Published in final edited form as:

Diabetes Obes Metab. 2018 April ; 20(4): 974–984. doi:10.1111/dom.13184.

Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical parameters only observed in electronic health records

E Patorno¹, C Gopalakrishnan¹, JM Franklin¹, KG Brodovicz², E Masso-Gonzalez³, DB Bartels^{3,4}, J Liu¹, and S Schneeweiss¹

¹Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

²Global Epidemiology, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA

³Corporate Department Global Epidemiology, Boehringer Ingelheim GmbH, Ingelheim, Germany

⁴Hannover Medical School, Inst. for Epidemiology, Social Medicine and Health Systems Research, Hannover, Germany

Abstract

Background: Healthcare claims databases can provide information on the effects of type 2 diabetes (T2DM) medications as used in routine care, but often do not contain data on important clinical characteristics, which may be captured in electronic health records (EHR).

Objectives: To evaluate the extent to which balance in unmeasured patient characteristics was achieved in claims data, by comparing against more detailed information from linked EHR data.

Methods: Within a large US commercial insurance database and using a cohort design, we identified T2DM patients initiating linagliptin or a comparator agent within class (i.e., other DPP-4 inhibitors) or outside class (i.e., (pioglitazone or sulfonylureas) between 05/2011-12/2012. We focused on comparators used at a similar stage of diabetes as linagliptin. For each comparison, 1:1 propensity score (PS) matching was used to balance over 100 baseline claims-based characteristics, including proxies of diabetes severity and duration. Additional clinical data from EHRs was available for a subset of patients. We assessed representativeness of the claims-EHR linked subset, evaluated the balance of claims- and EHR-based covariates before and after PS-matching via standardized differences (SD), and quantified the potential bias associated with observed imbalances.

Results: From a claims-based study population of 166,613 T2DM patients, 7,219 (4.3%) patients were linked to their EHR data. Claims-based characteristics between the EHR-linked and EHR-unlinked patients were comparable ($SD < 0.1$), confirming representativeness of the EHR-linked subset. The balance of claims-based and EHR-based patient characteristics appeared to be reasonable before PS-matching and generally improved in the PS-matched population, to be

SD<0.1 for most patient characteristics and SD<0.2 for select laboratory results and BMI categories, not large enough to cause meaningful confounding.

Conclusion: In the context of pharmacoepidemiologic research on diabetes therapy, choosing appropriate comparison groups paired with a new user design and 1:1 PS matching on many proxies of diabetes severity and duration improves balance in covariates typically unmeasured in administrative claims datasets, to an extent that residual confounding is unlikely.

Keywords

type 2 diabetes; glucose-lowering medications; administrative data; linkage; electronic medical records

Background

Over 29 million people in the United States (U.S.) currently have diabetes, with approximately 1.7 million new cases every year.¹ Diabetes is associated with several serious complications and is a leading cause of death, with 200,000 related deaths annually.^{1,2}

Most diabetes drugs enter the market based on relatively small, short, placebo-controlled randomized controlled trials (RCT) that often use surrogate outcomes as endpoints, and in which many patients who would receive the drug(s) under routine care are generally underrepresented. In 2008, the U.S. Food and Drug Administration (FDA) mandated larger post-marketing cardiovascular outcome trials (CVOT) for all new antidiabetic drugs to rule out excess cardiovascular (CV) risk.^{3–9} Though CVOTs provide additional information on the effects of new diabetes therapies, clinicians caring for patients with diabetes continue to face challenges in choosing the best glucose-lowering treatment for their patients because: (1) key information on long term safety and effectiveness is often unavailable for a considerable time after approval; (2) the safety of new drugs is poorly established in populations underrepresented in RCTs; and (3) RCTs do not usually perform head-to-head comparisons across clinically relevant glucose-lowering options.

Large pharmacoepidemiologic studies based on longitudinal insurance claims data, routinely generated in the provision of healthcare for millions of patients, are increasingly utilized to fill these gaps and provide information on the comparative effectiveness and safety of glucose-lowering agents in real-world populations.^{10–13} However, these studies are often criticized because of the lack of information on critical clinical characteristics, such as body mass index (BMI), exact duration of diabetes, and hemoglobin A1C (HbA1C) levels, which in the absence of randomization could remain unbalanced across comparison groups, even after adjustment, and thus lead to biased treatment effect estimates.

It is often assumed that the absence of this information in pharmacoepidemiologic studies could be largely addressed by the application of state-of-the-art study design and analytic choices. For example, in the context of safety and effectiveness research on diabetes therapy, using a new-user study design enhanced by the proper choice of an active comparator drug, which tends to be used by patients at a similar stage of diabetes, could better distinguish drug effects from diabetes disease effects.^{14,15} Additionally, adjusting analyses via

propensity score (PS) matching could leverage the vast information recorded in large claims databases to estimate treatment effects in a population with clinical equipoise regarding many aspects of care, including characteristics that may act as proxies for unmeasured information.^{16,17} These strategies combined are thought to improve balance in patient characteristics, and thus mitigate the potential confounding in non-interventional pharmacoepidemiologic studies.

With the growing proliferation of digital information in healthcare, subsets of administrative data can be successfully linked to electronic health records (EHR), which routinely collect important clinical information, to assess the balance of these characteristics across exposure groups identified in claims. Thus, we sought to evaluate the extent to which balance in unmeasured patient characteristics was achieved in claims data, by comparing against more detailed information from EHR data.

Methods

Overview of monitoring program

This study was preemptively conducted in the context of an ongoing multi-year monitoring program aiming at assessing linagliptin, a DPP-4 inhibitor, compared with 3 groups of oral antidiabetic drugs, with regard to several safety and effectiveness outcomes (NCT02197078, EUPAS5790). The primary outcome of the monitoring program is a composite cardiovascular disease (CVD) outcome comprised of acute coronary syndrome, ischemic or hemorrhagic stroke, and coronary revascularization. Secondary outcomes include, but are not limited to, the individual components of the composite CVD outcome, hospitalization for heart failure, renal outcomes, and malignancy. The monitoring program involves repeated outcome evaluations over time, starting on May 1, 2011 (consistent with the availability of linagliptin in the U.S.).

Data source

Data were collected from the Truven Health MarketScan claims database,¹⁸ a nationwide U.S. insurance dataset covering commercially insured persons and patients enrolled in Medicare Advantage plans, employer sponsored coverage of seniors, and Medicare supplemental insurance. For each participant, the database contains demographic information, health plan enrollment status, and longitudinal records of reimbursed medical services, including inpatient and outpatient medical encounters coded using the International Classification of Diseases, Ninth Revision, Clinical Modification codes (ICD-9-CM) and the Current Procedural Terminology (CPT-4) classifications, and filled medications, including the National Drug Code (NDC) numbers, quantity dispensed, and days' supply. For a subset of the population, claims data were linked to EHRs from select clinics providing care to MarketScan beneficiaries.

The Institutional Review Board of the Brigham and Women's Hospital approved the study.

Formation of claims-based study population

The study population included three pairwise cohorts of patients aged 18 years or older who initiated linagliptin or a comparator, i.e., other DPP-4 inhibitors (alogliptin, saxagliptin, or sitagliptin), 2nd generation sulfonylureas (glimepiride, glipizide, or glyburide), or pioglitazone, between May 1, 2011 and December 31, 2012. These comparators were chosen as they represent therapeutic strategies used at comparable stages of diabetes progression as linagliptin, which is not a first-line recommended strategy for diabetes, to enhance clinical equipoise across treatment groups and reduce confounding. Patients entered the study on the day of a first filled prescription of any of the drugs above (i.e., cohort entry date), defined for each pair-wise cohort as no prior use of linagliptin or the specific comparator class or agent in the previous six months, and were required to have six or more months of continuous enrollment prior to the cohort entry date. Patients who met inclusion criteria could contribute to multiple cohorts. We restricted to patients with a diagnosis of type 2 diabetes mellitus (T2DM), defined as an inpatient or outpatient ICD-9CM diagnosis code of 250.x0 or 250.x2 at any point prior to drug initiation. We excluded patients with a diagnosis of T1DM (ICD-9CM diagnosis code 250.x1 or 250.x3) at any point prior to cohort entry, or history of secondary or gestational diabetes, malignancy, end-stage renal disease, human immunodeficiency virus, organ transplant, or a nursing home admission in the previous six months.

To control for imbalances in patient characteristics between treatment groups, in three separate multivariable logistic regression models we estimated exposure propensity scores (PS) as the predicted probability of receiving the treatment of interest (i.e. linagliptin vs. each comparator) conditional upon over 100 claims-based subjects' baseline characteristics,¹⁹ identified during the six months before and including the cohort entry date. Emphasis was placed on the identification of claims-measured proxies of diabetes severity and duration (e.g., number of glucose-lowering medications at index date and specific past or concurrent diabetes therapy, diabetic nephropathy, neuropathy, retinopathy, diabetic foot, number of HbA1c or glucose tests ordered, etc.). Other patient characteristics included demographics, presence of other comorbidities, use of medications, and indicators of health care utilization as proxy for overall disease state and care intensity. (See Table-1 and eTable-1)

Comorbidities were defined using ICD-9 codes and CPT-4 codes. Exposure groups were 1:1 matched on their PS using nearest neighbor matching without replacement with a maximum caliper of 0.05.²⁰ Matching was performed within calendar quarters to account for changing prescribing behaviors over time, as selective prescribing of a new medication may be strong in the early marketing period and characteristics of patients receiving the new agent may shift quickly.²¹

Claims-EHR linkage and EHR-based clinical characteristics

For a subset of patients enrolled in the claims-based study population before and after PS-matching, insurance claims were enriched with additional data obtained through linkage with EHRs, performed by Truven Health Analytics®. EHR information was contributed by select clinics providing care to MarketScan beneficiaries. In order to maintain patient confidentiality, the linkage was performed through variables such as the patient's gender, month, year of birth and three-digit ZIP code of residence. In large ZIP code areas,

additional criteria such as dates of office visits were used to discriminate true from false matches.²² Within the linked subset, several EHR-based clinical characteristics related to T2DM treatment that may also predict the primary CVD outcome of the monitoring program were identified and used in the analyses. EHR-based covariates were captured prior to cohort entry and included health behaviors (smoking status and BMI), duration of diabetes (the earliest record for a T2DM diagnosis in the EHR using all available information prior to treatment initiation), laboratory test results (baseline HbA1c, creatinine, estimated glomerular filtration rate [eGFR],²³ and lipid levels), and blood pressure (systolic and diastolic levels). If multiple recording of EHR-based covariates were available, we only considered the value closest to the day of cohort entry.

Statistical analysis

To assess whether the claims-EHR linked subset was representative of the overall study population, we compared claims-based characteristics among study participants for whom EHR data were available to patients without EHR data available, and evaluated covariate balance using standardized differences. Similarly, to assess the presence of potential confounding associated with unmeasured clinical characteristics in the claims-based study population, we cross-tabulated baseline patient characteristics by each pair of linagliptin or its comparator and evaluated the balance of EHR-based covariates between exposure groups, before and after PS-matching, via standardized differences. Meaningful imbalances in standardized differences were defined as differences greater than 0.1.²⁴

In a secondary analysis, we quantified the potential bias associated with observed imbalances, through hypothetical scenarios built upon varying assumptions of exposure-outcome and confounder-outcome associations.²⁵

Results

After applying inclusion and exclusion criteria, we identified a claims-based study population of 166,613 T2DM patients who initiated either linagliptin or a comparator: 70,427 patients who initiated either linagliptin (N=5,732) or another DPP-4 inhibitor (N=64,695), 77,382 patients who initiated either linagliptin (N=3,436) or a sulfonylurea (N=73,946), and 18,804 patients who initiated either linagliptin (N=4,441) or pioglitazone (N=14,363) (Figure-1). The study included 5,697 unique linagliptin initiators. After 1:1 PS-matching, we identified 5,689 pairs of patients initiating linagliptin or another DPP-4 inhibitor, 3,410 pairs of patients initiating linagliptin or a sulfonylurea, and 3,963 pairs of patients initiating linagliptin or pioglitazone. From the total claims-based study population (N =166,613 patients), 7,219 (4.3%) T2DM initiators were successfully linked to EHRs (598 initiators of linagliptin, 3,041 other DPP-4 inhibitors, 3,050 sulfonylureas, and 530 pioglitazone). From the overall EHR-linked T2DM initiators, 1,159 were PS-matched patients (Figure-1).

Claims-based patient characteristics between study participants for whom EHR data were available and patients without available EHR data were well balanced (standardized difference <0.1), suggesting the EHR-linked subset was representative of the overall study population (eTable-2). Independently of EHR data availability, drug initiators were

approximately 55-year-old, were more likely to be male, to be on therapy with one other antidiabetic medication, and to have low prevalence of diabetes complications and other comorbidities. Minor imbalances were noted for a few characteristics, i.e., total number of distinct medications prescribed, number of physician visits, and number of lab tests ordered, all of which were more prevalent amongst the EHR-linked subset.

Within the overall study population, claims-based patient characteristics between exposure groups in each comparison appeared reasonably well balanced even before PS-matching, with initiators of linagliptin and comparator agents having similar mean age and prevalence of diabetes complications (Table-1 and eTable-1). Compared with initiators of other agents, linagliptin initiators had similar mean age (approximately 55 years), gender distribution (approximately 60% males), use of antidiabetic medications (over 50% started therapy in augmentation to one other antidiabetic medication), and prevalence of most diabetes complications, e.g., diabetic nephropathy, neuropathy, retinopathy, and diabetic foot. Yet, we also noted a few imbalances. Specifically, compared with initiators of other agents linagliptin initiators were more likely to have a greater overall burden of comorbidities, as measured by the Romano modification of the Charlson comorbidity score,²⁶ a higher prevalence of select comorbidities, e.g., kidney disease and congestive heart failure, a lower utilization of metformin combination therapy, and a higher utilization of overall medications (Table-1 and eTable-1). After PS-matching, all claims-based characteristics were well balanced between linagliptin initiators and initiators of comparator agents.

As for claims-based characteristics in the overall population, within the EHR-linked subset, EHR-based covariates appeared reasonably well balanced before PS-matching (Table-2 and eTable-3). Smoking status, BMI, diabetes duration, and blood pressure, were largely well balanced across comparisons, with most of the population being never smoker, being obese or severely obese, having diabetes duration under 3 years, and being normotensive. A few imbalances with standardized differences smaller than 0.2 were noted within the comparison of linagliptin vs. pioglitazone for BMI, with higher proportions of obese or severely obese patients among linagliptin initiators, and for laboratory test results such as HbA1c, eGFR, and selected lipid levels, particularly within the comparisons of linagliptin vs. other DPP-4 inhibitors and vs. pioglitazone. After PS-matching, smoking status and diabetes duration were well balanced within each PS-matched exposure group (similar to that observed before matching) and the balance of laboratory test results meaningfully improved across all comparisons, with most standardized differences smaller than 0.1. Sporadic minor imbalances resulting in standardized differences largely smaller than 0.2 remained across comparisons.

The quantification of the potential bias associated with observed imbalances revealed that hypothetical exposure-outcome relative risks are fairly robust even under extreme assumptions of the confounder-outcome association. If the duration of diabetes of a year or longer would truly increase the risk of the outcome by 50% ($RR_{CD}=1.5$) then the observed RR in claims data would truly be 1.52 instead of 1.50 (Figure 2, Panel A left). Similarly, if a 1 percent point increase in HbA1c would result in a 25% increase in outcome risk then the correct RR in claims would be 1.53 instead of 1.50 (Figure 2, Panel A right).

Discussion

Using a claims-based study population of 166,613 T2DM patients linked to EHR data for a subset of 7,219 (4.3%) drug initiators, we observed overall balance of claims-based patient characteristics and EHR-based clinical variables across three pair-wise cohorts of patients who initiated either linagliptin or another antidiabetic agent after PS-matching. Both claims-based and EHR-based covariates were reasonably well-balanced even before PS-matching owing to appropriate study design choices; PS-matching further improved their balance. The comparability between the EHR-linked subset and the non-linked study population and the random nature of the EHR sampling from the claims-based study population (i.e., the availability of EHR data among patients enrolled in the claims-based study population is expected to be at random) indicates that the observed balance of important clinical information should be expected in the overall study population.

Our study findings indicate that proper study design choices can meaningfully reduce the likelihood of residual confounding from typically unmeasured variables such as duration of diabetes, BMI, smoking status, and HbA1C levels in pharmacoepidemiologic studies of diabetes treatment. The choice of clinically meaningful comparators, i.e., initiators of other drugs that are available at the time of market authorization of a newly marketed medication of interest and that are used at a similar stage of diabetes progression, can mitigate bias due to confounding by indication and to differences in health system use.^{14,27} An incident user design, which includes patients initiating a treatment of interest or comparator agents, avoids the problem of studying prevalent users who are ‘survivors’ of early adverse effects, allows evaluating drug effects that vary over time, and ensures that baseline covariates are assessed before treatment initiation and are not affected by treatment itself.^{12,28} These choices resulted in fair balance in EHR-based covariates even before PS-matching, as a testament of the major role played by a proper study design for the achievement of overall study validity.^{14,15} Once adequate study design choices have been made, our findings also suggest that the inclusion of a large number of patient characteristics in the estimation of the propensity score can further lead to balance of unmeasured but correlated variables by proxy, and as a result, mitigate confounding by the same unmeasured covariates, as previously observed.²⁹ In our study, we considered many aspects of care, including several proxies of diabetes progression, in the calculation of the propensity score, and used the estimated score to match study participants with expected clinical equipoise on these characteristics, i.e., excluding patients who will always or never receive therapy because of indications or contraindications.¹⁶

Sporadic imbalances resulting in standardized differences between 0.1 and 0.2 remained with respect to a few specific EHR-based covariates across propensity-score matched comparisons. However, these imbalances are not expected to meaningfully affect confounding in a monitoring program on diabetes treatment because 1) from a clinical perspective these imbalances are considered minor at most; as an example, the numerical differences observed in HbA1c values are within the FDA-recommended non-inferiority margin of 0.3 to 0.4 percentage points³⁰; 2) a secondary bias analysis revealed that the resulting changes are minor even under fairly extreme assumptions and may even cancel each other out depending on the correlation between the observed clinical parameters; 3)

they could be partly driven by the limited sample size of the claims-EHR-linked population paired with a random sampling strategy that selected claims-EHR-linked patients independently of the PS-matching process, as hinted by the presence of less optimal balance among less prevalent covariate categories, and by the worsening of the balance of a few covariates in the PS-matched population.

The primary strength of the current study was the ability to empirically confirm the balance of clinical covariates typically unmeasured in claims-based pharmacoepidemiologic studies, within the same population that is currently being used for a monitoring program comparing second-line treatments among T2DM patients as treated in routine care. Findings can be leveraged to inform targeted bias analyses of any residual imbalances in real-time. Another strength is that EHR information on potential confounders was recorded prior to diabetes therapy initiation, and thus the accuracy and completeness of the data were unlikely to be related to treatment initiation.

We were able to evaluate covariate balance only within an EHR-linked subset corresponding to the 4.3% of the overall study population, which could have undermined the generalizability of our findings to the main population. However, the comparability in claims-based covariate distribution between the EHR-linked subset and the overall study population, and the random nature of the EHR sampling, suggests balance in important clinical information should be expected for the overall study population of T2DM patients, and should be generalizable to populations included in the Truven Health MarketScan claims database, i.e., commercially-insured patients and Medicare-eligible patients subscribing to Medicare Advantage Plans, Medicare Prescription Drug Plans and Medicare Special Needs Plans. Another limitation is that potential misclassification of EHR-based covariates (e.g. under-diagnosis of conditions such as kidney disease) cannot be ruled out. We also observed a proportion of patients with missing EHR information (ranged from approximately 25% to 60% depending on the variable), though the mechanism underlying the missingness appeared to be non-differential between exposure groups. Finally, EHR linkage was not available for all matched pairs in the PS-matched population essentially breaking the 1:1 matching for a small proportion of the EHR-linked subset which could have contributed to minor residual imbalances.

In conclusion, in a claims-based study population of T2DM patients linked to EHR data for a representative subset, we found that choosing appropriate comparison groups with a new user design and using PS matching on many proxies of diabetes progression substantially improves balance in diabetes risk factors typically unmeasured in claims datasets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank John D. Seeger and Olesya Zorina for their support and input at various stages of this research.

Source of Support: EP was supported by a career development grant K08AG055670 from the National Institute on Aging. This research was supported by a research contract with Boehringer-Ingelheim. The research contract granted Brigham and Women's Hospital right to publication of results as well as final wording of the manuscript.

References

1. Prevention CfDCA. National diabetes statistics report. 2014; <http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf>. Accessed May 3, 2016.
2. Organization WH. Diabetes Fact Sheet. <http://www.who.int/mediacentre/factsheets/fs312/en/>. Accessed May 3, 2016.
3. Administration UFaD. Guidance for Industry: diabetes mellitus — evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes. 2008; <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf>. Accessed May 3, 2016.
4. Scirica BM, Bhatt DL, Braunwald E, et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. The New England journal of medicine. 10 3 2013;369(14):1317–1326. [PubMed: 23992601]
5. White WB, Cannon CP, Heller SR, et al. Alogliptin after acute coronary syndrome in patients with type 2 diabetes. The New England journal of medicine. 10 3 2013;369(14):1327–1335. [PubMed: 23992602]
6. Pfeffer MA, Claggett B, Diaz R, et al. Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome. The New England journal of medicine. 12 3 2015;373(23):2247–2257. [PubMed: 26630143]
7. Green JB, Bethel MA, Armstrong PW, et al. Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes. The New England journal of medicine. 7 16 2015;373(3):232–242. [PubMed: 26052984]
8. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. The New England journal of medicine. 11 26 2015;373(22):2117–2128. [PubMed: 26378978]
9. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. The New England journal of medicine. 7 28 2016;375(4):311–322. [PubMed: 27295427]
10. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. Journal of clinical epidemiology. 4 2005;58(4):323–337. [PubMed: 15862718]
11. Schneeweiss S, Seeger JD, Jackson JW, Smith SR. Methods for comparative effectiveness research/patient-centered outcomes research: from efficacy to effectiveness. Journal of clinical epidemiology. 8 2013;66(8 Suppl):S1–4. [PubMed: 23849143]
12. Schneeweiss S A basic study design for expedited safety signal evaluation based on electronic healthcare data. Pharmacoepidemiology and drug safety. 8 2010;19(8):858–868. [PubMed: 20681003]
13. Schneeweiss S Improving therapeutic effectiveness and safety through big healthcare data. Clinical pharmacology and therapeutics. 3 2016;99(3):262–265. [PubMed: 26659268]
14. Patorno E, Patrick AR, Garry EM, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. Diabetologia. 11 2014;57(11):2237–2250. [PubMed: 25212258]
15. Patorno E, Garry EM, Patrick AR, et al. Addressing limitations in observational studies of the association between glucose-lowering medications and all-cause mortality: a review. Drug safety. 3 2015;38(3):295–310. [PubMed: 25761856]
16. Glynn RJ, Schneeweiss S, Sturmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic & clinical pharmacology & toxicology. 3 2006;98(3):253–259. [PubMed: 16611199]
17. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology. 7 2009;20(4):512–522. [PubMed: 19487948]

18. Hansen L The Truven Health MarketScan Databases for life sciences researchers. 2017; <https://truvenhealth.com/Portals/0/Assets/2017-MarketScan-Databases-Life-Sciences-Researchers-WP.pdf>. Accessed August 22, 2017.
19. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Annals of internal medicine*. 10 15 1997;127(8 Pt 2):757–763. [PubMed: 9382394]
20. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. *Pharmacoepidemiology and drug safety*. 5 2012;21 Suppl 2:69–80. [PubMed: 22552982]
21. Patorno E GC, Zorina OI, Schneeweiss S, Bartels DB, Liu J, Seeger JD. Dynamic channeling among initiators of a recently marketed medication for type 2 diabetes mellitus (T2DM). *Pharmacoepidemiol and Drug Safety*. 2015;24:565–566.
22. Huse DM. Linking insurance claims and medical records for outcome research. White paper available at Dan.Huse@truvenhealth.com.
23. Rule AD, Larson TS, Bergstralh EJ, Slezak JM, Jacobsen SJ, Cosio FG. Using serum creatinine to estimate glomerular filtration rate: accuracy in good health and in chronic kidney disease. *Annals of internal medicine*. 12 21 2004;141(12):929–937. [PubMed: 15611490]
24. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine*. 11 10 2009;28(25):3083–3107. [PubMed: 19757444]
25. Schneeweiss S Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology and drug safety*. 5 2006;15(5):291–303. [PubMed: 16447304]
26. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *Journal of clinical epidemiology*. 10 1993;46(10):1075–1079; discussion 1081–1090. [PubMed: 8410092]
27. Schneeweiss S Developments in post-marketing comparative effectiveness research. *Clinical pharmacology and therapeutics*. 8 2007;82(2):143–156. [PubMed: 17554243]
28. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *American journal of epidemiology*. 11 1 2003;158(9):915–920. [PubMed: 14585769]
29. Eng PM, Seeger JD, Loughlin J, Clifford CR, Mentor S, Walker AM. Supplementary data collection with case-cohort analysis to address potential confounding in a cohort study of thromboembolism in oral contraceptive initiators matched on claims-based propensity scores. *Pharmacoepidemiology and drug safety*. 3 2008;17(3):297–305. [PubMed: 18215000]
30. Administration UFaD. Guidance for Industry - diabetes mellitus: developing drugs and therapeutic biologics for treatment and prevention. 2008 Accessed August 9, 2017.

Figure 1.
Flowchart of study population

A) Linagliptin versus Other DPP-4 inhibitors **B) Linagliptin versus Pioglitazone** **C) Linagliptin versus Sulfonylureas**

Figure 2. Quantitative bias analysis based on the observed residual differences in key clinical parameters

Using the observed residual difference in key clinical parameters and assuming a hypothetical observed (apparent) relative risk (ARR) of 1.5 from the claims data analysis these graphs plot the changes in RR for a range of associations between the clinical parameter observed in the EMR data and the hypothetical outcome (RR_{CD}). The RR_{CD} values reach from a non-association of $RR_{CD}=1$ to strong associations of 3.25 for each clinical parameter. For example, if the duration of diabetes of a year or longer would truly increase the risk of the outcome by 50% ($RR_{CD}=1.5$) then the observed RR in claims data

would truly be 1.52 instead of 1.50 (Panel A left). Similarly, if a 1 percent point increase in HbA1c would result in a 25% increase in outcome risk then the correct RR in claims would be 1.53 instead of 1.50 (Panel A right).

Overall the resulting changes are minor even under fairly extreme assumptions and may even cancel each other out depending on the correlation between the observed clinical parameters, for example the duration of existing diabetes may be correlated with the chance of being obese.

Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Safety* 2006;15:291-303. URL: [http://www.drugapi.org/dope-downloads/#Sensitivity Analysis](http://www.drugapi.org/dope-downloads/#Sensitivity%20Analysis)

Table 1.
Selected baseline claims-based characteristics of study participants before and after PS-matching

Claims-based patient characteristics	BEFORE PS-matching			AFTER PS-matching			BEFORE PS-matching			AFTER PS-matching		
	Linsiglipin (N=5,732)	Other DPP-4 inhibitors (N=64,095)	% or Mean (SD)	Linsiglipin (N=5,689)	Other DPP-4 inhibitors (N=5,689)	% or Mean (SD)	Linsiglipin (N=3,436)	Sulfonylureas (N=73,946)	% or Mean (SD)	Linsiglipin (N=3,410)	Sulfonylureas (N=3,410)	% or Mean (SD)
Demographics												
Age (mean, SD)	55.4 (10.9)	55.0 (11.3)		55.4 (10.9)	55.0 (11.0)		54.5 (10.9)	54.7 (11.9)		54.5 (10.9)	54.4 (10.8)	
Female	40.3	41.5		40.4	40.5		42.2	41.2		42.3	41.6	
Features of medication initiation												
Monotherapy	26.3	17.0		26.2	26.7		36.9	34.5		36.8	37.7	
Dual therapy	43.1	49.8		43.2	42.3		52.2	56.2		52.3	51.8	
Therapy with > 2 agents	30.5	33.2		30.6	31.1		10.9	9.3		10.9	10.5	
Dual therapy with metformin ¹	29.3	40.2		29.4	28.5		43.7	50.8		43.9	43.7	
Concomitant initiation of other antidiabetic agents	8.9	23.0		8.9	9.4		10.1	22.7		10.1	9.7	
Concomitant initiation of metformin	6.4	20.2		6.3	6.7		8.7	20.9		8.7	8.5	
Concomitant initiation of insulin	0.8	1.1		0.8	0.7		0.9	1.6		0.9	0.9	
Current use of other antidiabetic agents ²	67.1	65.4		67.2	67.1		54.3	44.8		54.4	54.0	
Current use of metformin	50.3	51.4		50.4	49.9		45.1	38.7		45.3	44.7	
Current use of insulin	0.8	1.1		0.8	0.7		0.9	1.6		0.9	0.9	
Comorbidities at Baseline												
Charlson comorbidity score (mean, SD)	1.4 (0.9)	1.3 (0.8)		1.4 (0.9)	1.3 (0.9)		1.3 (0.9)	1.2 (0.8)		1.3 (0.9)	1.3 (0.8)	
Diabetic nephropathy	3.9	2.5		3.9	4.1		3.1	2.5		3.1	3.1	
Diabetic retinopathy	3.0	2.8		3.0	3.0		2.6	2.2		2.5	2.5	
Diabetic neuropathy	5.8	4.9		5.8	6.1		4.7	4.4		4.8	5.0	
Peripheral vascular disease	1.6	1.5		1.6	1.8		1.3	1.3		1.3	1.1	
Erectile dysfunction	2.1	1.8		2.1	2.2		2.0	1.6		2.0	1.9	
Diabetic foot	1.0	0.9		1.0	0.9		0.6	1.0		0.6	0.6	
Skin infections	4.4	4.4		4.3	4.3		4.3	4.9		4.2	4.1	
Hypoglycemia	6.1	6.1		6.1	6.1		5.4	7.2		5.3	5.2	
Hypertension	35.9	32.9		35.9	36.3		34.7	28.0		34.7	34.0	
Hyperlipidemia	30.8	28.3		30.8	30.9		29.7	21.4		29.8	29.2	
Coronary atherosclerosis	8.3	7.7		8.3	8.3		7.6	7.3		7.5	7.9	
Acute myocardial infarction	0.6	0.8		0.6	0.7		0.6	0.9		0.6	0.4	
Old myocardial infarction	0.7	0.7		0.7	0.6		0.5	0.7		0.5	0.6	

PS: propensity score; SD: standard deviation; CABG: coronary artery bypass grafting; PTCA: percutaneous transluminal coronary angioplasty; ACE: angiotensin converting enzyme; ARBs: angiotensin receptor blockers.

1 Dual therapy with metformin was defined as concomitant initiation or current use of metformin on the day of linagliptin or comparator initiation, i.e., having metformin days' supply overlapping with the
day of drug initiation, and no use of other glucose-lowering agents.

2 Current use was defined as having days' supply available for an agent on the day of linagliptin or comparator initiation without a grace period.

Table 2.

Baseline EHR-based characteristics of study participants before and after PS-matching

EHR-based patient characteristics	BEFORE PS-matching			AFTER PS-matching			BEFORE PS-matching			AFTER PS-matching			BEFORE PS-matching			AFTER PS-matching		
	Linagliptin (N=243)	% or Median (25th-75th IQR)	Other DPP-4 inhibitors (N=3,041)	Linagliptin (N=240)	% or Median (25th-75th IQR)	Other DPP-4 inhibitors (N=271)	Linagliptin (N=150)	% or Median (25th-75th IQR)	Sulfonylureas (N=3,050)	Linagliptin (N=148)	% or Median (25th-75th IQR)	Sulfonylureas (N=159)	Linagliptin (N=205)	% or Median (25th-75th IQR)	Pioglitazone (N=530)	Linagliptin (N=176)	% or Median (25th-75th IQR)	Pioglitazone (N=165)
Health Behaviors																		
Smoking Status (%)																		
Current	9.1		7.6	8.8		8.9	11.3		9.4	10.8		10.1	10.7		7.9	9.7		7.9
Past	12.3		12.2	12.1		11.1	10.7		10.9	10.8		8.2	12.2		10.2	11.9		12.7
Never	32.9		35.5	33.3		38.4	34.7		33.3	35.1		37.1	32.7		30.8	32.4		33.9
Unknown	5.8		4.9	5.8		4.1	6.0		6.5	6.1		5.7	6.3		5.1	6.8		6.1
Missing	39.9		39.8	40.0		37.6	37.3		39.8	37.2		39.0	38.0		46.0	39.2		39.4
BMI (%)																		
Normal or underweight (<25 Kg/m ²)	2.5		4.0	2.5		5.2	3.3		4.1	3.4		6.3	2.4		3.4	2.3		1.8
Overweight (25 to 29.9 Kg/m ²)	11.5		11.8	11.3		10.7	10.7		14.0	10.8		12.6	13.7		15.5	12.5		16.4
Obese (30 to 39.9 Kg/m ²)	35.4		35.1	35.8		37.3	36.0		34.2	36.5		40.3	35.1		29.6	34.1		35.2
Severe Obesity (>=40 Kg/m ²)	16.5		15.5	16.3		18.5	16.0		15.4	15.5		9.4	16.1		10.9	15.3		10.9
Missing	34.2		33.6	34.2		28.4	34.0		32.3	33.8		31.4	32.7		40.6	35.8		35.8
Duration of diabetes² (%)																		
Less than 1 year	11.9		13.7	12.1		10.7	16.0		15.4	16.2		15.1	12.7		11.1	11.4		17.6
1.00-2.99 years	11.5		14.2	11.7		14.4	12.7		13.6	12.8		15.1	11.2		10.0	12.5		12.1
3.00-4.99 years	7.8		7.8	7.9		9.2	7.3		8.4	7.4		7.5	7.8		8.3	8.0		9.1
5.00-6.99 years	4.9		5.1	5.0		5.2	4.0		4.8	4.1		5.7	3.9		5.7	4.0		5.5
7+ years	5.3		5.3	5.4		4.1	5.3		4.8	5.4		5.7	5.4		5.8	5.7		5.5
Missing	58.4		53.9	57.9		56.5	54.7		53.0	54.1		50.9	59.0		59.1	58.5		50.3
Laboratory test results³ (Median, IQR)																		
HbA1c, %	7.8 (7.0-9.0)		8.1 (7.1-9.5)	7.8 (7.0-9.0)		7.9 (7.1-9.1)	7.8 (7.0-8.8)		8.1 (7.1-9.7)	7.8 (7.0-8.9)		8.1 (7.2-9.6)	7.9 (7.1-9.1)		8.2 (7.1-9.9)	8.0 (7.1-9.1)		8.2 (7.1-9.9)
Creatinine, mg/dL	0.9 (0.8-1.1)		0.9 (0.7-1.0)	0.9 (0.8-1.1)		0.9 (0.7-1.1)	0.9 (0.7-1.1)		0.9 (0.7-1.0)	0.9 (0.7-1.1)		0.9 (0.7-1.0)	0.9 (0.7-1.1)		0.9 (0.8-1.0)	0.9 (0.8-1.1)		0.9 (0.7-1.0)
eGFR, mL/min/1.73 m ²	98 (84-108)		102 (92-115)	101 (92-115)		103 (92-116)	99 (86-109)		102 (91-116)	102 (93-117)		102 (94-115)	99 (90-110)		102 (92-119)	102 (93-116)		104 (96-118)
Total cholesterol level, mg/dL	168 (146-197)		171 (147-201)	168 (146-196)		171 (152-200)	175 (146-209)		174 (150-206)	171 (146-200)		182 (154-220)	168 (146-196)		174 (147-206)	167 (145-196)		170 (145-195)
HDL, mg/dL	40 (34-49)		42 (35-50)	40 (34-49)		42 (36-49)	41 (35-50)		41 (35-50)	41 (35-50)		40 (34-48)	40 (33-49)		41 (35-48)	40 (33-49)		41 (33-46)
LDL, mg/dL	100 (76-122)		95 (73-120)	100 (76-121)		97 (70-120)	104 (80-127)		98 (76-123)	104 (80-124)		102 (74-126)	100 (76-121)		94 (73-119)	97 (73-116)		97 (79-115)
Triglycerides, mg/dL	171 (110-233)		156 (105-226)	171 (109-233)		156 (103-216)	171 (115-231)		161 (109-240)	172 (115-232)		180 (125-399)	172 (114-235)		158 (102-236)	171 (110-235)		152 (111-263)

EHR-based patient characteristics	BEFORE PS-matching		AFTER PS-matching		BEFORE PS-matching		AFTER PS-matching		BEFORE PS-matching		AFTER PS-matching	
	Linagliptin (N=243)	Other DPP-4 inhibitors (N=3,041) % or Median (25th-75th IQR)	Linagliptin (N=240)	Other DPP-4 inhibitors (N=271) % or Median (25th-75th IQR)	Linagliptin (N=150)	Sulfonylureas (N=3,050) % or Median (25th-75th IQR)	Linagliptin (N=148)	Sulfonylureas (N=159) % or Median (25th-75th IQR)	Linagliptin (N=205)	Pioglitazone (N=530) % or Median (25th-75th IQR)	Linagliptin (N=176)	Pioglitazone (N=165) % or Median (25th-75th IQR)
Blood pressure³ (Median, IQR)												
Systolic BP, mmHg	130 (118–140)	128 (120–140)	130 (118–140)	127 (118–138)	130 (118–140)	130 (120–140)	130 (118–140)	128 (120–140)	130 (120–140)	130 (120–141)	130 (120–140)	129 (120–140)
Diastolic BP, mmHg	80 (72–86)	80 (72–85)	80 (72–86)	78 (70–82)	80 (72–86)	80 (72–86)	80 (72–86)	80 (72–84)	80 (72–86)	80 (72–86)	80 (72–86)	80 (72–85)
Patients with EHR information available												
Laboratory test results (%)												
HbA1c	58.4	62.7	58.8	69.0	59.3	63.0	59.5	67.3	61.5	58.7	61.9	60.6
Creatinine	63.8	67.5	63.8	72.7	64.0	67.4	64.2	73.0	66.8	63.4	65.9	67.9
Total cholesterol level	58.8	65.2	58.8	69.4	60.0	63.7	60.1	71.7	61.5	62.1	61.4	64.2
HDL	49.0	54.3	49.2	59.4	50.0	53.4	50.0	57.2	51.7	50.6	51.7	53.9
LDL	49.0	54.3	49.2	59.4	50.0	53.4	50.0	57.2	51.7	50.6	51.7	53.9
Triglycerides	58.4	64.9	58.3	69.0	59.3	64.0	59.5	71.7	61.0	61.7	60.8	64.2
Blood pressure (%)												
Systolic BP	77.8	78.9	77.9	81.2	78.7	80.0	79.1	80.5	79.5	76.6	77.8	81.8
Diastolic BP	77.8	78.8	77.9	81.2	78.7	80.0	79.1	80.5	79.5	76.6	77.8	81.8

PS: propensity score; IQR: interquartile; BMI: body mass index; HbA1c: hemoglobin A1c; eGFR: estimated glomerular filtration rate; HDL: high-density lipoprotein; LDL: low-density lipoprotein; BP: blood pressure.

¹ Captured recording closest to cohort entry, within two years prior to and including the date of cohort entry.

² Duration of diabetes was calculated by capturing the earliest record of a type 2 diabetes mellitus (T2DM) diagnosis in the EHR using all available information prior to treatment initiation. Diagnosis of T2DM is recorded as a structured field in the EHR along with an “onset date”.

³ Captured recording closest to cohort entry, prior to and including the date of cohort entry.