**SUPPLEMENTARY MATERIAL**



Figure S1: Chronological change in concomitant diuretic use in SGLT2i new users

Footnotes:

P-dataset: From a pharmacy claims database using data from over 800 pharmacies nation-wide which provided a coverage of approximately 2% of all outpatient prescriptions.

H-dataset: From a hospital-based administrative database constructed from data for inpatients and outpatients from 287 Diagnosis Procedure Combination (DPC) hospitals.

I-dataset: From an insurance claims database containing medical and prescription claims of 3.8 million employees and their dependents who were mostly aged ≤65 years.









Figure S2: Distribution of number of pre-index and concomitant antidiabetics in SGLT2i new users identified in in the full cohort and in the sensitivity cohort

(a) Pre-index antidiabetics in the P-dataset

(b) Pre-index antidiabetics in the H-dataset

(c) Concomitant antidiabetics in the P-dataset

(d) Concomitant antidiabetics in H-dataset.

Footnotes:

P-dataset: From a pharmacy claims database using data from over 800 pharmacies nation-wide which provided a coverage of approximately 2% of all outpatient prescriptions.

H-dataset: From a hospital-based administrative database constructed from data for inpatients and outpatients from 287 Diagnosis Procedure Combination (DPC) hospitals.

The insurance claims database included enrollment information, allowing for the requirement of continuous enrollment during the pre-index period. This information, however, was not available for the hospital-based database and pharmacy claims database. For example, approximately 40% of patients in the I-dataset had no pre-index OAD versus approximately 60% of patients in the H-dataset and P-dataset. Considering the younger age distribution in insurance claims database, this is unlikely to be true. In H-dataset and P-dataset, approximately 25% of patients had no record at all during the pre-index period. These patients might have either been 1) true new users without having any reason to come to hospital or pharmacy or 2) visited other hospital or pharmacy during this period of time. The case #1 should have been included as new users whereas case #2 should not have been included since they were actually the prevalent users. However, these were not distinguishable in this study setting. Therefore, we conducted sensitivity analyses for the hospital data and pharmacy claims data by limiting the population to only those who had any record during the pre-index period.

By excluding patients without having any record during the pre-index period, the number of new users reduced from 14,861 to 14,156 (-5%, P-dataset) and 27,039 to 26,594 (-2%, H-dataset), respectively. Patient demographics and baseline characteristics for the full study cohort were described elsewhere [30]. Demographic factors in the sensitivity cohort were similar to the full cohort. Comorbidities, complications, HbA1c levels, body weight, BMI, and eGFR values were also similar among the full cohort and sensitivity cohort.

Table S1: Patient demographics and baseline characteristics in the full cohort and sensitivity cohort

|  |  |  |
| --- | --- | --- |
|  | **P-dataset** | **H-dataset** |
|  | Full cohort | Sensitivity cohort | **Full cohort** | Sensitivity cohort |
|  |  |  |  |  |
| N | 14,861 | 14,156 | 27,039 | 26,594 |
| Age, Mean (SD), years | 58.7 (12.6) | 59.0 (12.6) | 57.7 (12.6) | 57.8 (12.6) |
| Female, n (%) | 5,649 (38.0) | 5,389 (38.1) | 10,200 (37.7) | 10,045 (37.8) |
| No. of beds at prescription site, n (%) |  |  |  |  |
| 0-19 | 6786 (45.7) | 6428 (45.4) | 0 | 0 |
| 20-199 | 3755 (25.3) | 3629 (25.6) | 3020 (11.2) | 2950 (11.1) |
| 200-399 | 1952 (13.1) | 1858 (13.1) | NA | NA |
| 200-499 | NA | NA | 16248 (60.1) | 16007 (60.2) |
| 400+ | 2368 (15.9) | 2241 (15.8) | NA | NA |
| 500+ | NA | NA | 7771 (28.7) | 7637 (28.7) |
| Specialty, n (%) |  |  |  |  |
| General Internal Medicine | 7559 (50.9) | 7176 (50.7) | 15617 (57.8) | 15373 (57.8) |
| Diabetologist | 721 (4.9) | 674 (4.8) | 2952 (10.9) | 2905 (10.9) |
| Endocrinology and metabolism | NA | NA | 3109 (11.5) | 3045 (11.4) |
| Cardiology | 631 (4.2) | 608 (4.3) | 2768 (10.2) | 2727 (10.3) |
| Others | 1907 (12.8) | 1824 (12.9) | 2253 (8.3) | 2205 (8.3) |
| Unknown | 4032 (27.1) | 3863 (27.3) | 340 (1.3) | 339 (1.3) |
| Multiple | 11 (0.1) | 11 (0.1) | 0 (0.0) | 0 (0.0) |
| Comorbidity, n (%) |  |  |  |  |
| Hypertension | 7631 (51.3) | 7464 (52.7) | 18259 (67.5) | 18048 (67.9) |
| Hypercholesterolemia | 7900 (53.2) | 7735 (54.6) | 19873 (73.5) | 19664 (73.9) |
| Elixhauser comorbidity score | NA | NA |  |  |
| Mean (SD) |  |  | 5.8 (6.9) | 5.8 (6.9) |
| Median (IQR) |  |  | 3 (0-11) | 3 (0-11) |
| Complications (DCSI), n (%) | NA | NA |  |  |
| Retinopathy |  |  | 5679 (21.0) | 5640 (21.2) |
| Nephropathy |  |  | 5701 (21.1) | 5647 (21.2) |
| Neuropathy |  |  | 3445 (12.7) | 3415 (12.8) |
| Cerebrovascular disease |  |  | 3161 (11.7) | 3135 (11.8) |
| Cardiovascular disease |  |  | 9440 (34.9) | 9359 (35.2) |
| Peripheral vascular disease |  |  | 701 (2.6) | 697 (2.6) |
| Metabolic disease |  |  | 138 (0.5) | 136 (0.5) |
| DCSI score | NA | NA |  |  |
| mean (SD) |  |  | 1.4 (1.5) | 1.5 (1.5) |
| median (IQR) |  |  | 1 (0-2) | 1 (0-2) |
| HbA1c, % | NA | NA |  |  |
| n |  |  | 3120 | 3095 |
| mean (SD) |  |  | 8.4 (1.4) | 8.4 (1.4) |
| Body weight, kg | NA | NA |  |  |
| n |  |  | 336 | 330 |
| mean (SD) |  |  | 74.4 (19.5) | 74.5 (19.5) |
| BMI, kg/m2 | NA | NA |  |  |
| n |  |  | 336 | 330 |
| mean (SD) |  |  | 27.6 (6.1) | 27.7 (6.1) |
| eGFR, mL/min/1.73 m2 | NA | NA |  |  |
| n |  |  | 2556 | 2531 |
| mean (SD) |  |  | 77.8 (24.0) | 77.8 (24.1) |

P-dataset: From a pharmacy claims database using data from over 800 pharmacies nation-wide which provided a coverage of approximately 2% of all outpatient prescriptions.

H-dataset: From a hospital-based administrative database constructed from data for inpatients and outpatients from 287 Diagnosis Procedure Combination (DPC) hospitals.

Abbreviations: BMI=body mass index; DCSI=diabetes complication severity index; eGFR=estimated glomerular filtration rate; HbA1c=glycated hemoglobin; IQR=interquartile range; NA=not applicable; SD=standard deviation.