Published in final edited form as:

Med Care. 2014 March ; 52(3): 280–287. doi:10.1097/MLR.0000000000000064.

Assessing the Impact of Propensity Score Estimation and Implementation on Covariate Balance and Confounding Control Within and Across Important Subgroups in Comparative Effectiveness Research

Cynthia J. Girman^{1,2}, Mugdha Gokhale², Tzuyung Doug Kou^{3,*}, Kimberly G Brodovicz¹, Richard Wyss², and Til Stürmer²

¹Department of Epidemiology, Merck Sharp & Dohme, North Wales, PA

²Department of Epidemiology, University of North Carolina, Chapel Hill, NC

³Department of Global Pharmacovigilance&Epidemiology, Bristol Meyers Squibb, Hopewell, NJ

Abstract

Purpose—Researchers are often interested in estimating treatment effects in subgroups controlling for confounding based on a propensity score (PS) estimated in the overall study population.

Objective—To evaluate covariate balance and confounding control in sulfonylurea (SU) versus metformin (MET) initiators within subgroups defined by cardiovascular disease history (CVD) comparing an overall PS with subgroup-specific PSs implemented by 1:1 matching and stratification.

Corresponding Author: Cynthia J Girman, DrPH, Department of Epidemiology, UG1D-60, Merck Sharp & Dohme, 351 N Summerville Pike, North Wales, PA 19454.

*Dr. Kou was at Merck Sharp & Dohme when this work was performed

Parts of this research were presented as an oral presentation at the International Conference on Pharmacoepidemiology and Therapeutic Risk Management in Barcelona on August 24, 2012, and in poster form at the DECIDE Methods workshop, Agency for Healthcare Research and Quality, June 12–13, 2012.

Conflict of Interest:

Commercial Employment: **Merck: Girman, Brodovicz; BMS: Kou**

Advisory Board: **ASSESS Steering Comm (RTI, GSK): Stürmer**

Industry-sponsored grants: **GSK: Stürmer; Sanofi: Stürmer**

Stock ownership/options: **Merck: Girman, Brodovicz; BMS: Kou**

M Gokhale and R Wyss have no conflicts of interest to disclose

CONFLICT OF INTEREST:

- Project was funded by grants to UNC from:
 - **Merck Research Laboratories**
 - **NIA R01 AG023178 "Propensity Scores & Preventive Drug Use in the Elderly"**
- Personal/financial relationships relevant to presentation in past 12 months/during the conduct of the study:
 - Commercial Employment: **Merck: Girman, Brodovicz; BMS: Kou**
 - Advisory Board: **ASSESS Steering Comm (RTI, GSK): Stürmer**
 - Industry-sponsored grants: **GSK: Stürmer; Sanofi: Stürmer**
 - Stock ownership/options: **Merck: Girman, Brodovicz; BMS: Kou**

Methods—We analyzed younger patients from a US insurance claims database and older patients from two Medicare (Humana Medicare Advantage, fee-for-service Medicare Parts A, B and D) datasets. Confounders and risk factors for acute myocardial infarction (AMI) were included in an overall PS and subgroup PSs with and without CVD. Covariate balance was assessed using the average standardized absolute mean difference (ASAMD).

Results—Compared to crude estimates, ASAMD across covariates was improved 70–94% for stratification for Medicare cohorts and 44–99% for the younger cohort, with minimal differences between overall and subgroup-specific PSs. With matching, 75–99% balance improvement was achieved regardless of cohort and PS, but with smaller sample size. Hazard Ratios within each CVD subgroup differed minimally among PS and cohorts.

Conclusion—Both overall PSs and CVD subgroup-specific PSs achieved good balance on measured covariates when assessing the relative association of diabetes monotherapy with nonfatal MI. PS matching generally led to better balance than stratification, but with smaller sample size. Our study is limited insofar as crude differences were minimal, suggesting that the new user, active comparator design identified patients with some equipoise between treatments.

Keywords

Comparative effectiveness research; new user design; propensity scores

INTRODUCTION

In comparative effectiveness research, covariate selection is critical in deriving propensity scores (PSs) to control for confounding [1–5]. However, the effect of covariates on treatment may vary by subgroup. The impact of using a PS estimated in the overall cohort for pre-specified subgroup analyses is not well understood. There may be efficiency and validity gains in estimating separate PSs for different subgroups but subgroup-specific PS derivation within multiple subgroups may not be practical, and correct specification of PS with multiple subgroups may be difficult, resulting in greater imbalances.

A recent study assessed whether using an overall cohort PS in subgroup analyses was feasible, finding that a cohort PS resulted in similar effect estimates as a subgroup-specific PS when the overall cohort PS is correctly specified, especially in larger subgroups [6]. That study focused on simulations and a specific example of antipsychotics and short term mortality in two cohorts, however, and NSAIDs and severe gastrointestinal complications in one cohort. In patients with type 2 diabetes mellitus (T2DM), there is mixed evidence about the risk of myocardial infarction (MI) associated with sulfonylureas (SU) relative to metformin (MET) [7–9]. An important potential confounder for such analyses is history of cardiovascular disease (CVD) [10]. In this study, we extended the empirical work of Rassen [6] to T2DM using the design of the mini-Sentinel protocol [10]. We used three different administrative databases to assess the association of initiating SU vs MET monotherapy (new users) on nonfatal MI after conditioning (e.g., matching or stratifying) on PSs derived within an overall cohort compared to conditioning on CVD subgroup-specific PSs with different conditioning approaches.

METHODS

We studied younger (aged < 65 years) patients from a large, US-based administrative claims database of a commercially insured population and their dependents (Clinformatics DataMart, OptumInsight™ previously known as i3 InVision™ Data Mart, Ingenix, Inc). This database is used extensively for pharmacoepidemiological research and contains more than 45 million unique members spanning almost 10 years, of which more than 29 million patients were continuously enrolled for at least 12 months.

We also studied older adults from Humana (Humana MA) with Medicare Advantage Prescription Drug plans (MAPD), who had both pharmacy and medical claims, and prescription drug plan (PDP) claims. In 2010, Humana had over 3.2 million Medicare members including 2 million Medicare Advantage members, and there has typically been little plan turnover [11].

Finally, older adults from an approximate 20% random sample of Medicare beneficiaries 65 years or older enrolled in the stand-alone prescription drug plan (PDP) and fee-for service (FFS) components of Medicare claims data (Medicare FFS) from 1/1/2007 to 12/31/2009 were analyzed for this study [12–13].

Within the commercially insured younger cohort, we identified adult patients aged 25–64 years with T2DM who were initiating monotherapy with SU or MET. In Humana Medicare Advantage and Medicare FFS claims, we identified new users of the same therapies who were ≥65 years old. T2DM was identified as those with at least one of the following: a) at least one inpatient admission for which the principal diagnosis was recorded as diabetes; b) at least one inpatient admission for which any diagnosis was recorded as diabetes coupled with another criterion (oral antihyperglycemic agent (OHA) prescription, inpatient or outpatient type 2 diabetes diagnoses, or laboratory result); c) at least two outpatient facility or physician office visit claims for which a diagnosis was recorded as diabetes; d) at least one inpatient or outpatient claim of diabetes AND at least one prescription of OHA medication; or e) at least one laboratory result of blood glucose >200 mg/dl OR HbA_{1C} > 6.5% [14–15]. Data were extracted for patients with T2DM with claims between January 1, 2003 and December 31, 2010 for the commercially insured (6-month look-back period to July 1, 2002), Jan 1, 2008 to July 31, 2011 for Humana MA (6-month look-back to July 1, 2007) and from July 1, 2007 to Dec 31, 2009 in the Medicare FFS (6-month look-back period to Jan 1, 2007).

Patients were excluded if they met any of the following: a) type 1 diabetes (identified by the 5th digit number of ICD-9 code), juvenile diabetes, malnutrition-associated diabetes, drug-induced diabetes or gestational diabetes at any time in the database without subsequent type 2 diabetes code, or ketoacidosis; b) 25 years or younger at first notation of a diabetes code; c) diabetes ICD-9 code associated with a blood draw within 15 days with corresponding laboratory data indicating the glucose and HbA_{1C} results did not meet criteria for diabetes, and no prior or subsequent diabetes codes or OHA use; d) any OHA prescription within 7 months prior to index date (first prescription of SU or MET during the study window); e) not having both medical and pharmacy benefits. SU and MET were identified using NDC

codes recorded as part of paid prescription claims; hence, it is assured that the physician prescribed and the pharmacy filled the prescription.

Eligibility in terms of initiation of monotherapy required at least two consecutive prescriptions of MET or SU within 120 days. This improves the likelihood that the patient actually took the product given it was refilled. No patients meeting this criteria were excluded. Follow-up began at the 2nd qualifying prescription and continued until the patient experienced an MI or was censored at the earliest occurrence of: (1) end of patient's enrollment or death, or (2) end of the study window. Patients who transitioned to Medicare during follow-up in the younger cohort were censored at transition. In the primary analysis, patients were not censored if they switched, discontinued or augmented treatment with one of the drugs being compared. However, similar results were found in a secondary 'as treated' analysis where patients were censored when they discontinued drug, switched or augmented with SU or MET. Stopping was defined as having no additional prescription within the days supply plus a grace period of 30 days after the previous prescription.

The primary outcome of nonfatal acute myocardial infarction (AMI) was defined as a diagnosis based on hospital discharge ICD-9 code: 410.x0 or 410.x1 in the first or second position, or DRG codes 121, 122, 123 [10]. No length of stay criterion was applied. Other cardiovascular outcomes typically included in composite endpoints of cardiovascular outcomes trials are typically not available in databases (fatal MI, other cardiovascular death), and revascularizations, unless urgent (distinction not available in database), are typically considered a 'soft' outcome subject to substantial geographic variability and physician subjectivity. Hence, these outcomes were not studied.

The primary subgroup was defined as history of cardiovascular disease in the prior 6 months. History of cardiovascular disease included any cardiovascular disease, i.e., MI, CABG, PTCA, unstable angina and peripheral arterial disease (PAD) [10]. In addition, we included the following as prevalent CVD: lower extremity revascularization (38.18), endarterectomy (39.25, 39.29), lower extremity bypass (84.10 – 84.17), and procedure codes for carotid revascularization (38.11, 38.12), endarterectomy (00.61, 00.63), carotid bypass (39.28).

Baseline comorbidities were determined by ICD-9 codes in the 6-month period before the index date, and included complications of diabetes, cardiovascular disease, liver disease, lung disease, infection, cancer and other diseases. Confounders and risk factors for nonfatal MI included in the PS models were selected using prior clinical and epidemiologic knowledge based on published literature (*a priori*), and consisted of age, sex and diagnoses (coded at the 3-digit level), procedures, and drugs (coded based on generic drug name) that had a prevalence of >2% and <98% in the 6-month baseline period to allow evaluation of which individually had the greatest imbalance. Also included in the traditional Medicare analyses were nursing home residence, number of hospitalizations, emergency department visits, number of outpatient physician visits, and number of distinct prescription drugs dispensed. Interactions were not accounted for in derivations of PS, since the main effects PS model balanced the groups fairly well. In the base-case scenario, we only included

variables available in all three databases. In Medicare analyses, sensitivity analyses including variables unique to that database showed similar results.

Within each database population, we estimated three separate PS for predicting treatment initiation with SU versus MET: an overall PS (including prior CVD as a covariate) and 2 separate PSs for the prior CVD subgroups. We implemented these PS within each subgroup of CVD using a greedy algorithm for 1:1 matching [16] and equally-sized PS deciles (quintile for Medicare FFS) stratification after restricting to an overlapping PS range. Because the true effect (necessary to compare bias) is unknown in these empirical datasets, we primarily compared the balance within CVD subgroups based on the overall vs. subgroup-specific PS. We also compared overall vs. subgroup-specific PSs with respect to the point estimate of the hazard ratio (HR) and 95% confidence interval width for treatment with SU vs. MET. The average standardized absolute mean difference (ASAMD) was used to assess balance within CVD subgroups [17–19], with lower ASAMD indicating better balance. For the matched analyses, the ASAMD was calculated as the absolute value of the difference in means (percentages) divided by a pooled standard deviation for each covariate, averaged across covariates. For the stratified analyses, we first calculated the ASAMD for each covariate within each PS stratum and then combined differences within covariate across strata, finally averaging over all covariates. We used Kaplan Meyer plots to assess the proportional hazards assumption and fit Cox proportional hazards models for time to event (acute non-fatal MI) over the entire follow-up if appropriate (PH assumption met).

RESULTS

Younger cohort

Baseline characteristics of the younger cohort were generally comparable for MET and SU monotherapy new users, with the most pronounced differences at baseline before matching observed in gender, age, cardiovascular and renal diseases, non-skin cancers and retinopathy (Table 1).

There was minimal non-overlap between propensity scores overall (Figure 1), reflecting minimal imbalance in this younger cohort (crude average standardized absolute mean difference (ASAMD) in subgroups without CVD history: 0.073, and with prior CVD: 0.099). CVD history was related to SU vs MET treatment in the PS model (1.49 (1.45, 1.53)). With matches achieved for over 98% of the SU patients, baseline covariates were fairly well balanced with overall PS for the full cohort and within the subgroup without CVD history (ASAMD: 0.01 for both) as well as when the subgroup-specific PS was used (ASAMD: 0.01 for subgroup without CVD history) (Figure 2). Within the prior CVD subgroup where 92% of SU patients were matched, balance was only slightly better using the subgroup-specific PS (ASAMD: 0.01) than using the overall PS (ASAMD: 0.02).

With PS stratification, balance was achieved with both overall PS (ASAMD: 0.04) and subgroup-specific PS (ASAMD: 0.02) within the subgroup without CVD history, as well as in the subgroup with CVD history (ASAMD: 0.02 for both overall and subgroup-specific PS). ASAMD estimates were consistently lower for 1:1 matching than for stratification.

The crude HRs for non-fatal MI in the younger cohort with SU vs MET were 1.08 (1.05, 1.11) overall, 1.08 (1.04, 1.11) within the subgroup with no prior CVD history, and 1.05 (0.99, 1.13) for those with prior CVD (Figure 3). HR estimates were highly consistent (HR range: 1.1–1.15) between overall and subgroup-specific PS for 1:1 matching, differing by <1% in those without prior CVD, and only 3.5% in those with prior CVD. PS stratification showed similar patterns.

Older Medicare cohorts

Baseline characteristics of the older cohorts were less similar than for the younger cohort, with the most pronounced differences before matching observed in age, gender, heart disease (prior MI, peripheral vascular disease, pacemakers, congestive heart failure), renal disease, bone fractures, and diabetes complications, likely due to the higher proportion of older age (>80 years) in SU users (overall (ASAMD): 0.35 in Humana MA, 0.11 in Medicare FFS) (Table 2). CVD history was related to treatment in the PS model (1.60 (1.56, 1.64) for Humana MA; (1.27 (1.20, 1.33)) for Medicare FFS). Only 82% to 88% of the SU patients overall were matched. Balance was achieved with 1:1 matching in Humana MA for those without prior CVD where 90% of SU patients were matched (AMASD: overall PS: 0.01; subgroup-specific PS: 0.001), but imbalance was somewhat greater for patients with prior CVD where only 72% of SU patients were matched (overall PS: AMASD: 0.01, subgroup-specific PS: 0.01). PS distributions after matching were superimposable, however (not shown). For Medicare FFS, balance was achieved for those without prior CVD where 96% of SU initiators were matched (ASAMD using overall PS: 0.0185; ASAMD using subgroup-specific PS: 0.01) and among patients with prior CVD where only 84% of the SU users were matched (ASAMD overall PS: 0.01; subgroup-specific PS: 0.01). Again, PS distributions were superimposable. In both subgroups, balance was only minimally better with the subgroup-specific PS compared to overall PS.

For stratified analyses, balance in the Humana MA cohort using subgroup-specific PS (AMASD: 0.034) was almost identical to using overall PS (0.03) for subgroup without prior CVD as well as for the subgroup with CVD history (subgroup-specific PS AMASD: 0.03 and overall PS AMASD: 0.02). For Medicare FFS stratified analyses, balance was comparable among those without prior CVD (ASAMD overall PS: 0.03; subgroup-specific PS: 0.04) as well as among those with prior CVD (0.03 and 0.03 for overall and subgroup-specific PS). ASAMD estimates were lower for matching than for stratification across subgroups and cohorts.

Crude HR estimates ($\pm$ 95% CI) for non-fatal MI with SU vs. MET were higher for the older cohorts (HR: 2.1 (1.9–2.3) for Humana MA and HR: 2.2 (1.9–2.5) for Medicare FFS) than for the younger cohort (Figure 3). HR estimates were consistent between overall and subgroup-specific PS for 1:1 matching within those without prior CVD (HR: 1.4–1.9) and those with prior CVD (HR: 1.5–1.7), differing by 3–8% within the older cohorts. PS stratification showed similar patterns, differing by <1% within subgroups. Including race (not available in other datasets) in PS derivation for FFS did not change results.

DISCUSSION

Based on U.S. cohorts of younger and elderly T2DM patients initiating MET or SU, an overall PS balanced nearly as well in matching within subgroups as a subgroup-specific PS in assessing the association between monotherapy and MI. After PS implementation, subgroup-specific PS showed slightly better balance within the subgroup with prior CVD, but the differences were small, and HRs for MI were highly consistent across cohorts, regardless of whether matching or stratification was applied or overall vs subgroup-specific PS was used.

If the relationship between covariates and treatment varies by a strong confounder, a subgroup-specific PS could potentially provide better balance covariates in subgroup matched analyses, especially when the influence of covariates on treatment selection is not comparable across subgroups. In our case, particularly in the younger cohort, baseline differences between groups were minimal, perhaps reflecting the new user (initiators of therapy) design applied and the assessment of monotherapy, resulting in predominantly newly diagnosed T2DM with minimal complications and comorbidities. Even in the elderly diabetes cohorts, differences between groups were not pronounced despite greater prevalence of comorbidities and complications. The lower prevalence of cardiovascular disease in the younger cohort may help explain the somewhat greater imbalance observed with stratification methods, and the consistency in HRs for this cohort.

We hypothesized that CVD history could be a strong confounder given its strong predictive properties for nonfatal AMI and CV outcomes in general, and the potential to influence prescribing of oral anti-hyperglycemic agents [10]. A few studies have suggested that MET may reduce cardiovascular risk, while other studies suggest that mortality may be increased with use of SU [7–9]. We assessed the strength of the association of CVD history with SU and MET initiation, and found it to be moderately but significantly associated. CVD history may be of greater importance in the Humana cohort, given the significant reduction in ASAMD when stratifying by CVD subgroup.

This study included patients treated with SU or MET monotherapy who were naive to diabetes pharmacological treatment. However, treatment complexity in diabetes is strongly related to duration of diabetes and microvascular complications, and also to cardiovascular outcomes. We expect that a greater degree of confounding would be more likely in patients with longer duration diabetes and with more complex treatment regimens. In addition, we expected significant channeling bias due to concerns about hypoglycemia when adding SU to existing diabetes regimens in patients with longer term diabetes, especially in the elderly. We did not assess more complex regimens of SU or MET initiators on a background of other diabetes therapies as our objective was to study monotherapy and MET is overwhelmingly used first-line.

Our results are in agreement with Rassen et al [6], who showed that a correctly specified cohort PS can be applied in subgroup analyses, based on simulations and two empirical examples. This was supported by small differences (<10%) in estimated log odds ratios in empirical analyses when the full cohort PS was applied to a subgroup analysis, similar to our

findings. Differences were larger with small subgroups ($n < 1000$) or rare outcomes, which was confirmed in simulation analyses. However, the strength of confounding with the subgroups they assessed (gender, age and risk) was unclear, whereas we report the association of CVD in the PS model. In addition, their simulations were based on outcome model adjustment for PS, not on applying 1:1 PS matching and stratification, as in our analyses. Neither of our studies included interactions in modeling PS, an area that deserves further research. However, because of the consistency of our findings between overall and subgroup-specific PS, adding interactions would be unlikely to change our results. Finally, our results address an important methodological question in the area of diabetes, and one highly relevant to the mini-Sentinel protocol.

Certain limitations should be considered in interpreting these results. Our analyses are based on a younger and two elderly cohorts with T2DM, in three separate claims databases, which could not be combined due to inherent differences in the populations. Generalizability is hence limited to the insured in the younger group, and in the elderly, to those who choose the Medicare Advantage or FFS plans. We did not have information on ethnicity except in Medicare FFS where it appeared to not influence results. Balance was assessed using the commonly used ASAMD which has the advantage of being an overall balance metric, but does not account for strength of association to outcome. However, we included the most important confounders in PSs which indirectly assures inclusion of covariates associated with outcome. The small reductions in ASAMD with the various techniques are difficult to interpret without standards, and could benefit from further research. Finally, information on fatal MI, other cardiovascular deaths and other cardiovascular events were not available. Our primary goal to assess differences when applying an overall vs a subgroup-specific PS, not to assess differences in AMI between SU vs. MET *per se*. Results may be dependent on the disease, outcome and confounder as well as the exposures studied. Further research is needed through other empirical examples and simulations in which the covariates may have different relationships with the potential subgroup variable.

The strengths of our analyses are that three separate claims databases and two age cohorts were analyzed, and several approaches to control for confounding were used. These cohorts and the available information on medical history, prior and concomitant therapies, and comorbidities, reflect the types of studies in which propensity score methods are typically applied to control for confounding in comparative effectiveness research.

CONCLUSION

In this study, balance of baseline covariates between treatment groups was achieved with both overall cohort and subgroup-specific PSs in analyses of the effect of SU versus MET monotherapy with nonfatal MI in diabetes patients across three cohorts and within and across CVD subgroups, with only minimally better balance for subgroup-specific PSs. Matching achieved better balance than stratification on PS but with expected loss in sample size. Our study is limited insofar as crude differences were minimal, suggesting that the new user, active comparator design identified patients with some equipoise between treatments.

Acknowledgments

SPONSOR: This research was funded by grants from Merck & Co., Inc, and NIA R01 AG023178 "Propensity Scores & Preventive Drug Use in the Elderly"

Authors would like to acknowledge the programming efforts of Michael Senderak, in extracting and creating the diabetes cohorts in the younger United Health Care and Humana MA for these analyses. In addition, authors thank Virginia Pate for her help in creating the diabetes cohort for the analysis using Medicare FFS data..

References

1. Stürmer T, Schneeweiss S, Brookhart MA, Rothman KJ, Avorn J, Glynn RJ. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal anti-inflammatory drugs and short-term mortality in the elderly. *American Journal of Epidemiology*. 2005; 161:891–898. [PubMed: 15840622]
2. Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *Journal of Clinical Epidemiology*. 2006; 59:437–447. [PubMed: 16632131]
3. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *American Journal of Epidemiology*. 2006; 163:1149–1156. [PubMed: 16624967]
4. Brookhart Alan M, Sturmer Til, Glynn RJ, Rassen J, Schneeweiss S. Confounding Control in Healthcare Database Research: Challenges and Potential Approaches. *Med Care*. 2010; 48(6):114–120.
5. Patrick AR, Schneeweiss S, Brookhart MA, Glynn RJ, Rothman KJ, Avorn J, Stürmer T. The implications of propensity score variable selection strategies in pharmacoepidemiology: an empirical illustration. *Pharmacoepidemiology and Drug Safety*. 2011; 20:551–559. [PubMed: 21394812]
6. Rassen JA, Glynn RJ, Rothman KJ, Setoguchi S, Schneeweiss S. 2011 Applying propensity scores estimated in a full cohort to adjust for confounding in subgroup analyses. *Pharmacoepidemiology and Drug Safety*. 2012; 21:679–709.
7. Horsdal, Henriette Thisted; Søndergaard, Flemming; Johnsen, Søren Paaske; Rungby, Jørgen. Antidiabetic treatments and risk of hospitalisation with myocardial infarction: a nationwide case–control study. *Pharmacoepidemiology and Drug Safety*. Apr; 2011 20(4):331–337. [PubMed: 21442682]
8. Evans JM, Ogston SA, Emslie-Smith A, Morris AD. Risk of mortality and adverse cardiovascular outcomes in type 2 diabetes: a comparison of patients treated with sulfonylureas and metformin. *Diabetologia*. 2006; 49(5):930–6. [PubMed: 16525843]
9. Belcher G, Lambert C, Goh KL, et al. Cardiovascular effects of treatment of type 2 diabetes with pioglitazone, metformin and gliclazide. *Int J Clin Pract*. 2004; 58:833–837. [PubMed: 15529516]
10. Fireman B, Toh S, Butler MG, Go AS, Joffe HV, Graham DJ, Nelson JC, Daniel GW, Selby JV. A protocol for active surveillance of acute myocardial infarction in association with use of a new anti-diabetic pharmaceutical agent. *Pharmacoepidemiology and Drug Safety*. 2012; 21:282–290. [PubMed: 22262618]
11. Humana Inc. Healthcare claims for Medicare and commercial members from 2007–2011. 2012. Retrieved from Humana Enterprise Data Warehouse [Accessed January, 2011]
12. Centers for Medicare & Medicaid Services. Medicare program-general information: overview. Centers for Medicare & Medicaid Services; Web site <http://www.cms.gov/MedicareGenInfo/> [Accessed November 5, 2012]
13. Centers for Medicare & Medicaid Services. Prescription drug coverage-general information: overview. Centers for Medicare & Medicaid Services; Web site <http://www.cms.gov/PrescriptionDrugCovGenIn/> [Accessed November 5, 2012]

14. Girman CJ, Kou TD, Cai B, Alexander CM, O'Neill EA, Williams-Herman DE, Katz L. Patients with type 2 diabetes mellitus have higher risk for acute pancreatitis compared with those without diabetes. *Diabetes Obes Metab*. 2010; 12:766–771. [PubMed: 20649628]
15. Cai B, Katz L, Alexander CM, Williams-Herman D, Girman CJ. Characteristics of patients prescribed sitagliptin and other oral antihyperglycaemic agents in a large US claims database. *International J Clin Pract*. 2010; 64:1601–1608.
16. Parsons, LS. [Accessed January, 2012] Reducing bias in a propensity score matched-pair sample using greedy matching techniques. 2001. <http://www2.sas.com/proceedings/sugi26/p214-26.pdf>
17. Belitzer SV, Martens EP, Pestman WR, Groenwold RHH, de Boer A, Klungel OH. Measuring balance and model selection in propensity score methods. *Pharmacoepidemiology and Drug Safety*. 2011; 20:1115–1129. [PubMed: 21805529]
18. Groenwold RHH, de Vries F, de Boer A, Pestman WR, Rutten FH, Hoes AW, Klungel OH. Balance measures for propensity score methods: a clinical example on beta agonist use and the risk of myocardial infarction. *Pharmacoepidemiology and Drug Safety*. 2011; 20:1130–1137. [PubMed: 21953948]
19. Lunt M, Solomon D, Rothman K, et al. Different Methods of Balancing Covariates Leading to Different Effect Estimates in the Presence of Effect Modification. *Am J Epidemiol*. 2009 Apr 1; 169(7):909–17. Epub 2009 Jan 19. [PubMed: 19153216]

Key Points

- The active comparator, new user study design limited baseline covariate differences between sulfonylurea and metformin monotherapy initiators with type 2 diabetes mellitus prior to PS implementation.
- We found little indication for differential channeling between sulfonylureas and metformin in patients without CVD and with CVD.
- Balance of covariates across treatments was achieved with both overall and subgroup-specific PSs using matching and stratification in three different cohorts.

Figure 1.

Figure 2.

Figure 3.

TABLE 1
Demographic Characteristics of New Users of Metformin and Sulfonylureas in Younger Insured and Two Older Medicare Cohorts

	Younger Cohort						Humana Medicare Advantage						Medicare Fee for Service					
	MET only			SU only			MET only			SU only			MET only			SU only		
	n	%		n	%		n	%		n	%		n	%		n	%	
Total	173080	100.0		53187	100.0		63669	100.0		43501	100.0		25311	100		13652	100	
Age: Younger cohort(Older Cohorts																		
<40	22865	13.2		4679	8.8													
40-49 65-69	48111	27.8		13050	24.5		22373	35.1		9361	21.5		7837	30.96		2884	21.1	
50-59 70-79	73579	42.5		24020	45.2		31038	48.7		19832	45.6		12076	47.71		5555	40.7	
60-64 80+	28525	16.5		11438	21.5		10258	16.1		14308	32.9		5398	21.33		5213	38.2	
Gender																		
Female	86057	49.7		21021	39.5		28436	44.7		21723	49.9		15253	60.26		7965	58.3	
Male	87023	50.3		32166	60.5		35233	55.3		21778	50.1		10058	39.74		5688	41.7	
Prior cardiovascular disease	21901	12.7		9425	17.7		21263	33.4		19339	44.5		12905	51.0		8850	64.8	
Other characteristics																		
Alcoholism	173	0.1		118	0.2		222	0.3		332	0.8		61	0.24		43	0.31	
Amputation	107	0.1		117	0.2		9885	15.5		10228	23.5		n/a			n/a		
Carotid revascularization	141	0.1		86	0.2		2292	3.6		2094	4.8		64	0.25		54	0.4	
Surg revascular'n/coronary bypass graft	952	0.5		683	1.3		24644	38.7		23676	54.4		95	0.38		123	0.9	
Implanted pacemaker/cardioverter defibril.	11005	6.4		4194	7.9		12973	20.4		10491	24.1		5938	23.46		4795	35.1	
Blind/macular edema/cataract/retinopathy	13030	7.5		4900	9.2		8271	13.0		7443	17.1		9767	38.59		5150	37.7	
Bone fractures	6438	3.7		1801	3.4		3123	4.9		4692	10.8		1486	5.87		1097	8.0	
Cancer Non Skin	9793	5.7		3244	6.1		13099	20.6		9114	21.0		4242	16.76		2567	18.8	
Chronic kidney disease/renal failure	1898	1.1		2456	4.6		11850	18.6		13587	31.2		1231	4.86		2608	19.1	
Chronic lung diseases	22026	12.7		6540	12.3		4937	7.8		4627	10.6		6876	27.17		4738	34.7	
Chronic neurological cond.	256	0.1		103	0.2		53946	84.7		37894	87.1		2960	11.69		1799	13.2	
Congestive heart failure	2829	1.6		1886	3.5		1447	2.3		1250	2.9		1834	7.25		2070	15.2	
Connective tissue diseases	24498	14.2		6923	13.0		344	0.5		292	0.7		7397	29.22		4114	30.1	
Depression	21893	12.6		4679	8.8		2742	4.3		1933	4.4		3371	13.32		2064	15.1	

	Younger Cohort						Humana Medicare Advantage						Medicare Fee for Service					
	MET only			SU only			MET only			SU only			MET only			SU only		
	n	%		n	%		n	%		n	%		n	%		n	%	
Diabetic foot ulcers	12136	7.0		3673	6.9		16371	25.7		15433	35.5		1674	6.61		3096	22.7	
Hypertension	108849	62.9		35054	65.9		55350	84.7		36395	86.7		22259	87.94		12280	89.9	
Hypoglycemia	8048	4.6		1746	3.3		141	0.2		161	0.4		767	3.03		415	3.0	
Inflammatory gastrointestinal disease	1211	0.7		463	0.9		9189	14.4		9085	20.9		181	0.72		156	1.1	
Liver disease	13873	8.0		4149	7.8		1236	1.9		1441	3.3		1578	6.23		991	7.3	
Myocardial infarction	18939	10.9		8228	15.5		5985	9.4		11135	25.6		8416	33.25		6131	44.9	
Neuropathies	13813	8.0		4262	8.0		2402	3.8		2896	6.7		5329	21.05		3564	26.1	
Pancreatitis	1097	0.6		525	1.0		2662	4.2		2611	6.0		237	0.94		164	1.2	
PVD/clauidication	5057	2.9		2367	4.5		14449	22.7		9338	21.5		4084	16.14		3105	22.7	
Proteinuria	3015	1.7		1426	2.7		9742	15.3		7847	18.0		401	1.58		369	2.7	
Stroke/transient ischemic attack	2477	1.4		1049	2.0		22373	35.1		9361	21.5		1691	6.68		1368	10.0	
Upper respiratory infection	70957	41.0		18345	34.5		31038	48.7		19832	45.6		6956	27.48		3801	27.8	
Urinary Tract Infection	21233	12.3		5498	10.3		10258	16.1		14308	32.9		5286	20.88		3665	26.8	