
3 year	3.94 (3.56, 4.32)	Ref	Ref
4 year	5.26 (4.70, 5.74)	Ref	Ref
5 year	6.25 (5.51, 6.96)	Ref	Ref

TABLE 4.5. Risks, risk differences and relative risks for stroke accounting for competing risk by death: DPP-4i versus SU			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.92 (0.82, 1.02)	-0.11 (-0.23, 0.03)	0.89 (0.79, 1.03)
2 year	1.72 (1.55, 1.88)	-0.26 (-0.44, -0.07)	0.87 (0.78, 0.96)
3 year	2.46 (2.24, 2.69)	-0.30 (-0.58, -0.03)	0.89 (0.79, 0.99)
4 year	3.16 (2.84, 3.53)	-0.49 (-0.87, -0.05)	0.86 (0.76, 0.98)
5 year	3.63 (3.17, 4.13)	-0.81 (-1.30, -0.22)	0.81 (0.71, 0.95)
SU			
1 year	1.02 (0.96, 1.10)	Ref	Ref
2 year	1.97 (1.86, 2.09)	Ref	Ref
3 year	2.76 (2.60, 2.92)	Ref	Ref
4 year	3.65 (3.43, 3.88)	Ref	Ref
5 year	4.45 (4.14, 4.77)	Ref	Ref

TABLE 4.6. Risks, risk differences and relative risks for myocardial infarction accounting for competing risk by death: DPP-4i versus SU Subgroup without prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.64 (0.49, 0.80)	-0.10 (-0.26, 0.07)	0.85 (0.67, 1.09)
2 year	1.42 (1.20, 1.73)	-0.04 (-0.30, 0.27)	0.97 (0.80, 1.18)
3 year	1.79 (1.44, 2.14)	-0.41 (-0.83, 0.02)	0.81 (0.64, 1.00)
4 year	2.36 (1.96, 3.01)	-0.49 (-1.31, 0.20)	0.83 (0.65, 1.07)
5 year	3.49 (2.76, 4.58)	0.08 (-0.91, 1.11)	1.03 (0.75, 1.32)
SU			
1 year	0.74 (0.64, 0.83)	Ref	Ref
2 year	1.46 (1.33, 1.62)	Ref	Ref
3 year	2.22 (2.01, 2.48)	Ref	Ref
4 year	2.86 (2.58, 3.19)	Ref	Ref
5 year	3.43 (3.03, 3.81)	Ref	Ref

TABLE 4.7. Risks, risk differences and relative risks for stroke accounting for competing risk by death: DPP-4i versus SU subgroup without prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.48 (0.40, 0.63)	-0.21 (-0.34, -0.07)	0.69 (0.54, 0.88)
2 year	0.92 (0.72, 1.08)	-0.48 (-0.77, -0.30)	0.65 (0.48, 0.77)
3 year	1.36 (1.07, 1.64)	-0.76 (-1.19, -0.45)	0.64 (0.49, 0.77)
4 year	1.66 (1.37, 2.08)	-1.19 (-1.70, -0.75)	0.58 (0.46, 0.73)
5 year	2.17 (1.64, 2.82)	-1.39 (-2.10, -0.83)	0.60 (0.45, 0.76)
SU			
1 year	0.69 (0.61, 0.78)	Ref	Ref
2 year	1.41 (1.26, 1.56)	Ref	Ref
3 year	2.11 (1.93, 2.39)	Ref	Ref
4 year	2.85 (2.61, 3.24)	Ref	Ref
5 year	3.60 (3.25, 4.05)	Ref	Ref

TABLE 4.8. Risks, risk differences and relative risks for myocardial infarction accounting for competing risk by death: DPP-4i versus SU Subgroup with prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	1.41 (1.25, 1.54)	-0.44 (-0.64, -0.25)	0.76 (0.68, 0.85)
2 year	2.59 (2.33, 2.84)	-0.50 (-0.85, -0.19)	0.83 (0.73, 0.94)
3 year	3.54 (3.23, 3.91)	-0.91 (-1.36, -0.47)	0.79 (0.70, 0.88)
4 year	4.58 (4.09, 5.08)	-0.86 (-1.40, -0.33)	0.84 (0.75, 0.93)
5 year	6.04 (5.20, 7.06)	-0.37 (-1.43, 0.65)	0.94 (0.79, 1.10)
SU			
1 year	1.84 (1.72, 1.99)	Ref	Ref
2 year	3.11 (2.94, 3.30)	Ref	Ref
3 year	4.43 (4.18, 4.70)	Ref	Ref
4 year	5.47 (5.17, 5.84)	Ref	Ref
5 year	6.53 (6.02, 6.98)	Ref	Ref

TABLE 4.9. Risks, risk differences and relative risks for stroke accounting for competing risk by death: DPP-4i versus SU Subgroup with prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	1.18 (1.03, 1.35)	-0.05 (-0.25, 0.13)	0.96 (0.81, 1.10)
2 year	2.19 (1.90, 2.41)	-0.14 (-0.47, 0.14)	0.94 (0.81, 1.06)
3 year	3.08 (2.74, 3.42)	-0.12 (-0.56, 0.33)	0.95 (0.83, 1.11)
4 year	4.01 (3.40, 4.44)	-0.23 (-0.81, 0.46)	0.95 (0.81, 1.11)
5 year	4.30 (3.71, 4.84)	-0.74 (-1.48, 0.12)	0.85 (0.71, 1.02)
SU			
1 year	1.24 (1.12, 1.35)	Ref	Ref
2 year	2.36 (2.14, 2.51)	Ref	Ref
3 year	3.20 (2.93, 3.40)	Ref	Ref
4 year	4.22 (3.94, 4.55)	Ref	Ref
5 year	5.04 (4.58, 5.66)	Ref	Ref

TABLE 4.10. Risks, risk differences and relative risks for myocardial infarction accounting for competing risk by death: DPP-4i versus TZD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.87 (0.76, 1.00)	-0.12 (-0.36, 0.09)	0.88 (0.68, 1.12)
2 year	1.80 (1.61, 2.03)	-0.15 (-0.54, 0.22)	0.92 (0.76, 1.12)
3 year	2.54 (2.24, 2.85)	-0.43 (-0.94, 0.16)	0.86 (0.74, 1.06)
4 year	3.70 (3.19, 4.17)	-0.64 (-1.11, 0.33)	0.85 (0.68, 1.09)
5 year	4.79 (4.25, 5.58)	-0.13 (-0.99, 0.70)	0.98 (0.83, 1.31)
TZD			
1 year	1.01 (0.82, 1.16)	Ref	Ref
2 year	1.97 (1.66, 2.34)	Ref	Ref
3 year	2.94 (2.48, 3.36)	Ref	Ref
4 year	4.25 (3.56, 4.98)	Ref	Ref
5 year	4.85 (4.06, 5.77)	Ref	Ref

TABLE 4.11. Risks, risk differences and relative risks for stroke accounting for competing risk by death: DPP-4i versus TZD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.75 (0.65, 0.85)	-0.04 (-0.23, 0.12)	0.94 (0.75, 1.17)
2 year	1.47 (1.31, 1.71)	0.14 (-0.14, 0.44)	1.10 (0.89, 1.42)
3 year	2.08 (1.75, 2.42)	0.06 (-0.48, 0.49)	0.97 (0.79, 1.26)
4 year	2.48 (2.33, 3.31)	0.00 (-0.70, 0.75)	0.99 (0.79, 1.29)
5 year	3.45 (2.74, 4.16)	0.40 (-0.38, 1.26)	1.13 (0.89, 1.44)
TZD			
1 year	0.80 (0.65, 0.93)	Ref	Ref
2 year	1.33 (1.06, 1.55)	Ref	Ref
3 year	2.10 (1.82, 2.47)	Ref	Ref
4 year	2.83 (2.41, 3.35)	Ref	Ref
5 year	3.04 (2.57, 3.61)	Ref	Ref

TABLE 4.12. Risks, risk differences and relative risks for myocardial infarction accounting for competing risk by death: DPP-4i versus TZD subgroup without prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.75 (0.58, 0.89)	0.03 (-0.14, 0.24)	1.04 (0.79, 1.39)
2 year	1.54 (1.32, 1.81)	-0.03 (-0.42, 0.39)	0.98 (0.76, 1.27)
3 year	1.99 (1.66, 2.33)	-0.32 (-0.90, 0.15)	0.86 (0.64, 1.07)
4 year	2.71 (2.20, 3.33)	-0.60 (-1.60, 0.23)	0.82 (0.59, 1.08)
5 year	3.89 (3.10, 4.93)	0.34 (-1.02, 1.63)	1.09 (0.78, 1.48)
TZD			
1 year	0.71 (0.57, 0.86)	Ref	Ref
2 year	1.57 (1.29, 1.88)	Ref	Ref
3 year	2.29 (1.93, 2.80)	Ref	Ref
4 year	3.36 (2.73, 4.04)	Ref	Ref
5 year	3.62 (2.76, 4.60)	Ref	Ref

TABLE 4.13. Risks, risk differences and relative risks for stroke accounting for competing risk by death: DPP-4i versus TZD subgroup without prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	0.55 (0.44, 0.66)	-0.12 (-0.29, 0.05)	0.81 (0.62, 1.09)
2 year	1.15 (0.92, 1.35)	-0.06 (-0.34, 0.22)	0.95 (0.74, 1.22)
3 year	1.52 (1.22, 1.84)	-0.31 (-0.89, 0.09)	0.83 (0.59, 1.06)
4 year	2.33 (1.89, 2.92)	-0.15 (-0.72, 0.61)	0.94 (0.73, 1.30)
5 year	2.94 (2.23, 3.67)	0.21 (-0.60, 1.25)	1.07 (0.79, 1.54)
TZD			
1 year	0.67 (0.53, 0.82)	Ref	Ref
2 year	1.21 (0.97, 1.43)	Ref	Ref
3 year	1.83 (1.47, 2.23)	Ref	Ref
4 year	2.47 (1.98, 3.03)	Ref	Ref
5 year	2.74 (2.19, 3.27)	Ref	Ref

TABLE 4.14. Risks, risk differences and relative risks for myocardial infarction accounting for competing risk by death: DPP-4i versus TZD subgroup with prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	1.12 (8.56, 1.39)	-0.40 (-0.93, -0.01)	0.74 (0.49, 0.99)
2 year	2.40 (1.85, 2.90)	-0.30 (-1.10, 0.36)	0.88 (0.64, 1.51)
3 year	3.77 (3.00, 4.47)	-0.50 (-1.87, 0.74)	0.87 (0.64, 1.20)
4 year	5.86 (4.54, 7.28)	-0.48 (-2.42, 1.67)	0.92 (0.67, 1.29)
5 year	7.22 (5.72, 9.05)	-0.64 (-3.49, 1.88)	0.92 (0.65, 1.32)
TZD			
1 year	1.54 (1.18, 1.88)	Ref	Ref
2 year	2.71 (2.19, 3.34)	Ref	Ref
3 year	4.25 (3.54, 5.36)	Ref	Ref
4 year	6.28 (4.78, 7.87)	Ref	Ref
5 year	7.74 (5.95, 10.00)	Ref	Ref

TABLE 4.15. Risks, risk differences and relative risks for stroke accounting for competing risk by death: DPP-4i versus TZD subgroup with prior CVD			
	Risk per 100 patients	**Risk difference (95% CI)**	**Relative risk (95% CI)**
DPP-4i			
1 year	1.15 (0.89, 1.46)	0.00 (-0.44, 0.49)	0.99 (0.68, 1.56)
2 year	2.15 (1.73, 2.60)	0.59 (-0.10, 1.23)	1.37 (0.95, 2.00)
3 year	3.28 (2.67, 4.03)	0.47 (-0.49, 1.65)	1.17 (0.86, 1.81)
4 year	3.82 (3.04, 4.73)	0.19 (-1.09, 1.58)	1.05 (0.75, 1.58)
5 year	4.35 (3.28, 6.00)	0.70 (-0.79, 2.36)	1.20 (0.82, 1.82)
TZD			
1 year	1.17 (0.77, 1.54)	Ref	Ref
2 year	1.59 (1.11, 2.06)	Ref	Ref
3 year	2.79 (1.90, 3.60)	Ref	Ref
4 year	3.67 (2.63, 4.87)	Ref	Ref
5 year	3.70 (2.60, 4.90)	Ref	Ref

TABLE 4.16. Adjusted risks and risk differences per 100 patients, and relative risks for hospitalization for heart failure after adjusting for competing risk by death : DPP-4i versus TZD subgroup without prevalent HF or use of loop diuretics

	Risk (95% CI)	**RD (95% CI)**	**RR (95% CI)**
DPP (N = 26,198)			
1 year	0.50 (0.42, 0.61)	-0.20 (-0.40, 0.09)	0.67 (0.55, 0.82)
2 year	1.12 (1.01, 1.30)	-0.35 (-0.61, 0.22)	0.77 (0.65, 0.91)
3 year	1.91 (1.61, 2.34)	-0.10 (-0.40, 0.62)	0.84 (0.70, 0.99)
4 year	3.97 (3.49, 4.47)	-0.16 (-0.91, 0.56)	0.96 (0.80, 1.15)
5 year	4.71 (4.09, 5.37)	-0.87 (-2.10, 0.18)	0.84 (0.68, 1.04)
TZD (N = 18,842)			
1 year	0.70 (0.89, 1.19)	Ref	Ref
2 year	1.30 (1.72, 2.24)	Ref	Ref
3 year	2.03 (2.64, 3.46)	Ref	Ref
4 year	4.11 (3.59, 4.71)	Ref	Ref
5 year	5.58 (4.70, 6.662)	Ref	Ref

TABLE 4.17. Sensitivity analysis using cardiovascular death as a component of the composite outcome (nonfatal MI, stroke and cardiovascular death a) for DPP-4i versus TZD and DPP-4i versus SU

Comparison	**Treatment**	**Number of new-users**	**events**	**time to event in years interquartile range (median)**	**Total person-years**	**Incidence (per 100,000 person years)**	**Unadjusted HR (95% CI)** †	**Adjusted HR (95%CI)** §
DPP-4i vs SU	DPP-4i	44,771	4140	0.58 - 1.96 (1.04)	63,725	6496.67	0.72 (0.69, 0.74)	0.83 (0.81, 0.86)
	SU	119,436	17783	0.62 - 2.39 (1.21)	197,239	9015.97	1.00 (reference)	1.00 (reference)
DPP-4i vs TZD	DPP-4i	26,198	1750	0.58 - 1.96 (1.04)	39,103	4475.36	0.94 (0.87, 1.00)	0.95 (0.88, 1.01)
	TZD	18,842	1432	0.66 - 2.07 (1.08)	30,237	4735.92	1.00 (reference)	1.00 (reference)

[a] Cardiovascular death defined as death after excluding deaths in patients with metastatic cancer during follow-up

TABLE 4.18. Sensitivity analysis using Intent to treat approach - number of initiators[a], events (composite of nonfatal myocardial infarction, stroke and death), event rates and crude and adjusted hazard ratios for DPP-4i versus comparators in the entire population as well as subgroups based on prior cardiovascular disease

Comparison	**Treatment**	**Number of new-users**	**events**	**time to event in years interquartile range (median)**	**Total person-years**	**Incidence (per 100,000 person years)**	**Unadjusted HR (95% CI)** b	**Adjusted HR (95%CI)** c
DPP-4i vs SU	DPP-4i[d]	44,771	7358	0.86 - 3.09 (1.82)	95,204	7728.67	0.72 (0.70, 0.74)	0.84 (0.80, 0.85)
	SU	119,436	29924	0.87 - 3.59 (2.06)	277,496	10783.58	1.00 (reference)	1.00 (reference)
DPP-4i vs TZD	DPP-4i[d]	26,198	2762	0.83 - 3.18 (1.82)	56,429	4894.65	0.97 (0.93, 1.03)	0.98 (0.92, 1.03)
	TZD	18,842	3156	1.79 - 4.58 (3.32)	59,574	5297.61	1.00 (reference)	1.00 (reference)

DPP-4i - Dipeptidyl peptidase 4 inhibitors, TZD - thiazolidinediones, SU - sulfonylureas, CVD - cardiovascular disease

[a] Initiation defined as no dispensed prescriptions for DPP-4i or SU or TZD during the 6 months before initiation and filling a second prescription of the same drug/drug class within 6 months after the first prescription. Under the Intent to treat approach patients not censored for treatment changes.

[b] Hazard ratios and their 95 % confidence intervals from Cox proportional hazards models for the composite outcome with baseline treatment as the only independent covariate.

[c] Hazard ratios adjusted for variables in Table 1 and 2 using propensity score weighting (standardized to DPP-4i population).

[d] Number of DPP-4i initiators different in both cohorts because for the DPP-4i vs SU comparison, patients could be on any diabetes medication (including TZD) except for DPP-4i and SU during the washout period. Similarly for the DPP-4i versus TZD comparison, patients could be on any other drugs except DPP-4i and TZD. Further for the DPP-4i versus TZD analysis, patients with prevalent diagnosis of heart failure, hypertensive disease with heart failure, chronic kidney disease, edema and use of loop diuretics were excluded for all analyses.

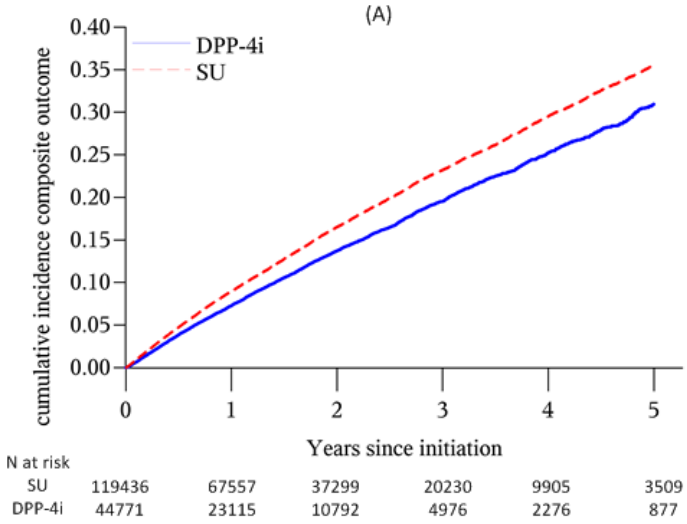
TABLE 4.19. Study population - DPP-4i versus SU				
	DPP-4i		**SU**	
	N	**% remaining**	**N**	**% remaining**
Prescription records meeting the new-use criteria after 6 months washout	128,327	100.0	195,237	100.0
Excluding records indicating prevalent use of the comparator drug during the washout period	62,715	48.9	173,288	88.8
Requiring at least 2 scripts of the same drug class within 180 days of initiation	51,150	39.9	140,245	71.8
Excluding records with patient age less than 66 years	47,393	36.9	124,409	63.7
Excluding records initiating DPP-SU dual therapy on the same day	46,261	36.0	123,235	63.1
Keeping only 1st record of patients meeting the new use criteria more than once	44,771	34.9	119,436	61.2

TABLE 4.20. Study population - DPP-4i versus TZD				
	DPP-4i		**TZD**	
	N	**% remaining**	**N**	**% remaining**
Prescription records meeting the new-use criteria after 6 months washout	128,327	100.0	65,561	100.0
Excluding records indicating prevalent use of the comparator drug during washout period	90,003	70.1	56,520	86.2
Requiring at least 2 scripts of the same drug class within 180 days of initiation	72,034	56.1	44,343	67.6
Excluding records with patient age less than 66 years	67,730	52.8	37,715	57.5
Excluding records initiating DPP-TZD dual therapy on the same day	67,154	52.3	37,113	56.6
Keeping only 1st record of patients meeting the new use criteria more than once	64,632	50.4	36,189	55.2
Excluding patients with baseline diagnosis of heart failure, hypertensive disease with heart failure, cardiomyopathy, arrhythmias/syncope/pacemaker, chronic kidney disease, edema and use of loop diuretics	26,198	20.4	18,842	28.7

FIGURE 4.1. Weighted cumulative incidence for the composite outcome (non-fatal myocardial infarction, stroke and all-cause mortality)

(A) Dipeptidyl peptidase-4 inhibitors (DPP-4i) versus sulfonylureas (SU)

(B) Dipeptidyl peptidase-4 inhibitors (DPP-4i) versus thiazolidinediones (TZD)



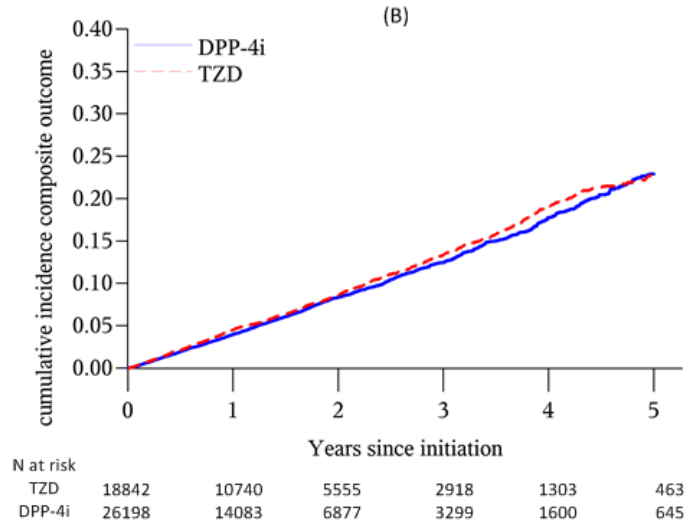


FIGURE 4.2. Weighted cumulative incidence for all-cause mortality (component of composite outcome): DPP-4i versus SU

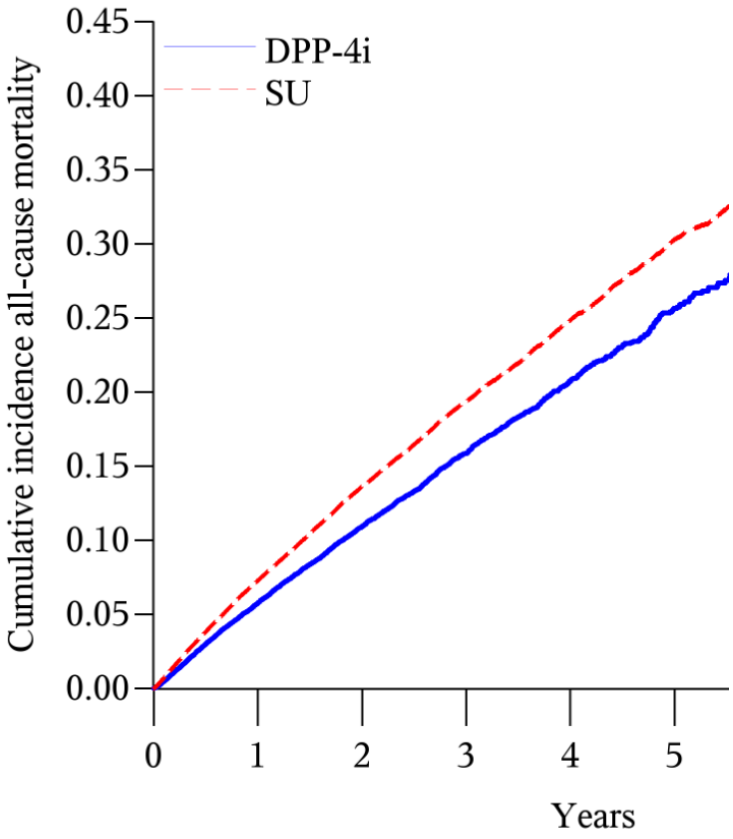


FIGURE 4.3. Weighted cumulative incidence for non-fatal myocardial infarction (MI) and stroke: DPP-4i versus SU and DPP-4i versus TZD

DPP-4i - dipeptidyl peptidase-4 inhibitors; SU – sulfonylureas; TZD – thiazolidinediones
(A) MI - DPP-4i versus SU, (B) Stroke - DPP-4i versus SU, (C) MI – DPP-4i versus TZD, (D) Stroke - DPP-4i versus TZD

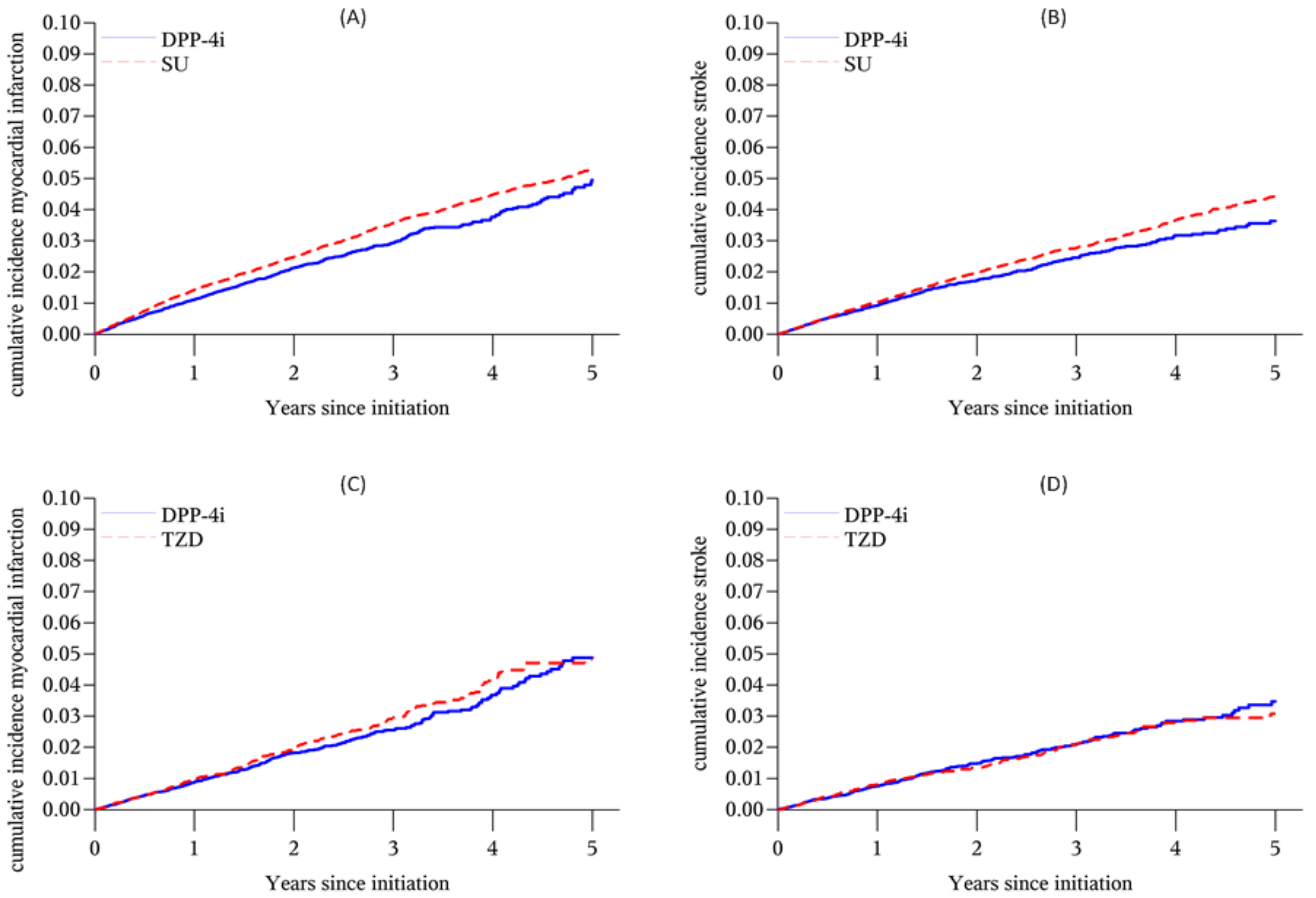


FIGURE 4.4. Weighted cumulative incidence curves for non-fatal myocardial infarction (MI) and stroke in subgroups based on prior CVD: DPP-4i vs SU

DPP-4i - dipeptidyl peptidase-4 inhibitors; SU – sulfonylureas

(A) MI - Subgroup with prior CVD; (B) MI – Subgroup without prior CVD;

(C) Stroke – Subgroup with prior CV; (D) Stroke - Subgroup without prior CVD

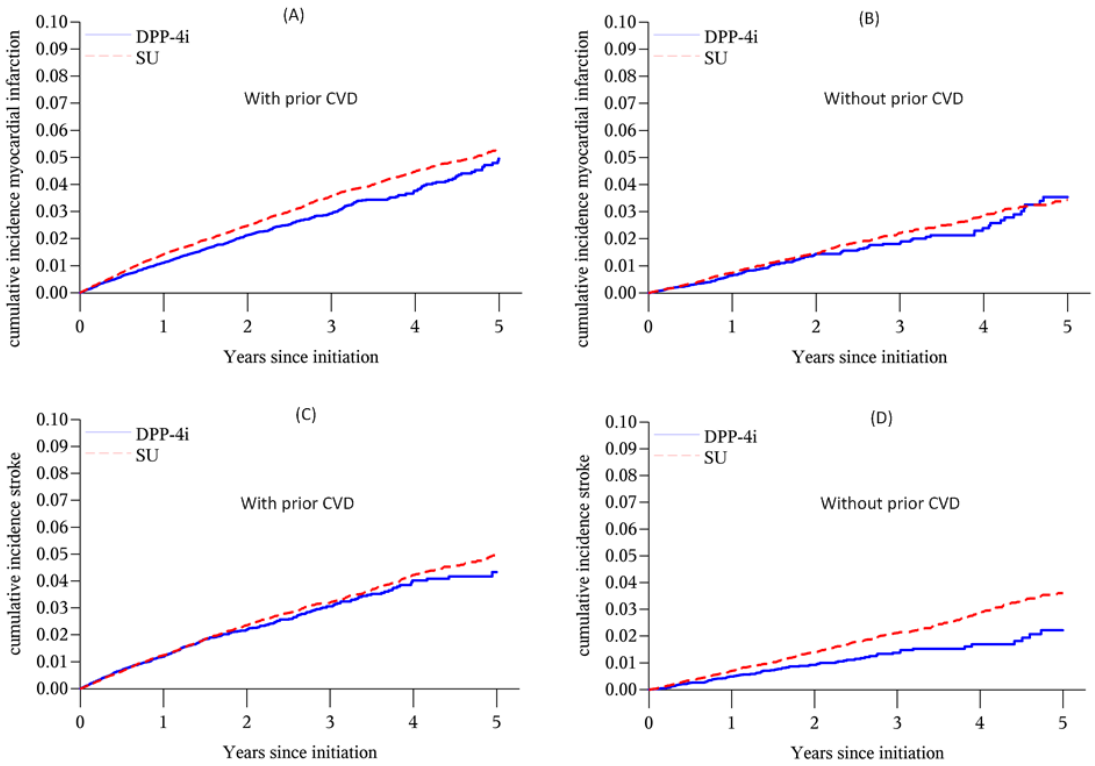
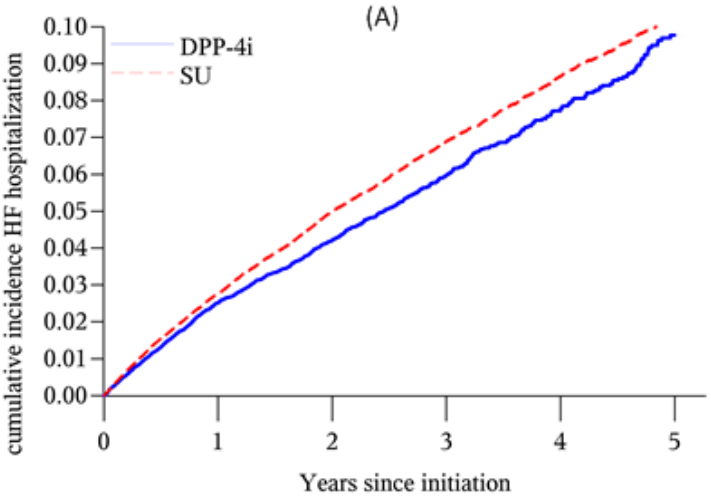


FIGURE 4.5. Weighted cumulative incidence curves for hospitalization for heart failure (HF) (A) DPP-4i vs SU; (B) DPP-4i vs TZD



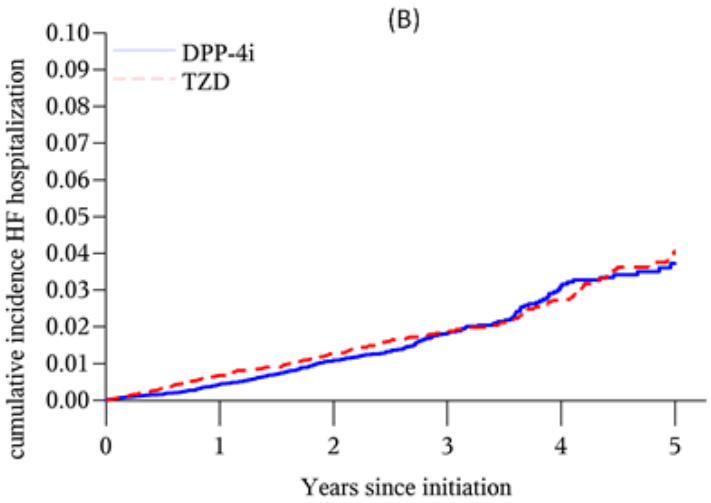
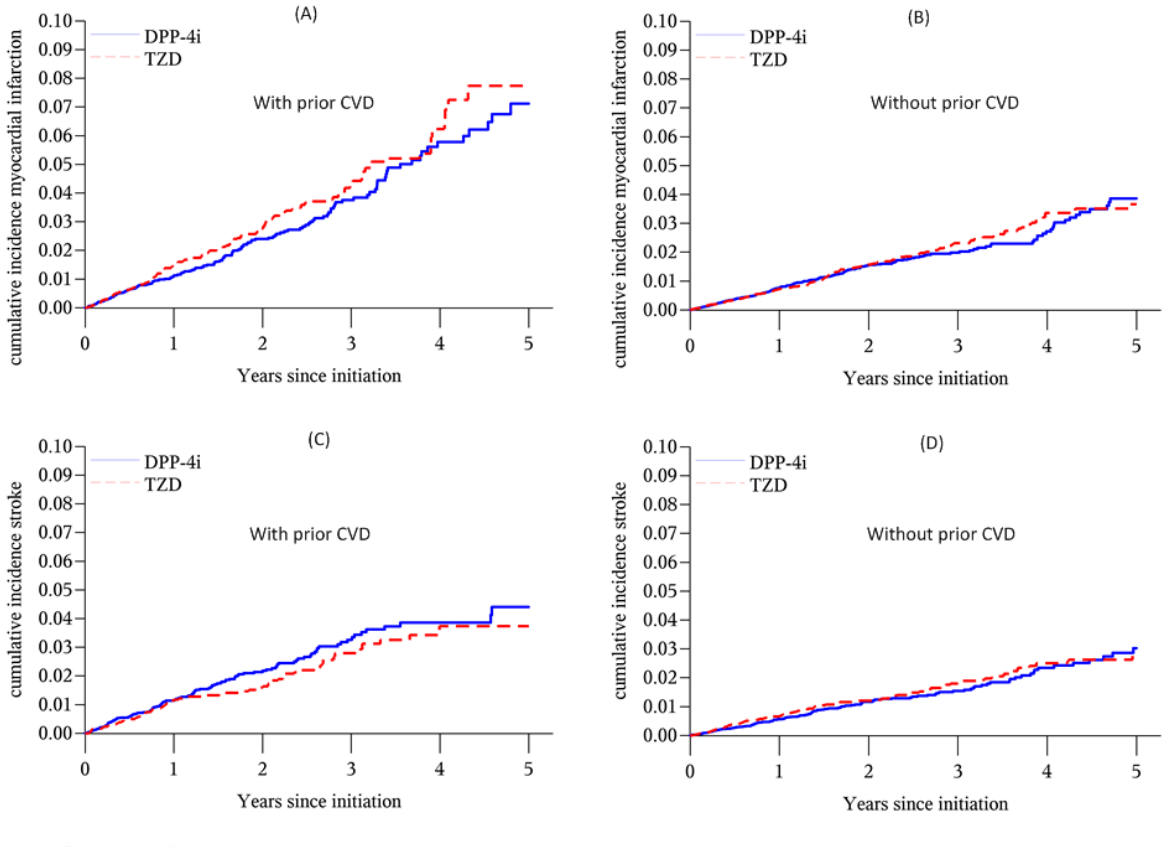


FIGURE 4.6. Weighted cumulative incidence curves for non-fatal myocardial infarction (MI) and stroke in subgroups based on prior CVD: DPP-4i vs TZD

DPP-4i - dipeptidyl peptidase-4 inhibitors; TZD - thiazolidinediones

(A) MI - Subgroup with prior CVD; (B) MI – Subgroup without prior CVD;

(C) Stroke – Subgroup with prior CV; (D) Stroke - Subgroup without prior CVD



CHAPTER 5. CALENDAR TIME AS AN INSTRUMENTAL VARIABLE IN ASSESSING THE RISK OF HEART FAILURE WITH ANTIHYPERGLYCEMIC DRUGS

I. Introduction

The prevalence of diabetes continues to increase around the world. In the United States about 9.3% of the adult population had diabetes in 2012.[61] The use of antihyperglycemic drugs has increased over the past few years and accounts for about 12% of the cost of diabetes.[13, 62] The antihyperglycemic drug market has seen introduction of several new drugs in the last decade – the most notable being an amylin analog, glucagon-like peptide-1 receptor agonists, dipeptidyl peptidase-4 inhibitors (DPP-4i) and sodium glucose transport protein-2 inhibitors. Along with introduction of new drugs, the last decade also saw emerging safety concerns about thiazolidinediones, an established class of antihyperglycemic drugs. In 2007, a meta-analysis raised concerns over the cardiovascular safety of rosiglitazone [63] after which it was pulled from the European market [46] and in September 2010 the Food and Drug Administration (FDA) announced that its use would be restricted in the US.[64] In 2010 the FDA also released a drug safety communication that it was evaluating data from an epidemiologic study examining whether pioglitazone increased the risk of bladder cancer.[65] In 2011 the FDA approved updates to the labelling pioglitazone products to include "an increased risk of bladder cancer in pioglitazone users."[66]

According to a recent analysis [13] of the IMS health data examining the market trends of antihyperglycemic drugs, the newly approved drugs, particularly the DPP-4i class quickly gained significant market share and was the most commonly prescribed new drug class by 2012. On the

other hand, the use of thiazolidinediones decreased by 2012 with the use of pioglitazone reaching only half its peak use in 2008 and negligible use of rosiglitazone.

In observational studies, the use of an active comparator greatly reduces the potential for confounding by indication and results in treatment cohorts that are comparable with respect to measured and unmeasured confounders.[67] DPP-4i and TZD are both generally used as second-line antihyperglycemic treatment, with intermediate (DPP-4i) to high (TZD) efficacy, low risk of hypoglycemia[14] and comparable costs before the generic form of pioglitazone was introduced in August 2012.[15] The DPP-4i and TZD drugs are therefore expected to be in therapeutic equipoise in a population excluding anyone with clear contraindications. This was demonstrated in our previous work comparing DPP-4i and TZD in a population of patients without heart failure and related conditions[68] where the measured baseline characteristics were well-balanced among initiators of DPP-4i and TZD which increases the confidence that unmeasured confounders were also well-balanced by design, although this cannot be proven. The increase in the use of DPP-4i with a simultaneous decrease in TZD use over a short period of time also provides an opportunity to use calendar time as an instrumental variable (IV) as suggested previously in other settings.[16]

While IV methods are generally indicated in studies where substantial unmeasured confounding is expected[28, 69, 70] (likely in cases where an active comparator does not exist), according to our knowledge, this is a unique setting where there exists both an active comparator and also a potential IV driven by the antihyperglycemic drug market which enables a direct comparison of the two methods and their estimators. Using hospitalization for heart failure (HF) as a positive control outcome (TZDs are known to increase the risk of HF whereas DPP-4i drugs sitagliptin and alogliptin do not)[3, 4, 71] we explored the use of calendar time as an IV and

compared this approach to an active comparator new-user study comparing DPP-4i versus TZD (actual treatment received) in sample of Medicare beneficiaries without a history of HF or related conditions.

II. Methods

Study Population

We used a 20% random sample of Medicare beneficiaries >65 years with fee-for-service Part A (hospital coverage), B (outpatient care) and D (dispensed prescription drugs) enrollment in at least one month during a calendar year from January 1, 2007 (2006 for Part A and B) to December 31, 2013. This dataset contains information about demographics, enrollment, diagnoses, procedures and prescription drugs for each enrollee.[49, 50]

From this population, we identified new-users of DPP-4i or TZD. Prevalent users of DPP-4i or TZD in the 6 months before initiation were excluded, but patients were allowed to be on other antihyperglycemic treatments during the washout period. Initiation was defined as the first prescription of the drug with the index date defined as the date of dispensing. Since TZDs are contraindicated in patients with existing heart failure [33.54] (which can lead to intractable confounding by contraindication if not properly accounted for), we further excluded patients with diagnoses of heart failure, hypertensive disease with heart failure and closely related conditions of cardiomyopathy, arrhythmias, chronic kidney disease, edema and use of loop diuretics. Patients needed to have at least 6 months of continuous Part D enrollment and at least 12 months parts A and B enrollment pre-index. To ensure that patients were actually started on the drugs, we restricted our cohorts to patients with a second prescription for the same drug class dispensed within 6 months after the index date.

Instrumental variable (IV)

An IV is an observed variable associated with the variation in exposure like randomized assignment in a clinical trial.[28] To be valid, an IV should be: 1) associated with the treatment, 2) unrelated to patient characteristics and 3) related to the outcome only through its association with the treatment (exclusion restriction). Additionally, in order to estimate the average treatment effect in the 'compliers' an additional assumption of 'monotonicity' (no 'defiers') should hold. Monotonicity requires that the instrument affects the treatment deterministically in one direction.[28]

We defined the IV as a binary variable (post versus pre-period) anchored around the time of the crossover of the drug initiation curves (June – September 2010) which was also the time when the FDA issued communications about restriction of rosiglitazone and concerns about the safety of pioglitazone (September 2010). In our study, 'compliers' are patients whose initial treatment is determined by the calendar time in which they initiated treatment – i.e, TZD if initiating before the crossover of the drug initiation curves and DPP-4i if initiating after the crossover. We identified an 'optimal' IV by evaluating the percentage of compliers, i.e, the strength of the instrument's effect on the received treatment.[16] We examined the instrument in relation to IV assumptions using measured covariates, falsification tests and expert knowledge. The binary measure of calendar time was decided before examining the effect estimates.

Outcome

The outcome used in this study was hospitalization for HF defined using *International Classification of Diseases, Ninth Revision* (*ICD-9*) code 428.xx in the primary position in inpatient claims. This definition has a specificity >98% but a very low sensitivity of 21% in a Medicare population.[32] TZDs are known to be associated with an increased risk of heart failure

[34,71] and while saxagliptin is suspected to be associated with an increased risk of heart failure, recently reported placebo-controlled trials indicate no increased incidence of HF with DPP-4i drugs sitagliptin and alogliptin (which is >80% of our cohort).[2, 3, 35] HF hospitalization is therefore a positive control outcome in our study enabling the comparison of performance of the propensity score weighting method comparing the actual treatments initiated and IV analysis comparing levels of the calendar time instrument. Patients were followed up from the second fill date till the earliest of the following – outcome (HF hospitalization), death, end of enrollment or 2 years after the index date.

Confounding control and analysis

We used weighted Kaplan Meier methods to estimate the 1- and 2-year risk difference (RD) for HF hospitalization under the two analytic approaches, i.e, comparing the actual treatment initiated (DPP-4i vs TZD) and levels of the calendar time IV ('post' versus 'pre' periods). Under the 'treatment received' approach, propensity scores (PS) were used to control for measured confounding. Using variables listed in table 1 measured before initiation, we predicted the probability for initiating DPP-4i versus TZD for each patient (the PS) [36] using a logistic regression model without including the year of initiation in the model as it can increase the variability of effect estimates and can increase bias in case of unmeasured confounding if it acts as an instrument.[40, 41, 36] We then assigned a weight of 1/PS to DPP-4i and a weight of (1/(1-PS)) to TZD and stabilized both groups by marginal prevalence of the treatment actually received. [42] Such inverse probability of treatment (IPTW) weighting creates a pseudo-population in which the association between covariates and treatment is removed by the weighting described above. Under the assumption of no unmeasured confounding this method allows us to contrast two scenarios: "what would have happened if the entire population initiated DPP-4i' versus

‘“what would have happened if the entire population initiated TZD’.[42] Using IPTW weighted Kaplan Meier curves we then estimated the RDs for HF hospitalization using treatment as the strata variable.

The effect of the IV levels on HF hospitalization was observed by the IV estimates of RD scaled by the strength of the instrument to estimate the average treatment effect among the ‘compliers’. Next we generated covariate-adjusted IV estimates of the RD using weighted Kaplan-Meier methods analogous to the IPTW weighted method outlined above and scaled this by the strength of the instrument. Under this approach, we first predicted the probability (PP) of the instrument as a function of baseline covariates and assigned a weight of 1/PP to the ‘post’ period and 1/(1-PP) to the ‘pre’ period and stabilized the weights by the marginal prevalence of the instrument.

These methods are appreciably different and results apply to different patient populations (IV estimates apply to the marginal patients while IPTW methods apply to the whole population). Propensity score models used the full study population and we performed sensitivity analyses in the reduced instrumental variable cohort to evaluate selection differences that may have been introduced based on instrumental variable exclusions. [16] Next we performed a sensitivity analysis restricting the study cohort to patients who initiated DPP-4i or TZD between July 2007 and December 2011, with follow-up allowed through December 2013. This approach excluded anyone who initiated therapy during or after January 2012, who would not have the potential for a 2-year follow-up before the end of the study (and therefore potentially violate the assumption of ‘non-informative’ censoring i.e., the assumption that patients who drop out of the study should do so for reasons unrelated to HF).[43] Finally, we performed additional analyses with different

definitions of the instrument (IV defined using a cut-point and by comparing periods with maximum separation).

III. Results

As seen in FIGURE 5.1, calendar time greatly affected treatment initiation. The 'optimal' binary IV compared the periods from October 2010 through December 2013 (hereafter referred to as 'post'; N = 22,696) with the period from January 2008 to May 2010 ('pre'; N = 20,283). This definition excludes patients in the region where the curves intersect (June 2010 through September 2010) which was the time around which the drug initiation curves were close together and also when the FDA issued drug safety communications about TZDs. With this IV definition, the DPP-4i treatment rates were 78% and 38% in the post and pre periods and the strength of the IV was 40%.

TABLE 5.1 presents the baseline covariates comparing the actual treatment initiated (DPP-4i versus TZD) and the levels of the IV (post versus pre). While the measured covariates are well balanced by design across the treatment received (DPP-4i versus TZD), the balance was slightly better across the levels of the IV as illustrated by the average standardized absolute mean difference (average SAMD, 4.5 vs 3.0%). This indicates that the instrument is independent of the measured risk factors of the outcome and potentially also unmeasured covariates. The IV particularly improved the balance for race, diabetes complications and health seeking behaviors like statin use, influenza vaccinations, and lipid panels. For example, the SAMD for lipid panels comparing DPP-4i versus TZD was 11.8% which reduced to 5.8% across the levels of the IV.

TABLE 5.2 shows the number of initiators, events, median time on treatment and incidence rates for HF hospitalization across levels of the IV and treatment initiated. In the full population comparing DPP-4i versus TZD, the crude incidence rate per 100,000 was 563 in

DPP-4i and 676 in TZD initiators. The 1- and 2-year risks for HF hospitalization were 0.5% and 1.2% for DPP-4i and 0.7% and 1.3% for TZD respectively (FIGURE 5.2). In the reduced IV population, crude incidence rate per 100,000 person-years was 509 in the post period and 695 in the pre period. The 1- and 2-year risks for HF hospitalization were 0.4% and 1.1% in the post period and 0.7% and 1.4% in the pre period respectively.

FIGURE 5.3 shows the 1- and 2-year RDs per 100 patients for HF hospitalization obtained using different analyses (TABLE 5.3). The crude RD was not very different from the IPTW weighted RD obtained by comparing DPP-4i versus TZD (expected because of the relative balance of baseline covariates by design). The IV estimates of the RDs scaled by the strength of the instrument were consistent overall with the IPTW weighted RDs indicating a protective effect with the DPP-4i but relatively less precise and farther from the null. The 2-year RD per 100 patients obtained using the IV analysis scaled by IV strength was -0.62 (-0.99, -0.25) indicating that of the 1000 compliant patients treated with DPP-4i, 6 additional patients had no HF hospitalization till 1 year compared with those treated with TZD. Covariate adjusted IV estimates of the RD were virtually similar (FIGURE 5.3) (TABLE 5.3).

Sensitivity analyses in a population restricted to patients initiating on or before December 2011, analyses using IPTW weighting in the reduced IV population (TABLE 5.4) and analyses using different definitions of the instrumental variable yielded similar results (FIGURE 5.4, FIGURE 5.5).

IV. Discussion and Conclusions

In this study we compared the performance of two analytic methods - IPTW weighting comparing DPP-4i versus TZD initiators and the IV analysis comparing initiators of either DPP-4i or TZD during later years with initiators of these drug classes in earlier years in assessing the

risk of HF hospitalization, a positive control outcome, in DPP-4i versus TZD initiators. In a population of patients without evidence of prior HF or related conditions, DPP-4i and TZD, both second-line antihyperglycemic drugs are in therapeutic equipoise and measured patient characteristics were well balanced across the two cohorts as seen in table 1. This also indicates lesser chance of unmeasured confounding, although this cannot be ruled out completely. While IV methods are generally employed in cases where substantial unmeasured confounding is expected, according to our knowledge, this is a unique situation where both methods were expected to be equally valid and this was confirmed by our results. Both approaches indicated a decreased risk of HF hospitalization with DPP-4i relative to TZD, an effect that was consistent across a number of sensitivity analyses.

The absolute 1-year risks of HF hospitalization reported in the randomized trials on sitagliptin (>80% of our DPP-4i cohort) and pioglitazone (>90% of our TZD cohort) are approximately 1% and 2% respectively.[4, 71] These trials compared sitagliptin and pioglitazone with placebo and found no increased risk with sitagliptin and 0.5% increased risk with pioglitazone relative to placebo at 1 year, but there are no data about RDs comparing DPP-4i and TZD. In our study, the ICD-9 code based definition of HF had a high specificity, but a low sensitivity of 21%[32] which led to an underestimation of absolute HF risks and therefore RDs with both the IPTW and IV methods. While we could have chosen to estimate relative risks (expected to be unbiased because of near-perfect specificity of the HF definition), IV methods used to derive relative measures require additional strong assumptions about the homogeneity of risks within the levels of the treatment and measured covariates.[28, 72] These assumptions are often implicit in IV methods based on structural equation regression models that yield estimates of absolute measure of effect (RD) and this is what we used in this study.[28, 72]

While the RDs obtained using both analytic approaches generally agreed, the IV estimates had wider confidence intervals as expected and the magnitudes of the RDs differed because the two analytic methods apply to potentially different populations and are based on different assumptions. The IPTW weighted RD is the average treatment effect in the entire population indicated for treatment. However, the scaled RD from the IV analysis is the average treatment effect in the 40% compliers from the reduced IV population. Scaling of the IV estimate of the RD involves dividing the effect of the instrument on the outcome (measured as risk difference) by the IV strength (rates of DPP-4i treatment in the post minus pre periods). The weaker the instrument, the larger the scaled estimate gets relative to the unscaled RD, potentially multiplying any biases present in the unscaled RD.[28] In a sensitivity analysis we defined the instrument by a cut-point (before and after July 2010) which made this a weaker instrument (difference in DPP-4i treatment rates post – pre = 25%), and the scaled 1-year RD was -0.73 (-1.74, 0.27) which is higher in magnitude compared to the RD estimate scaled with a stronger instrument in the main analysis, but in the same direction. In another sensitivity analysis we compared the periods with maximum separation (December 2011 – December 2012 versus June 2008 – June 2009). This definition of the IV led to the IV strength of 50% (stronger than the main analysis) and as expected the scaled 1-year RD was closer to the unscaled estimate (-0.45, CI:-0.92, 0.05). Thus, scaling could explain some of the differences in magnitude of the IV estimates of RDs compared to the IPTW weighed RDs seen in the main analysis. We compared the characteristics of the compliers with that of the full population and found that while the populations were generally similar, the compliers were slightly less likely to be on other antidiabetic drugs like metformin and sulfonylureas, but the difference was not big enough to explain our results (TABLE 5.5).

It is also possible that instrumental variable assumptions are violated. While we were able to empirically observe that the instrument strongly affected the treatment initiation between 2008 and 2013 (assumption 1), the assumption of calendar time being unrelated to the outcome through other mechanisms may be violated. A recent study reports a decline in the rates of HF hospitalizations from 1999 to 2011 among Medicare beneficiaries with and without a history of heart failure[73] which could be due to several reasons including increases in the use of statins, decreased smoking and increasing outpatient management of HF. The Hospital Readmissions Reduction Program (HRRP) started in late 2012 requires CMS to penalize hospitals with excess readmissions for certain conditions including HF, which could be somewhat responsible to the decreasing trend in HF hospitalizations, but this is not expected to affect earlier years of the data used for this study.[74] To minimize the effect of time trends, we conducted a sensitivity analysis using a dichotomous cut-point of calendar time (described above) and estimated 1-year RD comparing the 12 months after versus before the cut-point but the results pointed in the same direction, although the magnitude was different. The assumption of IV being unrelated to patient characteristics is also not empirically verifiable, but it is upheld by the improvement in the covariate balance across the levels of the IV compared to the balance across the levels of the treatment initiated. The assumption of 'no defiers' also seems to be reasonable in this study as it is unlikely that a patient would initiate DPP-4i before June 2010 – September 2010 (region of crossover of the curves) while an identical patient would initiate TZD in the 'post' period in spite of the warnings about adverse effects associated with TZD. The generic form of pioglitazone was available starting August 2012, but this would potentially affect the last year of our study and not the earlier years. Given the difference in efficacy of DPP-4i versus TZD (intermediate versus high), it is possible that patients with high HbA1C initiated TZD in the post period, but

based on expert knowledge that is not expected to be a significant concern. There were no events that would preclude DPP-4i initiator in the 'pre' period from receiving DPP-4i in the 'post' period, especially given the positive reports about the lack of serious adverse events with these agents in the past few years.[6, 35, 51]

Some additional caveats should be considered. First, 'scaling by strength of the instrument' may have implications for bias amplification due to covariate imbalances across the levels of the instrument. While we compared the covariate balance achieved across the 'treatments initiated' to the balance across the levels of the IV as commonly recommended[28], this approach misses a key component (the scaling factor) of the relative bias comparing IV and non-IV approaches.[75] As proposed by Jackson, we calculated the bias component for the two analytic approaches - the prevalence difference for the treatments received and the prevalence difference multiplied by the scaling factor (1/0.40) for the IV levels (TABLE 5.6).[75] With this approach the differences across the IV levels somewhat increased for variables like younger age groups, statins and baseline use of sulfonylureas, indicating that not adjusting for these variables could bias our results. However our adjusted and unadjusted IV estimates are only slightly different indicating no significant concern for confounding by these variables in our main analysis. However, weak instruments (i.e., scaling factor closer to 0) can lead to greater imbalance across the levels of the IV and not adjusting for these variables might lead to bias. Second, our first-treatment carried forward / intent-to-treat (ITT) analysis approach could lead to exposure misclassification as it does not take into account treatment changes. However, the IV estimator is a measure of the association between the treatment assignment (i.e., treatment *intention*) and the outcome. Therefore, using a conceptually uniform ITT approach for the IPTW analysis comparing DPP-4i versus TZD initiation would facilitate direct comparison of the RDs

obtained by the two methods. Third, death is a competing event in the older population and not accounting for competing risks could theoretically overestimate the risk of HF. However, analyses with and without accounting for competing risks were virtually similar in a previous study using the same population indicating that competing risk by death is not a serious concern for this study. [68] Fourth, our use of the term compliance/compliers refers to prescriber behaviors rather than patient behaviors as in an RCT. Finally, it could be argued that unmeasured confounding exists in the DPP-4i versus TZD comparison by unmeasured factors like body mass index (BMI) or smoking. However, we used chronic obstructive pulmonary disease (COPD) as a proxy for smoking in our PS models and previously found using data from the Medicare Current Beneficiary Survey that BMI is not likely to affect the choice of initiation of DPP-4i versus TZD in these patients which reduces the concern of confounding by these variables.[60]

Strengths of our study included the ability to identify a strong instrument of calendar time driven by the dynamics of the antihyperglycemic drug market over a short period of time. This created a setting for a 'natural experiment' that creates an allocation of exposure similar to that of a randomized experiment. Use of an active comparator created a setting of treatment equipoise between DPP-4i and TZD as described before. It is often conservatively suggested to avoid IV methods when the assumptions are likely to be violated because relatively minor violations of the assumptions may lead to large biases[69], which can particularly be a problem when there is no valid comparison analytic method. However, in our setting, the IV and IPTW methods address different biases and are based on different assumptions some of which cannot be directly quantified; but the use of a positive control outcome and observing consistent results with either method implies that both methods may be equally valid, or supplement each other at the very least. In summary the use of calendar time as an IV in settings where real-world market

dynamics lead to profound changes in preferred treatments is worth consideration as previously suggested in other settings.[16]

V. Tables and Figures

TABLE 5.1. Characteristics of patients by treatment initiated and levels of the instrumental variable						
	By levels of the instrument			**BY actual treatment initiated**		
	Post (N = 22,696) Oct 2010 to Dec 2013	**Pre (N = 20,283) Jan 2008 to May 2010**	**SAMD**	**DPP-4i (N = 26,198)**	**TZD (N = 18,842)**	**SAMD**
	%	**%**	**%**	**%**	**%**	**%**
TZD	22.5	62.4				
DPP-4i	77.5	37.6				
Patient demographics						
66 to 75 years old	66.2	64.2	4.2	64.8	65.7	2.0
76 to 85 years old	27.6	29.6	4.3	28.9	28.3	1.1
86 years and above	6.1	6.2	0.3	6.4	5.9	1.8
Male	41.4	41.3	0.2	39.7	43.6	8.0
White	73.7	74.0	0.7	75.5	71.4	9.4
Black	9.2	9.9	2.3	8.9	10.5	5.3
Other	17.1	16.1	2.7	15.6	18.2	6.8
Baseline comorbidities						
Diagnosis of acute myocardial infarction	0.8	0.7	1.5	0.8	0.5	3.9
Diagnosis of old myocardial infarction	2.2	1.8	2.9	2.2	1.6	4.5
Angina	3.2	3.0	1.2	3.5	2.6	5.3
Diagnosis of Stroke or TIA	4.8	3.9	4.5	4.7	4.0	3.2
Other acute and subacute forms of ischemic heart disease	1.4	1.2	1.7	1.6	1.0	4.6
Obesity diagnosis	0.9	0.8	1.0	1.0	0.6	3.8
Diabetic nephropathy	3.0	2.8	1.3	3.0	2.8	0.8
Diabetic neuropathy	15.2	12.8	7.1	15.3	12.7	7.4
Diabetic retinopathy	14.4	12.4	5.9	13.8	12.9	2.7
Bone fractures	4.4	3.6	3.8	4.1	3.9	0.6
Chronic obstructive pulmonary disease	11.2	10.2	3.2	11.0	10.4	1.9
Cancer (Non-Skin)	14.9	13.0	5.6	15.1	12.4	7.8
Dementia	4.9	4.7	0.9	4.9	4.6	1.2
Foot ulcers	7.1	5.9	4.9	7.0	5.8	4.7
Inflammatory GI diseases	0.7	0.6	0.7	0.7	0.7	0.2

Liver disease	5.3	4.3	4.4	5.2	4.3	4.5
Pancreatitis	0.8	0.7	0.8	0.8	0.7	0.5
Alcoholism	0.1	0.1	0.6	0.1	0.1	0.6
Aortic aneurysm	1.5	1.3	1.6	1.6	1.2	3.1
Baseline medication use						
Metformin	71.3	73.3	4.3	73.9	70.5	7.4
Sulfonylureas	51.2	54.4	6.4	52.0	54.1	4.2
Glucagon-like peptide-1 agonists	1.7	1.8	0.4	1.8	1.6	1.2
Long acting insulin	11.9	10.9	3.0	11.3	11.5	0.8
Short acting insulin	4.3	4.2	0.5	4.1	4.4	1.2
Statins	70.2	65.5	10.1	70.2	64.7	11.8
Angiotensin converting enzyme inhibitors	39.7	40.1	0.8	38.5	41.9	6.9
Angiotensin receptor blockers	16.6	15.6	2.5	17.3	14.4	7.8
Beta blockers	34.9	35.1	0.4	36.7	32.4	9.0
Calcium channel blockers	27.1	25.3	4.1	27.1	25.2	4.3
Other diuretics	39.3	40.8	2.9	40.5	39.4	2.2
Health Care utilization						
Blood tests	8.5	9.4	3.2	9.2	8.6	2.3
Influenza vaccinations	55.9	52.0	7.8	56.2	51.0	10.4
Lipid panels	88.5	86.5	5.9	89.2	85.3	11.8
Mammogram	23.0	23.6	1.3	24.9	21.1	9.0
Average SAMD			3.0			4.5

Instrumental variable definition anchored around the time of the FDA drug safety communications about thiazolidinediones and the crossing of the drug initiation curves. This definition excludes patients in the region where the curves intersect (June 2010 through September 2010) where the curves were close together (N = 2061, 4.5% of the total population of 45,040).
DPP-4i - Dipeptidyl peptidase 4 inhibitors, TZD - thiazolidinediones, SAMD: standardized absolute mean difference.

TABLE 5.2. Number of new-users, heart failure hospitalizations, time on treatment, total person time and incidence for comparisons of levels of the instrument and treatment initiated

Comparison	**Level**	**Number of new-users**	**events**	**Duration of follow-up* in years interquartile range (median)**	**Total person-years***	**Incidence (per 100,000 person years)**
Post vs Pre^	Post	22,696	146	0.64 - 2.00 (1.36)	28,661	509.40
	Pre	20,283	255	2.00 - 2.00 (2.00)	36,682	695.16
DPP-4i vs TZD	DPP-4i	26,198	210	1.79 - 2.00 (2.00)	37,252	563.73
	TZD	18,841	215	1.68 - 2.00 (2.00)	31,791	676.29

DPP-4i - Dipeptidyl peptidase 4 inhibitors, TZD - thiazolidinediones

^Post - Oct 2010 to Dec 2013, Pre - Jan 2008 to May 2010

*Under the first treatment carried forward/intent-to-treat analysis patients were followed up from the second prescription of the initiated drug to earliest of the outcome occurrence, death or 2-years after initiation without accounting for treatment changes during follow-up

TABLE 5.3. Risk differences using instrumental variable and propensity score weighted approaches in the main analysis

	1-year Risk difference			**2-year Risk difference**		
	RD	**LCL**	**UCL**	**RD**	**LCL**	**UCL**
Unadjusted	-0.2	-0.34	-0.02	-0.17	-0.34	0.02
IV estimator unscaled	-0.24	-0.39	-0.10	-0.35	-0.58	-0.10
IV estimator scaled by IV strength of 40%	-0.62	-0.99	-0.25	-0.88	-1.46	-0.25
Adjusted IV estimator scaled by IV strength of 40%	-0.58	-1.00	-0.19	-0.76	-1.39	-0.12
PS IPTW weighted	-0.20	-0.33	-0.05	-0.18	-0.30	-0.03

PS - Propensity score, IPTW - Inverse probability of treatment weights, IV - Instrumental variable, RD - risk difference. LCL and UCL - lower and upper limits of the 95% confidence intervals respectively

TABLE 5.4. Risk differences with sensitivity analyses						
	1-year Risk difference			**2-year Risk difference**		
	RD	**LCL**	**UCL**	**RD**	**LCL**	**UCL**
Sensitivity analysis: IPTW weighted risk difference for DPP-4i vs TZD in the reduced IV population (N =)						
PS IPTW weighted in reduced IV population	-0.19	-0.26	0.02	-0.16	-0.32	0.15
Sensitivity analysis: IV analysis in a population where patients without potential for a 2 year follow-up, i.e. patients initiating therapy during or after January 2012 were excluded (N = 29,724: post = 9,442, pre = 20,282, IV strength 30%)						
IV estimator scaled by IV strength modified population	-0.76	-1.37	-0.11	-1.06	-1.93	0.15
PS: Propensity score, IPTW - Inverse probability of treatment weights, IV - Instrumental variable						

TABLE 5.5. Comparison of characteristics of compliers with full population						
	Full population (N=45,040)	**IV strength in subgroup (%)**	**Ratio of IV strength = IV strength in subgroup/IV strength overall**	**% of patients in subgroup = # in subgroup/IV population**	**Prevalence of subgroup in compliers = ratio of IV strength*% of patients in subgroup**	**Percent difference Compliers versus full population**
	%	%		%		
66 to 75 years old	65.19	39.87	1.00	65.29	65.41	0
76 to 85 years old	28.64	39.7	1.00	28.55	28.48	0
86 years and above	6.17	42.67	1.07	6.16	6.60	0
Male	41.31	39.82	1.00	41.34	41.36	0
White	73.76	38.71	0.97	73.82	71.79	2
Black	9.58	43.65	1.10	9.53	10.45	1
Other	16.66	43.16	1.08	16.65	18.06	1
Diagnosis of acute myocardial infarction	0.71	44.06	1.11	0.72	0.80	0
Diagnosis of old myocardial infarction	1.96	34.24	0.86	1.98	1.71	0
Angina	3.13	37.25	0.94	3.11	2.91	0
Diagnosis of Stroke or TIA	4.38	40.96	1.03	4.39	4.51	0
Other acute and subacute forms of ischemic heart disease	1.33	39.1	0.98	1.34	1.32	0
Obesity diagnosis	0.83	30.18	0.76	0.83	0.63	0

Diabetic nephropathy	2.90	37.41	0.94	2.91	2.73	0
Diabetic neuropathy	14.19	35.98	0.90	14.07	12.72	1
Diabetic retinopathy	13.46	39.72	1.00	13.42	13.39	0
Bone fractures	4.00	40.58	1.02	4.00	4.08	0
Chronic obstructive pulmonary disease	10.76	41.48	1.04	10.74	11.20	0
Cancer (Non-Skin)	14.00	39.58	0.99	13.98	13.90	0
Chronic lung diseases	10.57	39.24	0.99	10.57	10.42	0
Dementia	4.77	40.54	1.02	4.78	4.87	0
Depression	10.97	41.55	1.04	10.96	11.45	0
Foot ulcers	6.48	40.8	1.03	6.50	6.66	0
Inflammatory GI diseases	0.65	34.03	0.86	0.66	0.56	0
Liver disease	4.82	36.89	0.93	4.82	4.47	0
Other renal disorders	2.03	39.92	1.00	2.03	2.04	0
Pancreatitis	0.73	31.2	0.78	0.74	0.58	0
Peripheral vascular disease (claudication)	9.01	38.43	0.97	8.95	8.64	0
Proteinuria	2.12	38.05	0.96	2.13	2.03	0
Urinary tract infection/cystitis	17.41	42.95	1.08	17.39	18.77	1
Alcoholism	0.10	33.64	0.85	0.10	0.08	0
Aortic aneurysm	1.45	35.24	0.89	1.44	1.28	0
Metformin	72.47	38.75	0.97	72.26	70.35	2
Sulfonylureas	52.89	38.06	0.96	52.70	50.40	2
Glucagon-like peptide-1 agonists	1.72	20.92	0.53	1.73	0.91	1
Long acting insulin	11.39	39.49	0.99	11.41	11.32	0
Short acting insulin	4.22	42.65	1.07	4.22	4.52	0

Anticholinergic drugs	1.51	37.12	0.93	1.50	1.40	0
Statins	67.85	40	1.01	67.94	68.28	0
Angiotensin converting enzyme inhibitors	39.95	40.72	1.02	39.88	40.80	1
Angiotensin receptor blockers	16.06	37.94	0.95	16.13	15.37	1
Beta blockers	34.93	37.4	0.94	35.00	32.89	2
Calcium channel blockers	26.27	39.15	0.98	26.27	25.84	0
Other diuretics	40.04	38.82	0.98	40.00	39.02	1
Blood tests	8.93	38.73	0.97	8.88	8.64	0
Influenza vaccinations	53.99	38.25	0.96	54.04	51.93	2
Lipid panels	87.60	39.54	0.99	87.54	86.97	1
Mammogram	23.29	38.44	0.97	23.26	22.46	1
Pap Smear	6.65	37.23	0.94	6.62	6.19	0

TABLE 5.6. Bias components based on covariate imbalances calculated for the 'treatment received' and 'instrumental variable' approaches						
	By levels of the instrument			**BY actual treatment initiated**		
	Post (N = 22,696) Oct 2010 to Dec 2013	**Pre (N = 20,283) Jan 2008 to May 2010**	**Prevalence difference*1/ 0.40**	**DPP-4i (N = 26,198)**	**TZD (N = 18,842)**	**Prevalence difference**
	%	**%**	**%**	**%**	**%**	**%**
TZD	22.5	62.4				
DPP-4i	77.5	37.6				
Patient demographics						
66 to 75 years old	66.2	64.2	5	64.8	65.7	-0.9
76 to 85 years old	27.6	29.6	-5	28.9	28.3	0.6
86 years and above	6.1	6.2	-0.25	6.4	5.9	0.5
male	41.4	41.3	0.25	39.7	43.6	-3.9
white	73.7	74	-0.75	75.5	71.4	4.1
black	9.2	9.9	-1.75	8.9	10.5	-1.6
other	17.1	16.1	2.5	15.6	18.2	-2.6
Baseline comorbidities						
Diagnosis of acute myocardial infarction	0.8	0.7	0.25	0.8	0.5	0.3
Diagnosis of old myocardial infarction	2.2	1.8	1	2.2	1.6	0.6
Angina	3.2	3	0.5	3.5	2.6	0.9
Diagnosis of Stroke or TIA	4.8	3.9	2.25	4.7	4	0.7

Other acute and subacute forms of ischemic heart disease	1.4	1.2	0.5	1.6	1	0.6
Obesity diagnosis	0.9	0.8	0.25	1	0.6	0.4
Diabetic nephropathy	3	2.8	0.5	3	2.8	0.2
Diabetic neuropathy	15.2	12.8	6	15.3	12.7	2.6
Diabetic retinopathy	14.4	12.4	5	13.8	12.9	0.9
Bone fractures	4.4	3.6	2	4.1	3.9	0.2
Chronic obstructive pulmonary disease	11.2	10.2	2.5	11	10.4	0.6
Cancer (Non-Skin)	14.9	13	4.75	15.1	12.4	2.7
Dementia	4.9	4.7	0.5	4.9	4.6	0.3
Foot ulcers	7.1	5.9	3	7	5.8	1.2
Inflammatory GI diseases	0.7	0.6	0.25	0.7	0.7	0
Liver disease	5.3	4.3	2.5	5.2	4.3	0.9
Pancreatitis	0.8	0.7	0.25	0.8	0.7	0.1
Alcoholism	0.1	0.1	0	0.1	0.1	0
Aortic aneurysm	1.5	1.3	0.5	1.6	1.2	0.4
Baseline medication use						
Metformin	71.3	73.3	-5	73.9	70.5	3.4
Sulfonylureas	51.2	54.4	-8	52	54.1	-2.1
Glucagon-like peptide-1 agonists	1.7	1.8	-0.25	1.8	1.6	0.2
Long acting insulin	11.9	10.9	2.5	11.3	11.5	-0.2
Short acting insulin	4.3	4.2	0.25	4.1	4.4	-0.3
Statins	70.2	65.5	11.75	70.2	64.7	5.5
Angiotensin converting enzyme inhibitors	39.7	40.1	-1	38.5	41.9	-3.4

Angiotensin receptor blockers	16.6	15.6	2.5	17.3	14.4	2.9
Beta blockers	34.9	35.1	-0.5	36.7	32.4	4.3
Calcium channel blockers	27.1	25.3	4.5	27.1	25.2	1.9
Other diuretics	39.3	40.8	-3.75	40.5	39.4	1.1
Health Care utilization						
Blood tests	8.5	9.4	-2.25	9.2	8.6	0.6
Influenza vaccinations	55.9	52	9.75	56.2	51	5.2
Lipid panels	88.5	86.5	5	89.2	85.3	3.9
Mammogram	23	23.6	-1.5	24.9	21.1	3.8

Instrumental variable definition anchored around the time of the FDA drug safety communications about thiazolidinediones and the crossing of the drug initiation curves. This definition excludes patients in the region where the curves intersect (June 2010 through September 2010) where the curves were close together (N = 2061, 4.5% of the total population of 45,040).
DPP-4i - Dipeptidyl peptidase 4 inhibitors, TZD - thiazolidinediones
Calculation of bias component based on the method by Jackson et. al.

FIGURE 5.1. Initiation of dipeptidyl peptidase 4 inhibitors (DPP-4i) versus thiazolidinediones (TZD) across calendar time

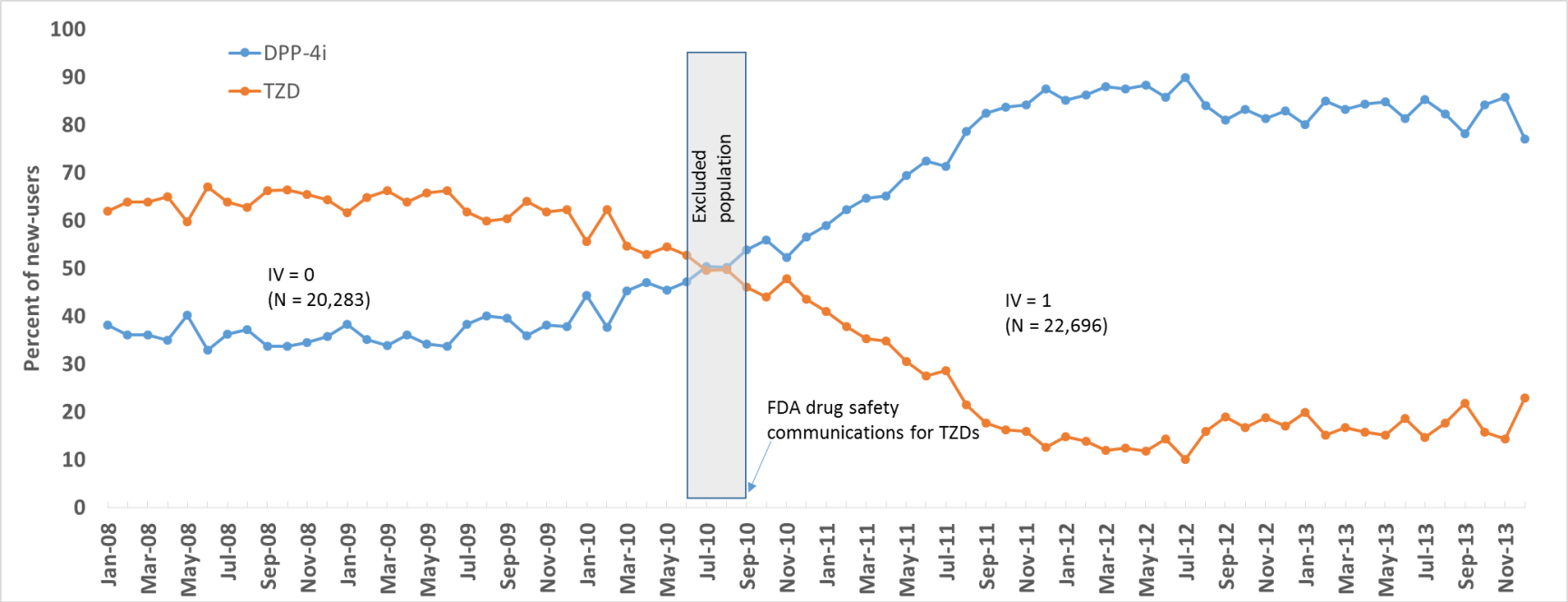


FIGURE 5.2. Probability of being event-free

(A) By treatment initiated (N = 45,040) and

(B) By levels of the calendar time instrument (N = 42,979). Patient assignment to instrumental variable category is based on month and year treatment was initiated. Oct 2010 to Dec 2013 (post) versus Jan 2008 to May 2010 (pre)

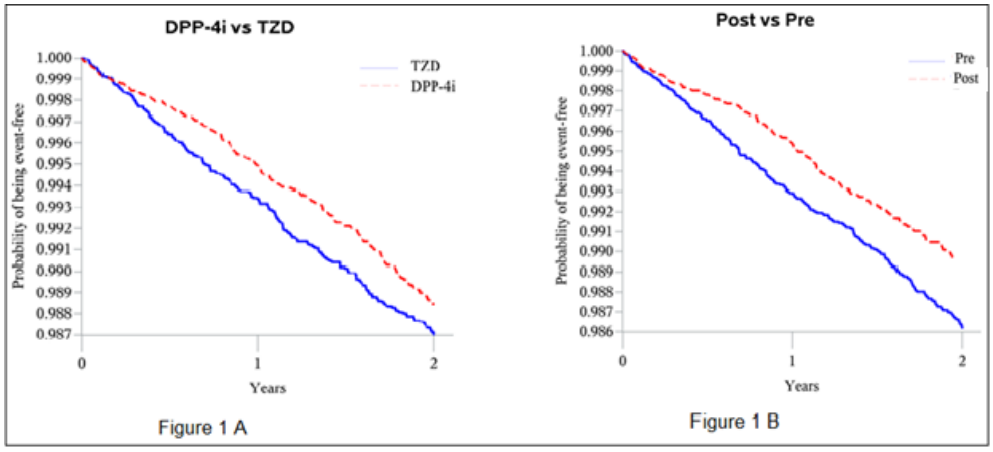


FIGURE 5.3. Comparison of risk differences for heart failure hospitalization. Estimates of RD are based on risks per 100 patients taken from Kaplan–Meier survival curves. The instrumental variable estimator is scaled by a compliance percentage of 40%. Adjusted estimates account for the variables presented in Table 1. IPTW = inverse probability of treatment weight.

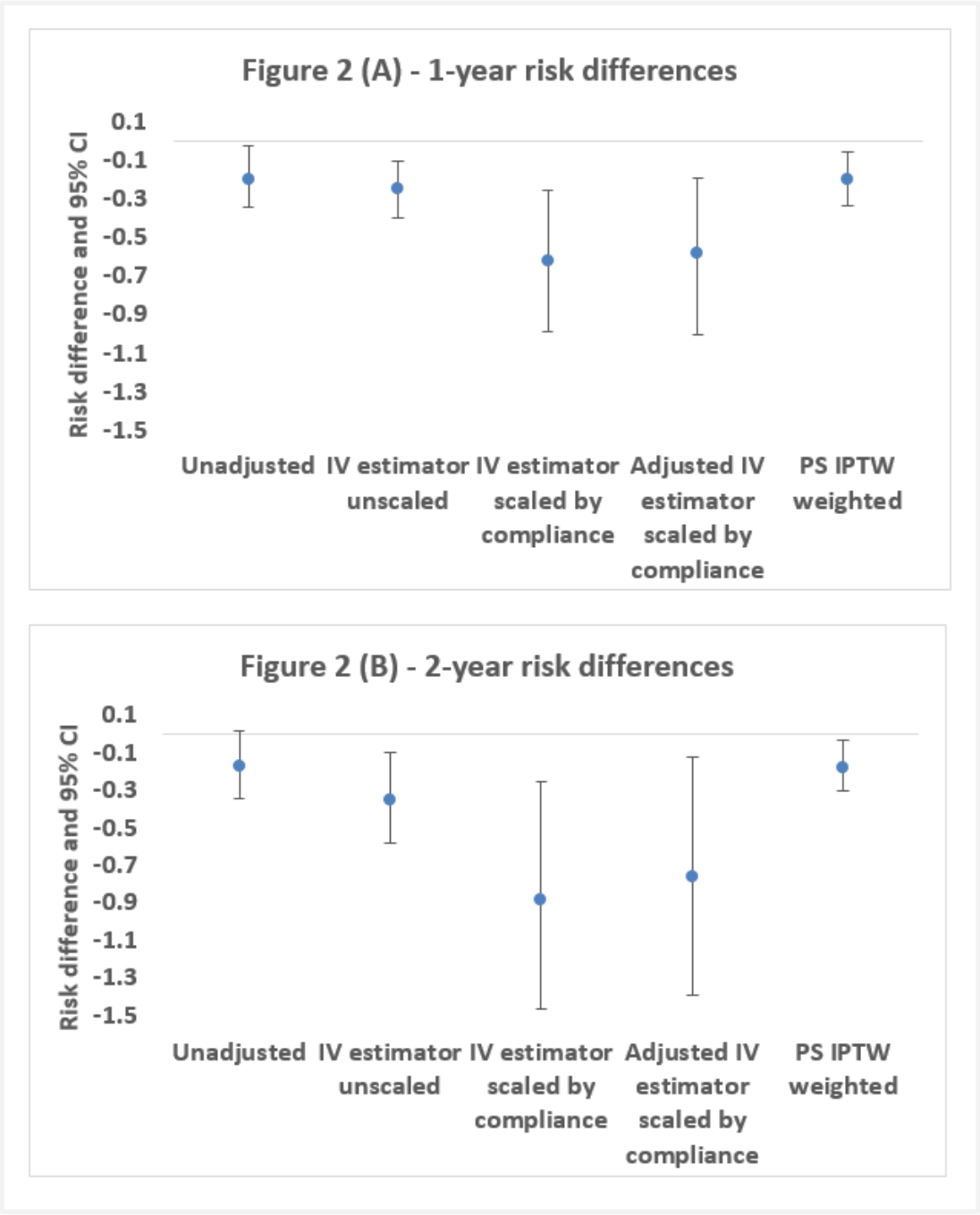


FIGURE 5.4. Sensitivity analyses comparing 1-year periods with maximum separation: December 2011 – December 2012 versus June 2008 – June 2009

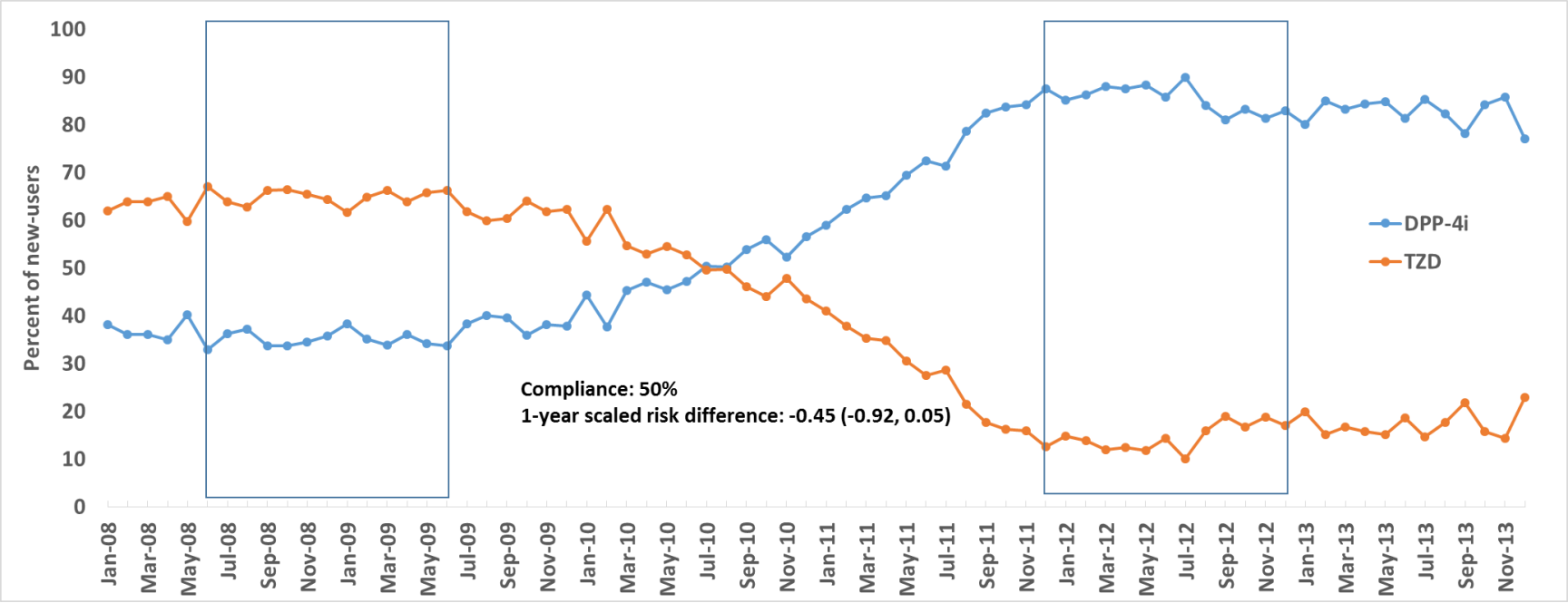
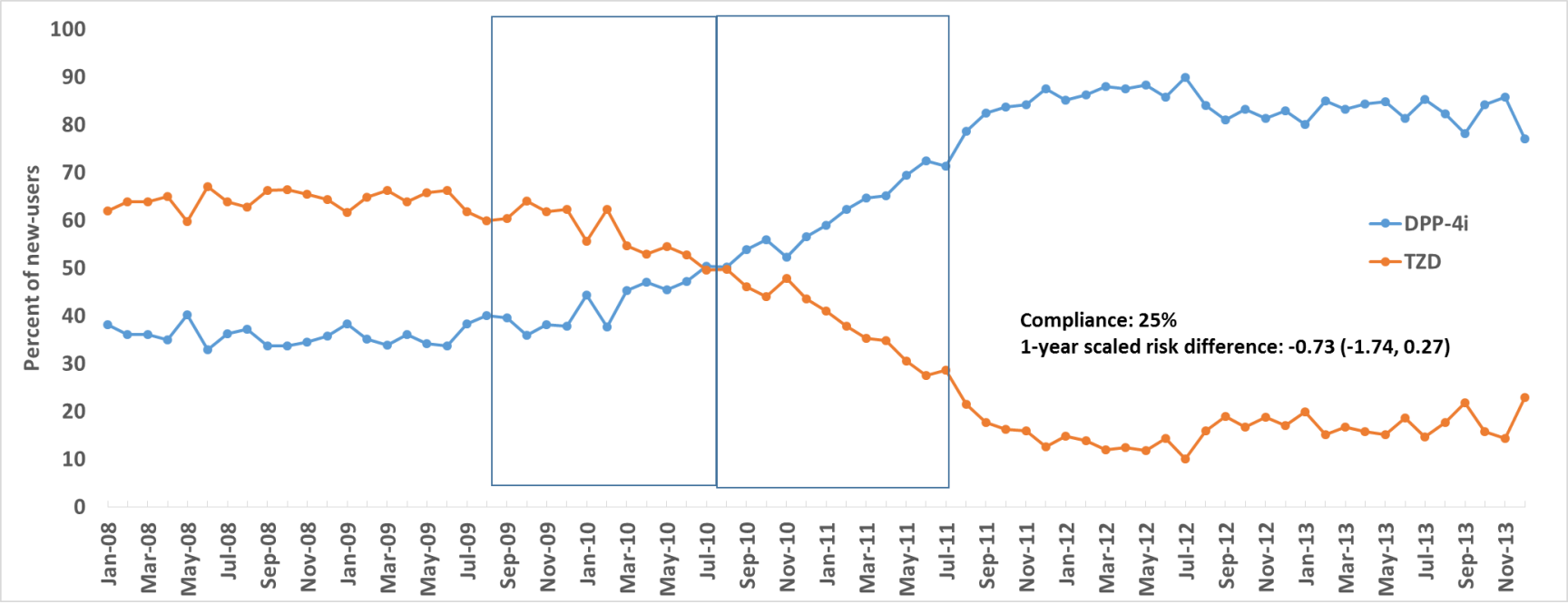


FIGURE 5.5. Sensitivity analyses comparing 1-year periods using cut-point definition of the instrumental variable: August 2010 – August 2011 versus July 2009 – July 2010



CHAPTER 6. CONCLUSIONS AND PUBLIC HEALTH SIGNIFICANCE

I. Summary of specific aims

Diabetes prevalence has been growing worldwide and over a quarter of the U.S. adults 65+ years of age have diabetes.[44] Cardiovascular disease is the leading cause of morbidity and mortality in these patients and the risk increases with age.[45] The use of antihyperglycemic drugs has been increased tremendously in the past few years accounting for approximately 12% of the total cost of diabetes.[62] The antihyperglycemic drug market has typically been very dynamic with a number of new agents being added in the last decade and more drug safety and effectiveness data is becoming available about the existing therapies. Cardiovascular safety of antihyperglycemic drugs is a question of great interest and regulatory agencies require a thorough assessment of cardiovascular risk in the antihyperglycemic drug development programs.[47]

In this study, dipeptidyl peptidase-4 inhibitors (DPP-4i) is the main drug class of interest. DPP-4i were first introduced in the U.S. in 2006 and are currently indicated as second line treatment of diabetes along with other oral drugs sulfonylureas (SU), thiazolidinediones (TZD) and sodium-glucose co-transporter 2 inhibitors (introduced in 2013). DPP-4i have intermediate efficacy, low risk of hypoglycemia and are weight neutral which is a desirable feature of antihyperglycemic drugs. In this study we focused on the cardiovascular effects of DPP-4i relative to therapeutic alternatives in older adults and examined some substantive and methodological aims as described below.

II. Summary of Aim 1

Randomized placebo-controlled clinical trials have examined the CV effects of dipeptidyl peptidase-4 inhibitors (DPP-4i), but there are limited data on the comparative incidence relative to therapeutic alternatives, including sulfonylureas (SU) and thiazolidinediones (TZD).[2-4] The existing observational studies report no increased relative risk of myocardial infarction (MI) and stroke with DPP-4i, but the evidence on heart failure (HF) is mixed.[7-12,25] All the studies mainly reported summary relative risk measures but not the absolute risk measures which may be important to put the issue in context. Further, several studies used a combined pool of non-DPP-4i drugs as the comparator making interpretation of the results difficult and less useful for physicians for making treatment decisions. Ours is the first epidemiologic study comparing the incidence of CV events with DPP-4i versus clinically relevant active comparators in a US population of older adults with a high prevalence of comorbidity and long duration of diabetes (both of which could affect the effects of DPP-4i on CV risk).

We found that there was no increased risk of MI, stroke and HF hospitalizations with DPP-4i relative to SU or TZD in this population of older adults. Since TZDs are associated with an increased risk of HF, we used HF as a positive control outcome in this study expecting no increased risk with DPP-4i and that is what we observed. Our results were robust to a number of sensitivity and subgroup analyses. The risk of the composite outcome (MI, stroke and all-cause mortality) seemed lower with DPP-4i then SU, but this was mainly due to death rather than MI or stroke for which the risk differences were small throughout the study.

Our study had several strengths including use of active-comparator new user cohort design which avoids biases related to prevalent user designs, synchronizes follow-up and minimizes imbalance of baseline covariates as the drugs being compared are for the same

indication.[27, 67] We used well defined comparators (SU and TZD in separate analyses) unlike few other observational studies that compared DPP-4i to a pooled group of comparators thus making interpretation of results difficult. Use of propensity score weighting further balanced the baseline covariates thereby minimizing confounding by these variables. Limitations of our study included a short treatment duration (which is mostly due to real-world drug use patterns), low sensitivity of the ICD-9 codes used to define HF, inability to capture cardiovascular death as cause specific death information is not present in the Medicare claims (but we used all-cause mortality as a proxy since this is responsible for more than half of the deaths in diabetic patients). Finally we could not adjust for smoking and BMI which can affect CV risk. However, we previously found that smoking and BMI do not meaningfully affect the choice of initiation of DPP-4i versus comparators thereby reducing concerns about unmeasured confounding by these variables.[60]

In summary, we did not observe an increased short-term risk of MI, stroke, mortality or HF hospitalization with DPP-4i versus relevant oral second line diabetes drugs in a population of older adults. Along with the RCT results, our results based on real world drug use and effects are relevant to physicians for making antihyperglycemic treatment choices.

III. Summary of Aim 2

A recent study [13] examining the market trends of antihyperglycemic drugs reported that newly approved drugs, particularly the DPP-4i class quickly gained significant market share and was the most commonly prescribed new drug class by 2012. On the other hand, the use of TZD decreased by 2012 potentially due to the safety concerns associated with this class in the last decade. DPP-4i and TZD are both second-line antihyperglycemic drugs, with intermediate (DPP-4i) to high (TZD) efficacy, low risk of hypoglycemia [14] and comparable costs before the generic form of pioglitazone was introduced in August 2012.[15] The DPP-4i and TZD drugs are therefore

expected to be in therapeutic equipoise in a population excluding anyone with clear contraindications as demonstrated in aim 1. The increase in the use of DPP-4i with a simultaneous decrease in TZD use over a short period of time also provides an opportunity to use calendar time as an instrumental variable (IV) as suggested previously in other settings.[16] Using hospitalization for heart failure (HF) as a positive control outcome we explored the use of calendar time as an IV and compared this approach to an active comparator new-user study comparing DPP-4i versus TZD (actual treatment received) in sample of Medicare beneficiaries without a history of HF or related conditions.

We found that calendar time greatly affected the initiation of DPP-4i and TZD and the two curves crossed in 2010 around the time when the Food and Drug Administration (FDA) issued drug safety communications about TZDs. In our study the IV compared the periods from October 2010 through December 2013 ('post') with the period from January 2008 to May 2010 ('pre') leading to IV strength of 40%. Covariate balance across the levels of the IV was slightly better than the covariate balance achieved across the treatments initiated. The IV estimates of the risk difference were similar in direction to the RDs obtained using the propensity score weighting methods, both indicating lesser risk of HF hospitalizations with DPP-4i as expected. The magnitude of the RDs obtained using the two methods were somewhat different which could be because the two analytic approaches are based on different populations and assumptions. Results were consistent across a number of sensitivity analyses.

Strengths of our study included ability to identify a strong instrument of calendar time driven by the dynamics of the antihyperglycemic drug market over a short period of time, thus creating a setting for a 'natural experiment'. Use of an active comparator created a setting of treatment equipoise between DPP-4i and TZD as described before. One of the limitations of our

study included potential violation of one of the IV assumptions which requires that time is not related to the outcome through pathways other than the treatment. However there have been reports about a decreasing trend in HF hospitalizations which could violate this assumption.[73] We conducted a sensitivity analyses to minimize the effect of time trends, but the results were virtually similar. Secondly while the assumption is monotonicity ('no defiers') is expected to hold in our study, it is possible that it is violated in the later periods when the generic form of pioglitazone was introduced in 2012 when patients could initiate TZD in the 'post' period.

While IV methods are generally indicated in settings where substantial unmeasured confounding is expected, according to our knowledge ours is a unique setting where both IV and 'treatment received approaches' may be equally valid. These two approaches address different biases and are based on different assumptions some of which cannot be directly quantified; but the use of a positive control outcome and observing consistent results with either method implies that both methods may supplement each other at the very least. In summary the use of calendar time as an IV in settings where real-world market dynamics lead to profound changes in preferred treatments is worth consideration.

REFERENCES

1. Petrie JR. The cardiovascular safety of incretin-based therapies: a review of the evidence. *Cardiovasc Diabetol*. 2013;12:130-2840-12-130. doi: 10.1186/1475-2840-12-130; 10.1186/1475-2840-12-130.

2. Scirica BM, Bhatt DL, Braunwald E, et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med*. 2013;369(14):1317-1326. doi: 10.1056/NEJMoa1307684; 10.1056/NEJMoa1307684.

3. White WB, Cannon CP, Heller SR, et al. Alogliptin after acute coronary syndrome in patients with type 2 diabetes. *N Engl J Med*. 2013;369(14):1327-1335. doi: 10.1056/NEJMoa1305889; 10.1056/NEJMoa1305889.

4. Green JB, Bethel MA, Armstrong PW, et al. Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2015;373(3):232-242.

5. Standl E. Saxagliptin, alogliptin, and cardiovascular outcomes. *N Engl J Med*. 2014;370(5):483. doi: 10.1056/NEJMc1313880#SA1; 10.1056/NEJMc1313880#SA1.

6. White WB, Zannad F. Saxagliptin, alogliptin, and cardiovascular outcomes. *N Engl J Med*. 2014;370(5):484.

7. Weir DL, McAlister FA, Senthilselvan A, Minhas-Sandhu JK, Eurich DT. Sitagliptin use in patients with diabetes and heart failure: a population-based retrospective cohort study. *JACC: Heart Failure*. 2014;2(6):573-582.

8. Yu OH, Filion KB, Azoulay L, Patenaude V, Majdan A, Suissa S. Incretin-based drugs and the risk of congestive heart failure. *Diabetes Care*. 2015;38(2):277-284. doi: 10.2337/dc14-1459 [doi].

9. Kim SC, Glynn RJ, Liu J, Everett BM, Goldfine AB. Dipeptidyl peptidase-4 inhibitors do not increase the risk of cardiovascular events in type 2 diabetes: a cohort study. *Acta Diabetol*. 2014;51(6):1015-1023.

10. Fu AZ, Johnston SS, Ghannam A, et al. Association Between Hospitalization for Heart Failure and Dipeptidyl Peptidase-4 Inhibitors in Patients With Type 2 Diabetes: An Observational Study. *Diabetes Care*. 2016. doi: dc150764 [pii].

11. Shih C, Chen H, Kuo S, Ou S, Chen Y. Cardiovascular outcomes of dipeptidyl peptidase-4 inhibitors in elderly patients with type 2 diabetes: A nationwide study. *Journal of the American Medical Directors Association*. 2016;17(1):59-64.

12. Velez M, Peterson EL, Wells K, et al. Association of antidiabetic medications targeting the glucagon-like peptide 1 pathway and heart failure events in patients with diabetes. *J Card Fail*. 2015;21(1):2-8.

13. Hampp C, Borders-Hemphill V, Moeny DG, Wysowski DK. Use of antidiabetic drugs in the U.S., 2003-2012. *Diabetes Care*. 2014;37(5):1367-1374. doi: 10.2337/dc13-2289 [doi].

14. Inzucchi SE, Bergenstal RM, Buse JB, et al. Management of hyperglycemia in type 2 diabetes, 2015: a patient-centered approach: update to a position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2015;38(1):140-149. doi: 10.2337/dc14-2441 [doi].

15. U.S. Food and Drug Administration. FDA approves first generic Actos to treat type 2 diabetes. http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm315951.htm. Accessed March 12, 2016.

16. De Filippo Mack C, Brookhart MA, Glynn RJ, et al. Comparative Effectiveness of Oxaliplatin Versus 5-flourouricil in Older Adults. *Epidemiology*. 2015;26(5):690-699.

17. Drucker DJ, Sherman SI, Gorelick FS, Bergenstal RM, Sherwin RS, Buse JB. Incretin-based therapies for the treatment of type 2 diabetes: evaluation of the risks and benefits. *Diabetes Care*. 2010;33(2):428-433. doi: 10.2337/dc09-1499; 10.2337/dc09-1499.

18. Ussher JR, Drucker DJ. Cardiovascular biology of the incretin system. *Endocr Rev*. 2012;33(2):187-215. doi: 10.1210/er.2011-1052; 10.1210/er.2011-1052.

19. Kannel WB, Kannel C, Paffenbarger RS,Jr, Cupples LA. Heart rate and cardiovascular mortality: the Framingham Study. *Am Heart J*. 1987;113(6):1489-1494.

20. Monami M, Ahren B, Dicembrini I, Mannucci E. Dipeptidyl peptidase-4 inhibitors and cardiovascular risk: a meta-analysis of randomized clinical trials. *Diabetes Obes Metab*. 2013;15(2):112-120. doi: 10.1111/dom.12000; 10.1111/dom.12000.

21. Patil HR, Al Badarin FJ, Al Shami HA, et al. Meta-analysis of effect of dipeptidyl peptidase-4 inhibitors on cardiovascular risk in type 2 diabetes mellitus. *Am J Cardiol*. 2012;110(6):826-833. doi: 10.1016/j.amjcard.2012.04.061; 10.1016/j.amjcard.2012.04.061.

22. FDA Drug Safety Communication: FDA to review heart failure risk with diabetes drug saxagliptin (marketed as Onglyza and Kombiglyze XR). http://www.fda.gov/Drugs/DrugSafety/ucm385287.htm. Accessed April,2014, .

23. McMurray JJ, Gerstein HC, Holman RR, Pfeffer MA. Heart failure: a cardiovascular outcome in diabetes that can no longer be ignored. *The Lancet Diabetes & Endocrinology*. 2014.

24. Best JH, Hoogwerf BJ, Herman WH, et al. Risk of cardiovascular disease events in patients with type 2 diabetes prescribed the glucagon-like peptide 1 (GLP-1) receptor agonist exenatide twice daily or other glucose-lowering therapies: a retrospective analysis of the LifeLink database. *Diabetes Care*. 2011;34(1):90-95. doi: 10.2337/dc10-1393; 10.2337/dc10-1393.

25. Wang K, Liu C, Chao T, et al. Sitagliptin and the risk of hospitalization for heart failure: a population-based study. *Int J Cardiol*. 2014;177(1):86-90.

26. Marx N, Rosenstock J, Kahn SE, et al. Design and baseline characteristics of the CARdiovascular Outcome Trial of LINAgliptin Versus Glimepiride in Type 2 Diabetes (CAROLINA(R)). *Diab Vasc Dis Res*. 2015;12(3):164-174. doi: 10.1177/1479164115570301 [doi].

27. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915-920.

28. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf*. 2010;19(6):537-554.

29. Morrish N, Wang S, Stevens L, Fuller J, Keen H, WHO Multinational Study Group. Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. *Diabetologia*. 2001;44(2):S14-S21.

30. Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. *Am Heart J*. 2004;148(1):99-104.

31. Tirschwell DL, Longstreth WT,Jr. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470.

32. Kucharska-Newton AM, Heiss G, Ni H, et al. Identification of Heart Failure Events in Medicare Claims: The Atherosclerosis Risk in Communities (ARIC) Study. *J Card Fail*. 2016;22(1):48-55.

33. Nesto RW, Bell D, Bonow RO, et al. Thiazolidinedione use, fluid retention, and congestive heart failure: a consensus statement from the American Heart Association and American Diabetes Association. October 7, 2003. *Circulation*. 2003;108(23):2941-2948. doi: 10.1161/01.CIR.0000103683.99399.7E [doi].

34. Hernandez AV, Usmani A, Rajamanickam A, Moheet A. Thiazolidinediones and risk of heart failure in patients with or at high risk of type 2 diabetes mellitus. *American Journal of Cardiovascular Drugs*. 2011;11(2):115-128.

35. Green JB, Bethel MA, Armstrong PW, et al. Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2015.

36. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.

37. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology*. 2003;14(6):680-686. doi: 10.1097/01.EDE.0000081989.82616.7d.

38. Kurth T, Walker AM, Glynn RJ, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol*. 2006;163(3):262-270. doi: kwj047 [pii].

39. Cole SR, Lau B, Eron JJ, et al. Estimation of the standardized risk difference and ratio in a competing risks framework: application to injection drug use and progression to AIDS after initiation of antiretroviral therapy. *Am J Epidemiol*. 2015;181(4):238-245. doi: 10.1093/aje/kwu122 [doi].

40. Myers JA, Rassen JA, Gagne JJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. *Am J Epidemiol*. 2011;174(11):1213-1222. doi: 10.1093/aje/kwr364 [doi].

41. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163(12):1149-1156. doi: kwj149 [pii].

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

43. Ranganathan P, Pramesh C. Censoring in survival analysis: potential for bias. *Perspect Clin Res*. 2012;3(1):40.

44. Kirkman MS, Briscoe VJ, Clark N, et al. Diabetes in older adults. *Diabetes Care*. 2012;35(12):2650-2664. doi: 10.2337/dc12-1801 [doi].

45. Halter JB, Musi N, McFarland Horne F, et al. Diabetes and cardiovascular disease in older adults: current status and future directions. *Diabetes*. 2014;63(8):2578-2589. doi: 10.2337/db14-0020 [doi].

46. European Medicines Agency. European Medicines Agency Recommends Suspension of Avandia, Avandamet and Avaglim, 2010. http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2010/09/news_detail_001119.jsp&mid=WC0b01ac058004d5c1. Accessed March 12, 2016.

47. FDA 2008. Guidance for Industry Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. http://www.fda.gov/downloads/Drugs/%20GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf. Accessed February 3, 2016.

48. White WB, Baker WL. Cardiovascular Effects of Incretin-Based Therapies. *Annu Rev Med*. 2016;67:245-260.

49. Medicare Program - General Information. http://www.cms.gov/Medicare/Medicare-General-Information/MedicareGenInfo/index.html. Accessed April, 2014, .

50. Brief Summaries of Medicare & Medicaid. 2011. http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MedicareProgramRatesStats/downloads/MedicareMedicaidSummaries2011.pdf. April, 2014.

51. Gokhale M, Buse JB, Gray CL, Pate V, Marquis MA, Stürmer T. Dipeptidyl-peptidase-4 inhibitors and pancreatic cancer: a cohort study. *Diabetes, Obesity and Metabolism*. 2014.

52. Brookhart MA. Counterpoint: the treatment decision design. *Am J Epidemiol*. 2015;182(10):840-845. doi: 10.1093/aje/kwv214 [doi].

53. Graham DJ, Ouellet-Hellstrom R, MaCurdy TE, et al. Risk of acute myocardial infarction, stroke, heart failure, and death in elderly Medicare patients treated with rosiglitazone or pioglitazone. *JAMA*. 2010;304(4):411-418.

54. Information for Healthcare Professionals: Pioglitazone HCl (marketed as Actos, Actoplus Met, and Duetact). http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm124178.htm. Accessed February 3, 2016.

55. Simpson SH, Lee J, Choi S, Vandermeer B, Abdelmoneim AS, Featherstone TR. Mortality risk among sulfonylureas: a systematic review and network meta-analysis. *The Lancet Diabetes & Endocrinology*. 2015;3(1):43-51.

56. Forst T, Hanefeld M, Jacob S, et al. Association of sulphonylurea treatment with all-cause and cardiovascular mortality: a systematic review and meta-analysis of observational studies. *Diab Vasc Dis Res*. 2013;10(4):302-314. doi: 10.1177/1479164112465442 [doi].

57. Chen D, Wang S, Mao C, et al. Sitagliptin and cardiovascular outcomes in diabetic patients with chronic kidney disease and acute myocardial infarction: a nationwide cohort study. *Int J Cardiol*. 2015;181:200-206.

58. Nikolaidis LA, Mankad S, Sokos GG, et al. Effects of glucagon-like peptide-1 in patients with acute myocardial infarction and left ventricular dysfunction after successful reperfusion. *Circulation*. 2004;109(8):962-965. doi: 10.1161/01.CIR.0000120505.91348.58 [doi].

59. Sokos GG, Nikolaidis LA, Mankad S, Elahi D, Shannon RP. Glucagon-like peptide-1 infusion improves left ventricular ejection fraction and functional status in patients with chronic heart failure. *J Card Fail*. 2006;12(9):694-699.

60. Gokhale M, Buse J, Sturmer T. Effect of Body Mass Index on Choice of Initiating Diabetes Therapies. *Pharmacoepidemiology and Drug Safety. Special Issue: Abstracts of the 30th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, October 24–27, 2014, Taipei, Taiwan*. ;23.

61. National Diabetes Statistics Report, 2014. http://www.cdc.gov/diabetes/pdfs/data/2014-report-estimates-of-diabetes-and-its-burden-in-the-united-states.pdf. Accessed March 12, 2016.

62. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. *Diabetes Care*. 2013;36(4):1033-1046. doi: 10.2337/dc12-2625 [doi].

63. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med*. 2007;356(24):2457-2471. doi: 10.1056/NEJMoa072761.

64. U.S. Food and Drug Administration. FDA significantly restricts access to the diabetes drug Avandia. http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm226956.htm. Accessed March 12, 2016.

65. U.S. Food and Drug Administration. FDA Drug Safety Communication: Ongoing Safety Review of Actos (pioglitazone) and Potential Increased Risk of Bladder Cancer After Two Years Exposure. http://www.fda.gov/Drugs/DrugSafety/ucm226214.htm. Accessed March 12, 2016.

66. U.S. Food and Drug Administration. FDA Drug Safety Communication: Updated drug labels for pioglitazone-containing medicines. http://www.fda.gov/Drugs/DrugSafety/ucm266555.htm. Accessed March 12, 2016.

67. Lund JL, Richardson DB, Stürmer T. The Active Comparator, New User Study Design in Pharmacoepidemiology: Historical Foundations and Contemporary Application. *Current Epidemiology Reports*. 2015;2(4):221-228.

68. Gokhale M, Buse JB, Jonsson Funk M, Lund JL, Simpson RS, Stürmer T. *Comparative incidence of cardiovascular events in older adults initiating Dipeptidyl peptidase 4 inhibitors versus Therapeutic alternatives.* [Doctorate]. University of North Carolina.

69. Swanson SA, Hernan MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology*. 2013;24(3):370-374. doi: 10.1097/EDE.0b013e31828d0590 [doi].

70. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006;17(3):268-275. doi: 10.1097/01.ede.0000193606.58671.c5 [doi].

71. Erdmann E, Charbonnel B, Wilcox RG, et al. Pioglitazone use and heart failure in patients with type 2 diabetes and preexisting cardiovascular disease: data from the PROactive study (PROactive 08). *Diabetes Care*. 2007;30(11):2773-2778. doi: dc07-0717 [pii].

72. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000;29(4):722-729.

73. Krumholz HM, Normand SL, Wang Y. Trends in hospitalizations and outcomes for acute cardiovascular disease and stroke, 1999-2011. *Circulation*. 2014;130(12):966-975. doi: 10.1161/CIRCULATIONAHA.113.007787 [doi].

74. Centers for Medicare and Medicaid Services. Readmissions Reduction Program (HRRP). https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html. Accessed March 12, 2016.

75. Jackson JW, Swanson SA. Toward a clearer portrayal of confounding bias in instrumental variable applications. *Epidemiology*. 2015;26(4):498-504. doi: 10.1097/EDE.0000000000000287 [doi].