PASS Study Report, Final (120-Month) Report	
Active substance	Dapagliflozin
Study Code	D1690R00007
Version number	1.0
Date	10 September 2025

Post-authorisation Observational Study, Final (120-Month) Report: Comparison of the Risk of Cancer Between Patients With Type 2 Diabetes Exposed to Dapagliflozin and Those Exposed to Other Antidiabetic Treatments

HMA-EMA Catalogue of RWD Studies: Study ID 49352
(https://catalogues.ema.europa.eu/node/3411/administrative-details)

Approved by:

[redacted]

Lia Gutiérrez, MPH — Date
Principal Investigator, RTI Health Solutions

[redacted]

[redacted]

Date

[redacted]

Date

PASS INFORMATION

Title	Post-authorisation Observational Study, Final (120-Month) Report: Comparison of the Risk of Cancer Between Patients With Type 2 Diabetes Exposed to Dapagliflozin and Those Exposed to Other Antidiabetic Treatments
Version identifier of the study report	1.0
Date of last version of the study report	10 September 2025
EU PAS Register number **HMA-EMA Catalogue of RWD Studies ID number**	EUPAS12116 Study ID 49352
Active substance	A10BK01 (dapagliflozin) A10BD15 (dapagliflozin + metformin) A10BD21 (saxagliptin + dapagliflozin)
Medicinal product	Dapagliflozin (Edistride, Forxiga [UK, EU]; Farxiga [US]) Dapagliflozin + metformin (Ebymect, Xigduo) Saxagliptin + dapagliflozin (Qtern)
Product reference	Forxiga EU/1/12/795/001-011 Edistride: EU/1/15/1052/001-010 Xigduo: EU/1/13/900/001-012 Ebymect: EU/1/15/1051/001-012 Qtern: EU/1/16/1108/001-004
Procedure number	Forxiga EMEA/H/C/002322 Edistride: EMEA/H/C/004161 Xigduo: EMEA/H/C/002672 Ebymect: EMEA/H/C/004162 Qtern: EMEA/H/C/004057

Marketing authorisation holder(s)	AstraZeneca AB
Joint PASS	No
Research question and objectives	This study is a multinational cohort study to estimate the risk of breast, bladder, and a composite of selected sex-specific cancers in patients with type 2 diabetes mellitus (T2DM) who are new users of dapagliflozin compared with those who are new users of antidiabetic drugs (ADs) other than sodium-glucose cotransporter 2 (SGLT2) inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy. The main study objectives are to compare, by insulin use at the index date, the following outcomes: • The incidence of breast cancer, by insulin use at the index date, amongst females with T2DM who are new users of dapagliflozin and females who are new users of other eligible comparator ADs • The overall and sex-specific incidence of bladder cancer, by insulin use and by pioglitazone use at the index date, amongst patients with T2DM who are new users of dapagliflozin and those who are new users of other eligible comparator ADs
Country (-ies) of study	United Kingdom (UK) The Netherlands United States of America (US)
Authors	Lia Gutiérrez, MPH [redacted] RTI Health Solutions [redacted] Heather Danysh, PhD [redacted] RTI Health Solutions [redacted] Anam Khan, PhD [redacted]

MARKETING AUTHORISATION HOLDER

Marketing authorisation holder(s)	AstraZeneca AB Södertälje SE-151 85 Sweden [redacted]
MAH contact person	[redacted]

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

1. ABSTRACT

Title

Post-authorisation Observational Study, Final (120-Month) Report: Comparison of the Risk of Cancer Between Patients With Type 2 Diabetes Exposed to Dapagliflozin and Those Exposed to Other Antidiabetic Treatments

Main authors

Lia Gutiérrez, MPH, RTI Health Solutions; Heather Danysh, PhD, RTI Health Solutions; Anam Khan, PhD, RTI Health Solutions

Date

10 September 2025

Keywords

Bladder cancer, cancer, dapagliflozin, female breast cancer, type 2 diabetes mellitus

Rationale and background

Dapagliflozin is a sodium-glucose cotransporter 2 (SGLT2) inhibitor approved by the European Medicines Agency (EMA) in 2012 for the treatment of type 2 diabetes mellitus (T2DM), with subsequent approvals for the treatment of symptomatic chronic heart failure with reduced ejection fraction in 2020 (extended to include heart failure with mildly reduced or preserved left ventricular ejection fraction in 2023) and chronic kidney disease in 2021. Due to numerical imbalances in the incidence of breast cancer and bladder cancer in the dapagliflozin premarketing clinical trials, the EMA required this post-authorisation safety study (PASS) at the time of the initial approval of dapagliflozin to further evaluate the risk of breast and bladder cancer in real-world settings.

Research question and objectives

The primary objective was to compare the incidence of female invasive breast cancer and the sex-combined and sex-specific incidence of in situ and invasive bladder cancer—(1) stratified by insulin use at the index date for all cancer outcomes and (2) stratified by pioglitazone use at the index date for the sex-combined bladder cancer outcome—amongst patients with T2DM who were new users of dapagliflozin compared with those who were new users of other eligible comparator antidiabetic drugs (ADs). Analyses were also conducted on the overall cohorts (ie, without stratification by insulin or pioglitazone use at the index date) due to the small number of insulin users and of pioglitazone users observed early on in the study (analyses on the overall cohorts were initiated during the third interim [72-month] analysis).

Secondary objectives were to (1) compare baseline characteristics of both exposure groups, (2) assess differential medical surveillance across exposure groups by comparing health care

resource utilisation (HCRU) during follow-up, and (3) compare the incidence of a composite of selected sex-specific cancer outcomes across the exposure groups.

Study design

This non-interventional cohort study used data from four longitudinal, population-based data sources. The study cohorts were defined by new use of an eligible study medication (ie, dapagliflozin or comparator AD). The study used a new-user design and defined the index date as the date of first prescription or dispensing of the study medication.

Patient characteristics and potential confounding variables recorded before or on the index date (ie, demographics, lifestyle, medical conditions, medications, and HCRU) were used to estimate propensity scores. Propensity score trimming and stratification were used to adjust for multiple possible confounding variables by making both exposure groups more comparable. Descriptive analyses were conducted for all cancer outcomes, before and after propensity score trimming in the overall cohorts and in the cohorts stratified by insulin use at the index date. For the sex-combined bladder cancer outcome, descriptive analyses were also stratified by pioglitazone use at the index date.

All incidence and comparative analyses were conducted on each cohort (the overall cohorts; insulin use–stratified cohorts; and, for sex-combined bladder cancer, the pioglitazone use–stratified cohorts) after propensity score trimming (ie, the propensity score–trimmed analysis samples). Crude and propensity score–adjusted incidence rates for each cancer outcome were calculated separately for each exposure group (using the respective dapagliflozin-exposed and comparator AD–exposed person-time). Incidence rate ratios (IRRs) were estimated comparing the incidence of each cancer outcome in new users of dapagliflozin with that of new users of comparator AD. Propensity score–adjusted IRRs for each cancer outcome were calculated using Mantel-Haenszel methods from propensity score–stratified IRRs for each of the overall and insulin use–stratified cohorts. Similarly, Mantel-Haenszel methods and the propensity score–stratified IRRs across all data sources were used to estimate a pooled adjusted IRR for each primary cancer outcome, both for the overall cohort and the insulin use–stratified cohorts. Sensitivity analyses were conducted in the primary cancer outcome cohorts to assess the impact of various study design elements on the propensity score–adjusted IRR estimates.

Setting

The study used data from the Clinical Practice Research Datalink (CPRD), specifically the CPRD General Practitioner Online Database (GOLD), in the United Kingdom; the PHARMO Data Network (PHARMO) in the Netherlands; and Carelon Research's Healthcare Integrated Research Database (HIRD®)[1] and the Medicare fee-for-service claims database in the United

[1] In March 2023, HealthCore, Inc. was rebranded to Carelon Research, and the "HIRD" was renamed Healthcare Integrated Research Database.

States. The study period began the day after regulatory approval of dapagliflozin in each country (ranging from November 2012 to January 2014) and ended at the last date of observation available at the time of data extraction in each data source (ranging from March 2021 to September 2023).

Subjects and study size, including cohort attrition

Depending on the data source, eligible individuals were required to be registered in the primary care practice, health data network, or health insurance plan for at least 180 days before the index date and, before or on the index date, to have no recorded diagnosis of type 1 diabetes mellitus, no prescription for dapagliflozin or other SGLT2 inhibitor, and no recorded diagnosis of an invasive cancer. After applying eligibility criteria and propensity score trimming for each cancer outcome cohort, the number of new users included in the propensity score–trimmed analysis sample (percentage of the full [untrimmed] sample that was retained after propensity score trimming) in the dapagliflozin and comparator AD exposure groups, respectively, were as follows:

- Female breast cancer cohort:
  - CPRD, 5,713 (75%) and 23,877 (69%)
  - PHARMO, 1,382 (61%) and 14,075 (77%)
  - The HIRD, 25,988 (84%) and 272,904 (85%)
  - Medicare, 31,656 (79%) and 503,227 (81%)
- Sex-combined bladder cancer cohort:
  - CPRD, 14,050 (74%) and 56,641 (68%)
  - PHARMO, 3,298 (50%) and 30,402 (75%)
  - The HIRD, 63,528 (84%) and 530,047 (89%)
  - Medicare, 62,856 (78%) and 884,302 (79%)

Variables and data sources

Occurrences of invasive breast cancer, in situ or invasive bladder cancer, and sex-specific invasive composite cancer diagnoses were identified using diagnosis codes and electronic algorithms tailored to each data source. Outcome validation was performed for a sample of breast cancer and bladder cancer cases in CPRD, the HIRD, and Medicare. In PHARMO, outcome validation using source medical records was not conducted because cancer cases were identified using the Netherlands Cancer Registry, the gold standard for identifying cancer diagnoses.

The primary exposure was newly initiated dapagliflozin. Comparator AD exposure was defined as new initiation of an eligible AD, including monotherapy or combination therapy

with alpha-glucosidase inhibitors, thiazolidinediones, dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide-1 (GLP-1) receptor agonists, meglitinides, and other blood glucose–lowering drugs (pramlintide, bromocriptine). SGLT2 inhibitors and monotherapy with insulin, metformin, or a sulfonylurea were not eligible comparators; however, combination therapy of an eligible AD with insulin, metformin, or a sulfonylurea, or combination therapy of free-dose metformin and sulfonylurea (with no other eligible AD) were included as eligible comparators.

Potential confounding variables were assessed at baseline and included medical conditions related to diabetes severity, cancer risk factors, other medical comorbidities, selected medications, lifestyle factors as available in each data source, and HCRU.

Results

Primary objective: compare the incidence of female breast cancer and bladder cancer across exposure groups

Female invasive breast cancer

The propensity score–adjusted IRR (95% confidence interval [CI]) for female breast cancer in the overall cohort, comparing dapagliflozin new users with comparator AD new users, was as follows: CPRD, 1.05 (0.75-1.46); PHARMO, 0.87 (0.46-1.63); the HIRD, 1.06 (0.90-1.24); and Medicare, 0.98 (0.86-1.11). The propensity score–adjusted IRR (95% CI) amongst insulin users ranged from 1.06 (0.80-1.41) in Medicare to 1.42 (0.95-2.10) in the HIRD; in PHARMO, the propensity score–adjusted IRR amongst insulin users was not estimable due to a single female breast cancer case (n = 1) in the dapagliflozin group. For each data source, most new users were insulin non-users and, therefore, the propensity score–adjusted IRR amongst insulin non-users was similar to the estimate in the overall cohort. The pooled adjusted IRR (95% CI) for female breast cancer across all four data sources was 1.01 (0.92-1.11) in the overall cohort, 1.17 (0.94-1.46) amongst insulin users (pooled across CPRD, the HIRD, and Medicare), and 0.98 (0.88-1.09) amongst insulin non-users.

Sensitivity analyses to assess the impact of various study design elements on the propensity score–adjusted IRR estimates for female breast cancer yielded results similar to those observed in the main analysis.

In situ and invasive bladder cancer

For sex-combined bladder cancer, the propensity score–adjusted IRR (95% CI) for the overall cohort, comparing dapagliflozin new users with comparator AD new users, was as follows: CPRD, 0.74 (0.45-1.21); PHARMO, 0.90 (0.48-1.68); the HIRD, 0.82 (0.55-1.24); and Medicare, 0.74 (0.59-0.94). The propensity score–adjusted IRR (95% CI) amongst insulin users ranged from 0.36 (0.16-0.77) in Medicare to 1.01 (0.33-3.09) in CPRD; in PHARMO, the propensity score–adjusted IRR amongst insulin users was not estimable due to few bladder cancer cases (n = 2) observed in the dapagliflozin group. For each data source, most new users

were insulin non-users and, therefore, the propensity score–adjusted IRR amongst insulin non-users was similar to the estimate in the overall cohort. The pooled adjusted IRR (95% CI) for sex-combined bladder cancer across all four data sources was 0.77 (0.64-0.92) in the overall cohort, 0.50 (0.28-0.87) amongst insulin users (pooled across CPRD, the HIRD, and Medicare), and 0.81 (0.66-0.98) amongst insulin non-users.

Amongst females, the propensity score–adjusted IRR (95% CI) for bladder cancer in the overall cohort was 0.64 (0.18-2.33) in CPRD and 0.72 (0.41-1.26) in Medicare; propensity score–adjusted IRRs were not estimable amongst females in PHARMO and the HIRD due to few bladder cancer cases. The propensity score–adjusted IRR (95% CI) for bladder cancer amongst female insulin users was estimable only in CPRD (2.00; 0.16-25.73), and amongst female insulin non-users was estimable only in Medicare (0.91; 0.52-1.60). The overall pooled adjusted IRR (95% CI) for female bladder cancer across CPRD and Medicare was 0.71 (0.42-1.18); the pooled adjusted IRRs were not estimable for the female insulin use–stratified cohorts.

Amongst males, the propensity score–adjusted IRR (95% CI) for bladder cancer in the overall cohorts was as follows: CPRD, 0.91 (0.54-1.53); PHARMO, 1.06 (0.55-2.04); the HIRD, 1.00 (0.65-1.54); and Medicare, 0.76 (95% CI, 0.59-0.98). The propensity score–adjusted IRR (95% CI) for bladder cancer amongst male insulin users was estimable only in CPRD (0.85; 0.27-2.75) and Medicare (0.44; 0.20-0.96). For each data source, most new users were insulin non-users and, therefore, the propensity score–adjusted IRR amongst insulin non-users was similar to the estimate in the overall cohort. The pooled adjusted IRR (95% CI) for male bladder cancer across all four data sources was 0.84 (0.69-1.02) in the overall cohort, 0.51 (0.27-0.98) amongst insulin users (pooled across CPRD and Medicare), and 0.86 (0.70-1.06) amongst insulin non-users.

Amongst those with pioglitazone use at the index date, the propensity score–adjusted IRR (95% CI) for sex-combined bladder cancer was 1.41 (0.35-5.60) in CPRD and 0.82 (0.41-1.67) in Medicare; the propensity score–adjusted IRR was not estimable amongst pioglitazone users in PHARMO or the HIRD due to few bladder cancer cases.

Sensitivity analyses to assess the impact of various study design elements on the propensity score–adjusted IRR estimates for bladder cancer yielded results similar to those observed in the main analysis.

Secondary objectives:

Baseline characteristics

Before propensity score trimming, dapagliflozin new users were younger (CPRD); were more likely to be male (PHARMO and the HIRD); had more frequent indicators of diabetes severity, including retinopathy and peripheral vascular disease (the HIRD and Medicare) and

coronary heart disease (PHARMO and the HIRD); had more frequent heart failure and chronic kidney disease (PHARMO and the HIRD); had greater concomitant use of insulin (CPRD) and previous AD classes (in all data sources); had a longer time since the first recorded diagnosis of T2DM (CPRD and PHARMO); and were more frequently overweight (CPRD, the only data source where this variable was evaluable in the full samples). After propensity score trimming and stratification, good balance was observed between exposure groups in the overall cohort and in the insulin use–stratified cohorts for all data sources, as indicated by absolute standardised difference values equal to or less than 0.20 for most variables within most propensity score strata.

Healthcare resource utilisation during follow-up

Across all data sources, there were no consistent differences in the means and rates of HCRU between dapagliflozin new users and comparator AD new users observed within each year of follow-up. In the time-to-event sensitivity analysis for all data sources, the cumulative incidence of female breast cancer and of bladder cancer in both exposure groups were similar following the initiation of the index therapy and did not indicate evidence of surveillance bias for either female breast cancer or bladder cancer.

Sex-specific composite cancers

For the female composite cancer outcome, the propensity score–adjusted IRR (95% CI) for the overall cohort, comparing dapagliflozin new users with comparator AD new users, was as follows: CPRD, 0.99 (0.75-1.30); PHARMO, 0.92 (0.59-1.41); the HIRD, 1.11 (0.96-1.28); and Medicare, 1.01 (0.89-1.15). Amongst insulin users, the propensity score–adjusted IRR (95% CI) for the female composite cancer outcome ranged from 0.85 (0.45-1.60) in CPRD to 1.59 (0.59-4.28) in PHARMO; amongst insulin non-users, the estimates ranged from 0.85 (0.52-1.41) in PHARMO to 1.11 (0.95-1.30) in the HIRD.

For the male composite cancer outcome, the propensity score–adjusted IRR (95% CI) for the overall cohort was as follows: CPRD, 0.89 (0.74-1.06); PHARMO, 0.74 (0.54-1.01); the HIRD, 0.88 (0.79-0.99); and Medicare, 0.98 (0.90-1.06). Amongst insulin users, the propensity score–adjusted IRR (95% CI) estimates for the male composite cancer outcome ranged from 0.70 (0.42-1.17) in CPRD to 1.08 (0.89-1.30) in Medicare; in PHARMO, the propensity score–adjusted IRR was not estimable amongst insulin users due to few male composite cancer outcome events. Amongst insulin non-users, the propensity score–adjusted IRR (95% CI) estimates for the male composite cancer outcome ranged from 0.66 (0.46-0.95) in PHARMO to 0.96 (0.88-1.05) in Medicare.

Discussion

The data source–specific and pooled adjusted IRRs for the overall cohorts (not stratified by insulin use at the index date) evaluating the association of dapagliflozin exposure with the incidence of cancer did not suggest an increased risk for any of the study cancer outcomes in

dapagliflozin new users compared with comparator AD new users; however, the small number of cancer cases in some of the overall cohorts (eg, the sex-specific bladder cancer cohorts) resulted in imprecise IRR estimates with wide 95% CIs. Due to the small number of dapagliflozin new users and small number of cancer cases in several data sources, the analyses amongst insulin users (and amongst pioglitazone users for the sex-combined bladder cancer outcome) do not allow for informative inferential conclusions.

Marketing Authorisation Holder

AstraZeneca AB

Names and affiliations of principal investigators

Lia Gutiérrez, MPH, RTI Health Solutions
Heather Danysh, PhD, MHS, RTI Health Solutions
Kimberly Daniels, PhD, MS, Carelon Research
Jetty Overbeek, PhD, PHARMO

2. LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
A1C	glycated haemoglobin
AD	antidiabetic drug
ATC	Anatomical Therapeutic Chemical (classification system)
BMI	body mass index (calculated as kg/m^2)
BRCA/*BRCA1*	breast cancer gene/breast cancer gene 1
CDC	Centers for Disease Control and Prevention
CI	confidence interval
CKD	chronic kidney disease
CMS	Centers for Medicare and Medicaid Services
COVID-19	coronavirus disease 2019
CPRD	Clinical Practice Research Datalink (UK)
CPT	Current Procedural Terminology
DC	District of Columbia
DCSI	Diabetes Complications Severity Index
DPP-4	dipeptidyl peptidase-4
EMA	European Medicines Agency
ER	extended release (medication)
ER-[positive or negative]	oestrogen receptor–[positive or negative]
ERM	oestrogen receptor modulator
EU	European Union
EU PAS Register	European Union Electronic Register of Post-Authorisation Studies
FDA	Food and Drug Administration (US)
GDPR	The General Data Protection Regulation (EU)
GIP	glucose-dependent insulinotropic polypeptide
GLP-1	glucagon-like peptide-1
GOLD	General Practitioner Online Database (of CPRD)
GP	general practitioner or general practice
GPI	Generic Product Identifier
HbA1c	glycated haemoglobin
HCPCS	Healthcare Common Procedure Coding System
HCRU	health care resource utilisation
HES	Hospital Episode Statistics
HIRD®	Healthcare Integrated Research Database

Abbreviation or special term	Explanation
HIV	human immunodeficiency virus
HMA-EMA Catalogue	Heads of Medicines Agencies–European Medicines Agency Catalogue of RWD Studies (replaced the EU PAS Register in February 2024)
HR	hazard ratio
ICD-9	International Classification of Diseases, 9th Revision
ICD-9-CM	International Classification of Diseases, 9th Revision, Clinical Modification
ICD-10	International Classification of Diseases, 10th Revision
ICD-10-CM	International Classification of Diseases, 10th Revision, Clinical Modification
ICD-O-3	International Classification of Diseases for Oncology, Third Edition
ICPC	International Classification of Primary Care
ID	identification
IPTW	inverse probability of treatment weights
IQR	interquartile range
IRB	institutional review board
IRR	incidence rate ratio
ISAC	Independent Scientific Advisory Committee
ISPE	International Society for Pharmacoepidemiology
KM	Kaplan-Meier
LVEF	left ventricular ejection fraction
NA	not applicable
NCR	Netherlands Cancer Registry
NDC	National Drug Code
NE	not estimable
NEN 7150	a Dutch standard for information security in the healthcare sector that provides a framework for managing and protecting patient data
NHL	non-Hodgkin lymphoma
NMSC	non-melanoma skin cancer
NOS	not otherwise specified
NR	not reportable
OR	odds ratio
PASS	post-authorisation safety study
PHARMO	PHARMO Data Network of the PHARMO Institute, part of Lumanity (the Netherlands)
PPV	positive predictive value
PS	propensity score
Q*n*	quintile

Abbreviation or special term	Explanation
Q*n* yyyy	quarter of the calendar year
RR	risk ratio
RTI	RTI International, of which RTI Health Solutions is a unit
RWD	real-world data
SAP	statistical analysis plan
SD	standard deviation
SEER	Surveillance, Epidemiology, and End Results (programme)
SES	socioeconomic status
SGLT2	sodium-glucose cotransporter 2
SGLT2i	sodium-glucose cotransporter 2 inhibitor
SR	sustained release
StDiff	absolute standardised difference
STIZON	Foundation for Information Provision for Care and Research [Stichting Informatievoorziening voor Zorg en Onderzoek] (the Netherlands)
T1DM	type 1 diabetes mellitus
T2DM	type 2 diabetes mellitus
UK	United Kingdom
US	United States of America
WHO	World Health Organization
XR	extended release

Glossary of Terms

Full sample (or "untrimmed sample"): for a specific cohort, the sample of patients after the eligibility criteria have been applied and before propensity score trimming has been implemented.

Index date: date of first prescription or dispensing of an eligible study drug during the study period.

Index study drug: dapagliflozin (single-entity dapagliflozin or a fixed-dose combination) or any of the antidiabetic drugs listed in Appendix B, Table 28, that were initiated during the study period after meeting all eligibility criteria, and for which no record of prescription or dispensing of that particular drug was found before the first prescription or dispensing that occurred during the study period, based on all available data, with a minimum of 180 days.

Insulin users cohort (or "insulin users"): patients who used insulin at the index date.

Insulin non-users cohort (or "insulin non-users"): patients who did not use insulin at the index date.

Overall cohort: cohort not stratified by insulin use at the index date

Pioglitazone users cohort (or "pioglitazone users"): patients who used pioglitazone at the index date.

Pioglitazone non-users cohort (or "pioglitazone non-users"): patients who did not use pioglitazone at the index date.

Propensity score–trimmed analysis sample (or "propensity score−trimmed sample"): for a specific cohort, the sample of patients after propensity score trimming has been applied to the full sample; the sample used for all incidence and comparative analyses.

Sex-combined bladder cancer cohort: bladder cancer outcome cohort that includes both females and males.

3. INVESTIGATORS

Collaborating Primary Institutions and Principal Investigators

RTI Health Solutions

Lia Gutiérrez, MPH
Senior Director, Epidemiology
[redacted]

Heather Danysh, PhD, MHS
Director, Epidemiology
[redacted]

Carelon Research, Formerly HealthCore, Inc.

Kimberly Daniels, PhD, MS
Research Scientist, Safety & Epidemiology
[redacted]

PHARMO Institute, part of Lumanity

Jetty Overbeek, PhD
Senior Research Manager, Head of PHARMO Research
[redacted]

AstraZeneca (Study Sponsor)

[redacted]
Associate Director, Safety Epidemiology
[redacted]

4. OTHER RESPONSIBLE PARTIES

The sponsor, AstraZeneca, was responsible for providing reasonable resources for study implementation and to assure study progress. AstraZeneca was also responsible for communicating with regulatory agencies about the study protocol, the progress of the study, and study findings.

The study investigators at RTI Health Solutions; PHARMO Institute, part of Lumanity (formerly PHARMO Institute for Drug Outcomes Research); and Carelon Research (formerly

HealthCore) shared responsibility with Bristol-Myers Squibb and AstraZeneca for the design of the study. The investigators at RTI Health Solutions were responsible for conducting the Clinical Practice Research Datalink and Medicare components in a manner that meets regulatory and methodologic standards, conducting analyses, and preparing scientific reports. The investigators at PHARMO and Carelon Research were responsible for implementing the PHARMO Data Network and the Healthcare Integrated Research Database (HIRD®) components, respectively, in a manner that meets regulatory standards, conducting analyses, providing results of these analyses to RTI Health Solutions, and reviewing scientific reports.

Other Study Team Members

RTI Health Solutions

Alicia Gilsenan, PhD, Vice President, Epidemiology
Becky MacKay, MPH, Research Epidemiologist
Anam Khan, PhD, Research Epidemiologist
Xabier García de Albeniz Martínez, MD, SCM, PhD, Director, Epidemiology
James Kaye MD, DrPH, Senior Director Epidemiology
Jaume Aguado, PhD, Director, Biostatistics
Jennifer Bartsch, BA, MStat, Senior Statistician
Brian Calingaert, MS, Director, Epidemiology Analysis
Diana Garbinsky, MS, Senior Statistician

PHARMO Institute

Brenda Baak, MSc, Research Manager
Louise Lemmens, PhD, Data Scientist
Ilona Verburg, PhD, Data Scientist
Fernie Penning-van Beest, PhD, Senior Research Quality Manager
Prof. Ron Herings, PhD, Scientific Director

Carelon Research, formerly HealthCore, Inc.

Daniel C. Beachler, PhD, MHS, Director, Safety & Epidemiology
Tracey Quimbo, MPH, Analyst, Biostatistics
Melissa Lamberti, MSW, Project Management
Kelsey Hall, MPH, Director, Program and Project Management
Stephan Lanes, PhD, MPH, Principal Scientist, Epidemiology

5. MILESTONES

Table 1 Milestones

Milestone	Planned date	Actual date	Comments
First ethics/IRB approval	Before data collection	20 November 2015 (RTI International IRB) 21 December 2015 (MHRA ISAC) 12 May 2016 (STIZON) 11 January 2018 (CMS privacy board) 19 March 2018 (New England IRB)	
Start of data collection	29 January 2016	05 January 2016	Date of extraction for PHARMO, first data source to extract data for first interim report
Registration in the HMA-EMA Catalogue of RWD Studies (formerly EU PAS Register)	None	18 January 2016	All studies in the EU PAS Register were migrated to the HMA-EMA Catalogue of RWD Studies in Q1 2024
Last ethics/IRB approval	None	25 April 2025 (RTI International IRB) 06 February 2025 (CMS privacy board) 18 February 2022 (New England IRB) 18 December 2017 (MHRA ISAC) 12 May 2016 (STIZON)	
Interim report 1 (24-month) data cut; descriptive	January 2016	08 February 2016 (CPRD) 03 February 2016 (HIRD) 05 January 2016 (PHARMO)	
Interim report 1 (24-month report) submitted to EMA	2016	26 December 2016	
Interim report 2 (48-month) data cut; comparative	January 2018	06 February 2018 (Medicare) 29 January 2018 (CPRD) 18 January 2018 (PHARMO) 08 January 2018 (HIRD)	
Interim report 2 (48-month report) submitted to EMA	2019	24 June 2019	

Milestone	Planned date	Actual date	Comments
Interim report 3 (72-month report) data cut; comparative	January 2020	27 July 2020 (PHARMO) 01 July 2020 (CPRD) 09 March 2020 (HIRD) 31 July 2019 (Medicare)[a]	Data extraction for this interim data cut was delayed to allow time for the statistical analysis plan to be amended based on observations made during the second interim analysis.
Interim report 3 (72-month report) submitted to EMA	2021	29 November 2021	Interim report 3 (72-month report) submitted to EMA
Interim report 4 (96-month) data cut; comparative	January 2022	24 October 2022 (CPRD) 14 June 2022 (Medicare) 01 February 2022 (HIRD) 14 January 2022 (PHARMO)	Interim report 4 (96-month) data cut; comparative
Interim report 4 (96-month report) submitted to EMA	2023	28 November 2023	Interim report 4 (96-month report) submitted to EMA
Final analysis data cut	January 2024	11 July 2024 (PHARMO) 24 January 2024 (HIRD) 05 December 2023 (Medicare) 24 October 2022 (CPRD)	The date of the final (120-month) analysis data cut for CPRD is the same as that for interim report 4 (96-month) analysis due to a delay in the availability of the updated linkage of CPRD GOLD and HES data at the time of the data extraction for the final (120-month) analysis; hence, the same CPRD GOLD and linked HES data used in the interim report 4 (96-month) analysis were used in the final (120-month) analysis.
End of data collection	30 January 2024	26 May 2025[b]	
Final report of study results	2025	01 July 2025	

[a] The July 2019 Medicare data cut was used because it included the same data that were available in January 2020, with the latest data available through 31 December 2017.

[b] Date the analytical data set required to perform the final statistical analyses for the primary objectives was available from all data sources.

CMS = Centers for Medicare and Medicaid Services; CPRD GOLD = General Practitioner Online Database of CPRD; EMA = European Medicines Agency; EU PAS Register = European Union Electronic Register of Post-Authorisation Studies; HES = Hospital Episode Statistics; HMA = Heads of Medicines Agencies; IRB = institutional review board; ISAC = Independent Scientific Advisory Committee; MHRA = Medicines Healthcare products Regulatory Agency (UK); Q*n* yyyy = quarter of the calendar year; RWD = real-world data.

6. RATIONALE AND BACKGROUND

6.1 Dapagliflozin Regulatory History

Dapagliflozin—Edistride, Forxiga (EU); Farxiga (US)—is a highly potent, selective, and reversible inhibitor of human renal SGLT2, the major transporter responsible for renal glucose reabsorption. Dapagliflozin received approval for the treatment of T2DM from the EMA on 12 November 2012 [52] and from the FDA on 08 January 2014 [47]. Dapagliflozin lowers plasma glucose by inhibiting renal reabsorption of glucose and by promoting its urinary excretion. The combination product of dapagliflozin + metformin (Xigduo, Xigduo XR, Ebymect) was approved in the EU in January 2014 [139] and in the US in October 2014 [140]. The combination product of saxagliptin + dapagliflozin (Qtern) was approved in the EU in July 2016 [105] and in the US in February 2017 [104]. The combination product of dapagliflozin, saxagliptin, + metformin was approved in the US (Qternmet XR) in May 2019 [106] and in the EU (Qtrilmet) in November 2019 [107] but has since been discontinued in the US [49], and authorisation in the EU has been withdrawn [40]. In May 2020, dapagliflozin was approved in the US for the treatment of adults with symptomatic chronic heart failure with reduced ejection fraction [7], and in the EU, this indication was approved in October 2020 [41]. In February 2023 in the EU, dapagliflozin approval was extended to also include the treatment of heart failure with mildly reduced or preserved left ventricular ejection fraction (LVEF > 40%) [43; 44]. In April 2021, dapagliflozin was approved in the US for the treatment of chronic kidney disease in adults [8], followed by an EU approval in June 2021 [42].

In 2013, the dapagliflozin clinical development programme reported numerical imbalances in the incidence of breast and bladder cancers between dapagliflozin-treated patients and controls. Pooled analysis of all phase 2b and phase 3 clinical studies was conducted to further evaluate the incidence of breast cancer (n = 12, 0.40 events per 100 patient-years in the dapagliflozin group; n = 3, 0.19 events per 100 patient-years in the control group; IRR, 2.47; 95% CI, 0.64-14.10) and of bladder cancer (n = 10, 0.15 events per 100 patient-years in the dapagliflozin group; n = 1, 0.03 events per 100 patient-years in the control group; IRR, 6.11; 95% CI, 0.83-272.02) diagnosed within 2 years of study start; the analysis yielded imprecise results. There was no imbalance in the number of all other malignancies [20]. The DECLARE trial—a randomised double-blind, placebo-controlled, phase 3 trial of dapagliflozin in patients with T2DM and established, or at risk of, atherosclerotic cardiovascular disease (defined as two or more risk factors for cardiovascular disease in addition to T2DM)—evaluated the risk of cardiovascular outcomes amongst 17,160 patients who were followed for a median of 4.2 years. As part of the evaluation of secondary outcomes, the HR for breast cancer was 1.02 with a wide CI (95% CI, 0.64-1.63); results for bladder cancer indicated a lower hazard in the dapagliflozin arm than in the placebo arm (HR, 0.57; 95% CI, 0.35-0.93) [136].

6.2 Epidemiology of Breast Cancer and Bladder Cancer Amongst Patients With T2DM

6.2.1 Breast Cancer

Diabetes and ADs have been previously evaluated as potential risk factors for breast cancer. Multiple studies have shown a higher risk of breast cancer amongst women with diabetes than amongst women without diabetes [16; 80; 83; 141; 144]; however, the risk varies by duration of T2DM; age; breast cancer molecular subtype; and, potentially, menopausal status. Amongst women with diabetes, studies have reported higher breast cancer risk in women within 10 years of their T2DM diagnosis [142] and in younger women (aged < 52 years) [119], and the risk is strongest for breast cancer negative for oestrogen, progesterone, and human epidermal growth factor receptors, known as triple-negative breast cancer [55; 66; 99]. The evidence is mixed regarding diabetes and differences in breast cancer risk based on menopausal status [67; 114]. When evaluating associations between use of specific ADs and breast cancer risk, studies assessing metformin use and risk of breast cancer have reported no increased risk of overall breast cancer [65; 80; 85; 99; 133], but one study reported a strong positive association between metformin use and triple-negative breast cancer [99]. Other ADs have been assessed, and no increased risk in breast cancer was reported, including GLP-1 receptor agonist medications [102] and pioglitazone [126]. In addition, studies evaluating users of SGLT2 inhibitors compared with users of DPP-4 inhibitors have reported no increased risk of breast cancer [25; 120; 132].

6.2.2 Bladder Cancer

Several observational studies have reported an elevated risk of bladder cancer amongst patients with T2DM [79], with increasing evidence suggesting that the association between T2DM and bladder cancer may vary by HbA1c level [100], race/ethnicity [15], and type of bladder cancer [130]. An association between ADs and bladder cancer has also been reported in the literature. Multiple observational studies have examined the association between use of pioglitazone and increased bladder cancer risk in patients with diabetes, with results from some studies suggesting that the risk increases with longer duration of pioglitazone use and higher cumulative doses [56; 122], whilst other studies show no clear patterns with longer treatment duration or increasing cumulative doses [74]. Studies evaluating users of SGLT2 inhibitors compared with users of GLP-1 receptor agonists and/or DPP-4 inhibitors have reported no positive association of bladder cancer risk with SGLT2 inhibitor use [1; 120]. However, a meta-analysis of 46 randomised controlled trials with a short-term duration (mean = 61 weeks), reported that users of SGLT2 inhibitors might have an increased odds of bladder cancer compared with placebo or users of other blood glucose–lowering treatments, although the estimates lacked precision due to the small number of events [121].

6.2.3 Rationale for the Dapagliflozin Cancer PASS

This present PASS was required by the EMA at the time of the initial approval of dapagliflozin for the treatment of T2DM in Europe and was conducted as part of the AstraZeneca Dapagliflozin Risk Management Plan to monitor the safety of dapagliflozin in real-world use and further evaluate the risk of cancer amongst patients using dapagliflozin, with a particular focus on breast cancer and bladder cancer. The final study protocol was approved on 27 June 2014 by Bristol-Myers Squibb Company, the study sponsor at the time, and amended 28 February 2017 and 06 June 2023 (see list of protocol amendments in Section 8). The eligibility criterion for identifying new users of dapagliflozin in this study was defined at a time when dapagliflozin's single approved indication in the US and the EU was the treatment of T2DM. Therefore, diagnosis of T2DM was not considered as an inclusion criterion for the study.

7. RESEARCH QUESTION AND OBJECTIVES

7.1 Research Question

The overall research questions are as follows:

- What is the risk of breast, bladder, and a composite of selected sex-specific cancers for patients with T2DM who are new users of dapagliflozin compared with those who are new users of other ADs?
- Is there differential medical surveillance (potential detection bias) for the diagnosis of these cancers for patients with T2DM who are new users of dapagliflozin compared with those who are new users of other antidiabetic treatments under study?

7.2 Study Objectives

7.2.1 Primary Objective

The primary objectives are as follows:

1 To compare the incidence of breast cancer, by insulin use at the index date, amongst females with T2DM who are new users of dapagliflozin and females who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

2 To compare the overall and sex-specific incidence of bladder cancer, by insulin use at the index date and by pioglitazone use, amongst patients with T2DM who are new users of dapagliflozin and those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

The index date is the date of the first prescription for each study medication. Antidiabetic drugs included in Appendix B, Table 28, were eligible for inclusion in the comparison group.

7.2.2 Secondary Objectives

The secondary objectives are as follows:

1 To compare during follow-up the frequency of several measures of HCRU (including outpatient visit frequencies and use of breast and bladder cancer screening and diagnostic tests), by insulin use at the index date, among patients with T2DM who are new users of dapagliflozin and those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

2 To compare baseline patient characteristics, by insulin use at the index date, amongst patients with T2DM who are new users of dapagliflozin and those who are new users of other ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy and to identify important prognostic variables that should be balanced between exposure groups and that should be included in the propensity scores used in the primary analyses.

3 To compare the composite incidence of selected cancers (prostate, colon/rectum, lung, stomach, NHL, and melanoma of skin), by insulin use at the index date, amongst males with T2DM who are new users of dapagliflozin and those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

4 To compare the composite incidence of selected cancers (colon/rectum, lung, corpus uteri, ovary, stomach, NHL, and melanoma of skin), by insulin use at the index date, amongst females with T2DM who are new users of dapagliflozin and those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

8. AMENDMENTS AND UPDATES

Minor updates to the cancer study protocol version 3.0 of 28 February 2017 were incorporated into amended protocol version 3.1 of 06 June 2023 (a list of stand-alone documents is presented in Appendix A). All substantial protocol amendments since the start of data collection are listed in Table 2.

Table 2 Amendments and Updates

Version number	Date	Section of study protocol	Amendment or update	Reason
3.1	06 Jun 2023	Section 7, Adverse Event Reporting	Updated adverse event language	Reporting of adverse events is not required for secondary database studies
3.0	28 Feb 2017	Synopsis; Section 3.2.3, Exclusion Criteria	Revised exclusion criteria: added exclusion of patients with previous non-dapagliflozin SGLT2 inhibitor use; clarified exclusion of patients with a prior history of any "invasive" cancer; clarified criteria that apply to bladder cancer only and to breast cancer only	Prevent potential issues with associations due to the SGLT2 inhibitor class effect; limit prior history of cancer as exclusion to invasive cancers only; incorporate edits to breast and bladder cancer exclusions
3.0	28 Feb 2017	Synopsis; Section 3.2.5, Selection of Subjects; Section 4.1.1, Propensity Score Approach	Revised sampling statement; removed restriction to limit selection of comparators to only four	Allow inclusion of all eligible comparators to minimise dropping of dapagliflozin users due to matching criteria
3.0	28 Feb 2017	Synopsis; Section 3.2.5, Selection of Subjects	Revised matching criteria: removed matching on duration of history/lookback time	Maximise matching efficiency for potential matching implementation
3.0	28 Feb 2017	Synopsis; Section 4.1.1, Propensity Score Approach	Calendar year of index date was added to propensity score models instead of fitting separate annual models	Analysis simplification
3.0	28 Feb 2017	Synopsis	Clarification of analysis of treatment switching: added text on analysis of comparator AD users if switching to dapagliflozin occurred after the index date	Ensure dapagliflozin was treated as the primary exposure of interest
3.0	28 Feb 2017	Synopsis; Section 4.1.7, Evaluation of Time-Varying Factors	New section: added analysis of time-varying variables (ie, severity of diabetes, changes in intensification of antidiabetic treatments)	Describe changes in time-varying variables during follow-up
3.0	28 Feb 2017	Section 3.2.6, Cohort Entry and Follow-up	Clarification of censoring of follow-up: added censoring of follow-up if a patient developed type 1 diabetes mellitus	Type 1 diabetes mellitus was an exclusion criterion

Version number	Date	Section of study protocol	Amendment or update	Reason
3.0	28 Feb 2017	Section 3.2.6, Cohort Entry and Follow-up; Section 4.1.9, Sensitivity Analysis	Added sensitivity analysis allowing continuing follow-up until first occurrence of each specific cancer endpoint	Further assess cancer risk regardless of occurrence of a prior study outcome during follow-up
3.0	28 Feb 2017	Section 3.3.1, Data Sources	Removed availability of free text in CPRD and added use of HES-linked data for identification of exclusions	Free text is no longer available in CPRD; HES-linked data is considered useful for identifying cancer diagnoses as exclusions
3.0	28 Feb 2017	Section 3.3.2, Data Source	Revised source of cancer data in PHARMO: removed Eindhoven Regional Cancer Registry and added Netherlands Cancer Registry	Access to Netherlands Cancer Registry became available
3.0	28 Feb 2017	Section 3.4.1.1, Electronic Case Identification	Added ICD-10 codes for identification of potential cases: added text on use of HES-linked data for case identification	HES-linked data were used to identify cases in CPRD
3.0	28 Feb 2017	Section 3.4.1.2, Validation of Outcomes	Removed review of free text for outcome validation in CPRD	Free text is no longer available in CPRD
3.0	28 Feb 2017	Section 3.4.3, Other Covariates/ Control Variables	Added covariates for inclusion in propensity score model: duration of lookback time and socioeconomic status	To make the list of covariates included in propensity score models more relevant
3.0	28 Feb 2017	Section 3.4.3, Other Covariates/ Control Variables; Table 1, Breast Cancer Covariates in Data Source	Revised list of covariates: removed oestrogen receptor status, *BRCA*/*BRCA1* mutations, and breast carcinoma in situ; added opioids as a covariate medication	To make the list of covariates included in propensity score models more relevant
3.0	28 Feb 2017	Section 3.4.3., Other Covariates/ Control Variables; Table 1, Breast Cancer Covariates; and Table 2, Bladder Cancer Covariates	Added opioids as a covariate medication	Recommended by clinical reviewer of interim report
3.0	28 Feb 2017	Section 4.1.1, Propensity Score Approach	Added the use of absolute standardised difference to assess balance of baseline covariates	This is a more standardised approach to propensity score modelling

Version number	Date	Section of study protocol	Amendment or update	Reason
3.0	28 Feb 2017	Section 4.1.1, Primary Objective 1: Calculation and Comparison of Female Breast Cancer Incidence	Clarified end of follow-up determinants: added date of first diagnosis of other invasive cancer as a determinant for end of follow-up in main analyses	To clarify analysis methods
3.0	28 Feb 2017	Section 4.1.9, Sensitivity Analyses	Added text on additional sensitivity analyses: added analyses on time since first exposure, lag time analysis, and all-cause mortality analysis	To further evaluate cancer risk in relation to exposure
3.0	28 Feb 2017	Section 4.1.10, Pooled Analyses	Specified analysis techniques to be used to pool data	To clarify analysis methods
3.0	28 Feb 2017	Section 4.2, Milestones	Specified criterion to trigger interim comparative analysis	To clarify methods
3.0	28 Feb 2017	Section 5.3, Other Sources of Bias	Expanded potential sources of bias	Acknowledged potential underestimate for secondary female composite cancer endpoint
3.0	28 Feb 2017	Section 5.4, Study Size	Revised estimates of person-years of dapagliflozin exposure at 10 years	Updated estimates of dapagliflozin person-years of exposure
3.0	28 Feb 2017	Appendix 1	Revised list of Read codes for outcome	Based on clinical review and updated CPRD browsers
3.0	28 Feb 2017	Appendix 2	Revised list of ICD-O-3 codes for outcome: added topography codes for skin cancer and trachea and removed codes for in situ malignancies as outcomes	To make the ICD-O-3 code list more comprehensive and ensure that in situ malignancies were not counted as study outcomes (except for bladder cancer)
3.0	28 Feb 2017	Appendix 3	Revised list of ICD-9-CM codes for outcome: removed codes for in situ malignancies as outcomes	In situ malignancies were not study outcomes
3.0	28 Feb 2017	Appendix 4	Revised list of Read codes for HCRU outcomes	Based on clinical review and updated CPRD browsers
3.0	28 Feb 2017	Appendix 4	Revised list of covariates for inclusion in propensity score model: removed covariates not applicable to cancer study and added other covariates of interest	To make the list of covariates included in propensity scores models more relevant
3.0	28 Feb 2017	Appendix 5	Revised list of comparator antidiabetic drugs	Based on new information

HES = Hospital Episode Statistics.

9. RESEARCH METHODS

9.1 Study Design

This is a non-interventional PASS conducted using data from four longitudinal population-based sources: CPRD—specifically, the GOLD data set and HES data—in the UK; PHARMO in the Netherlands; the HIRD® in the US; and the US Medicare research database for individuals aged 65 years or older and individuals with permanent disabilities. The study age range is 40 years or older in the CPRD and PHARMO, 40 to 64 years in the HIRD, and 65 years or older in Medicare.

This cohort study was conducted with an active-comparator, new-user design [81]. The primary exposure was dapagliflozin use newly initiated during the study period by eligible patients with or without concomitant use of insulin or any other AD. SGLT2 inhibitors other than dapagliflozin were excluded. Comparator AD exposure was defined as new use—initiated during the study period—of an AD other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy, with or without concomitant use of any other AD (other than SGLT2 inhibitors) at the index date (see Section 9.4.2.2 for details).

The primary study outcomes were female invasive breast cancer and in situ or invasive bladder cancer. The primary goal of the study was to estimate and compare the incidence rate for female invasive breast cancer and the sex-specific and sex-combined incidence of in situ and invasive bladder cancer amongst new users of dapagliflozin versus new users of eligible comparator ADs in the context of routine therapeutic use of these drugs.

Propensity scores were calculated for each cohort member within each data source for each outcome overall and by insulin use at the index date—as well as by pioglitazone use at the index date only for the bladder cancer cohorts—based on baseline information available any time before or on the index date, unless otherwise specified. Propensity scores incorporated measured potential predictors of the outcomes as independent variables and exposure cohort status as the dependent variable. Effect estimates were adjusted for possible confounding variables using propensity score stratification. At this final 120-month analysis, the study period covered up to 10 years, with interim analyses conducted approximately every 24 months. The cohort design and the long duration of the study allowed accumulation of sufficient sample size to assess the risk of cancer associated with exposure to study medications.

The study adhered to the *Guidelines for Good Pharmacoepidemiology Practices*, issued by the International Society for Pharmacoepidemiology [69] and the FDA's *Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data* [48]. This study also followed the guidance set forth in the EMA's *Guideline*

on Good Pharmacovigilance Practices: Module VIII – Post-Authorisation Safety Studies [39]. The study protocol is posted to the HMA-EMA Catalogue of RWD Studies (Study ID 49352) [64]. The contract between RTI Health Solutions and AstraZeneca includes independent publication rights.

The study protocol was reviewed and approved by the RTI International institutional review board (IRB) on 20 November 2015. The UK Medicines Healthcare products Regulatory Agency's Independent Scientific Advisory Committee (ISAC) approved the CPRD component of the study on 21 December 2015; the institutional review board of STIZON, Utrecht, the Netherlands, approved the PHARMO component of the study on 12 May 2016; the US Centers for Medicare and Medicaid Services (CMS) Privacy Board approved the Medicare component of the study on 11 January 2018; and the New England IRB approved the HIRD component of the study on 19 March 2018.

9.2 Setting

The study was conducted in four population-based health care data sources (see Section 9.5 for further details on these data sources):

- CPRD in the UK, specifically, CPRD GOLD
- PHARMO in the Netherlands
- The HIRD in the US
- Medicare in the US

The beginning of the study period varied by country and was defined as the day after regulatory approval of dapagliflozin in each country. The study period ended at the last date of observation available at the time of data extraction in each data source. The study period for the final analysis for each data source is shown in Table 3. The study period covered approximately 10 years, up to the end date of the last data cut in the HIRD, 9 years in PHARMO, and 8 years in CPRD and Medicare. For CPRD, at the time of data extraction for this final (120-month) analysis, a new release of the HES data set had not become available since the fourth interim (96-month) analysis to link to the latest available CPRD GOLD data. To limit the impact that the delayed release of the updated HES data set would have on study timelines and study completion, the same CPRD GOLD data and linked HES data set that were extracted for the fourth interim (96-month) analysis (data through March 2021) were used for the final (120-month) analysis; therefore, the study period for CPRD for the final (120-month) analysis was the same as the study period for the fourth interim (96-month) analysis.

Table 3 Study Period for Final (120-Month) Analysis in Each Data Source

Data source	Start date	Latest date of available data[a]
CPRD	13 November 2012	31 March 2021[b]
PHARMO	01 November 2013	31 December 2022
HIRD	09 January 2014	30 September 2023
Medicare	09 January 2014	31 December 2021

[a] The end date is the latest date that prescription (CPRD) or dispensing (PHARMO, the HIRD, and Medicare) data were available. The end date accounts for a lag time at the time of data extraction of approximately 17 months for linked CPRD GOLD and HES data, 18 months for PHARMO linkage to the Netherlands Cancer Registry, 3 months for HIRD data, and 2 years for Medicare data.

[b] At the time of data extraction for the final (120-month) analysis, the CPRD GOLD linkage to HES was only available for data through March 2021.

CPRD GOLD = General Practitioner Online Database of CPRD; HES = Hospital Episode Statistics.

9.3 Subjects

All patients newly initiating dapagliflozin (with or without concomitant use of any other AD) or newly initiating an eligible comparator AD (with or without concomitant use of any other AD) during the study period were evaluated for selection into the study cohorts individually for each data source (see Appendix C, List of Codes for Dapagliflozin). Eligible comparator ADs were medications other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy (see Appendix B, List of Relevant Antidiabetic Drugs).

New use (newly prescribed or dispensed) of a study medication was defined as the first recorded prescription or dispensing for dapagliflozin or the eligible comparator AD on or after the start of the study period, without any recorded prescription or dispensing for that medication before the index date, based on all available data, with a minimum of 180 days available before the initial prescription or dispensing. The *index date* was defined as the date the patient received a new prescription or dispensing of either dapagliflozin (single-entity dapagliflozin or a fixed-dose combination) or an eligible comparator AD on or after the beginning of the study period, after meeting all inclusion criteria.

The study analysis cohorts were created within each data source by first identifying all dates of new use of dapagliflozin and comparator ADs that occurred during the study period. For patients newly initiating more than one eligible comparator AD in the observation period, all new-use dates of eligible comparator ADs were considered eligible for entry into the comparator AD cohort. The inclusion criteria and exclusion criteria (Section 9.3.1) were then applied, resulting in a cohort of all potential eligible index dates. However, a patient could be selected into the comparator AD cohort only once, and the first eligible new-use date determined the patient's index date and index treatment.

Separate outcome cohorts were then created by applying the cancer outcome–specific exclusion criteria. All remaining patients were then analysed in separate cancer-specific cohorts.

The study design—including cohort selection and eligibility, covariate assessment, and follow-up—is illustrated in Figure 1.

Figure 1 **Study Design Schematic of Cohort Selection and Eligibility, Covariate Assessment, and Follow-up**

[a] Individuals were excluded if they did not have continuous enrolment in the data source for at least 180 days before and including the index date. Continuous enrolment was defined as follows: CPRD, continuous registration in an up-to-standard participating general medical practice; PHARMO, continuous up-to-standard data in the PHARMO Data Network; HIRD, complete pharmacy and medical coverage in a health

insurance plan with no enrolment gaps greater than 30 days; Medicare, those originally entitled to Medicare coverage due to age (65 years or older) enrolled in fee-for-service insurance in Parts A (hospital insurance), B (medical insurance), and D (prescription drug coverage) with no enrolment gaps greater than 30 days and no enrolment in a managed care plan.

[b] Applicable only to cohorts assessing the bladder cancer outcome.

[c] Applicable only to cohorts assessing the breast cancer outcome.

[d] Exclusions by age for each data source: CPRD and PHARMO, ages < 40 years; HIRD, ages < 40 years and > 64 years; Medicare, ages < 65 years.

[e] Medicare only. Assessed using data that were recorded most recently before the index date.

[f] Excluded invasive cancer other than non-melanoma skin cancer.

[g] For the time period before January 2020, calendar year was assessed as a covariate. To account for variability during the COVID-19 pandemic, the period of January to February 2020 was assessed as a covariate; starting in March 2020, 6-month time periods were assessed as covariates through the end of the study period.

[h] CPRD, the HIRD, and Medicare only.

[i] Includes CPRD (index of multiple socioeconomic deprivation), PHARMO (SES score based on postal code of residence), and Medicare (low-income subsidy). Assessed using data that were recorded most recently before the index date.

[j] Body mass index, smoking history, and alcohol use were available in CPRD GOLD data and in the GP subset of PHARMO data; these data elements were not available in the HIRD or Medicare.

[k] Individuals were followed until the occurrence of the cancer outcome of interest or they were censored due to any other cancer diagnosis, end of patient-specific data available in data source, initiation of dapagliflozin (applicable to comparator AD group only), initiation of a non-dapagliflozin SGLT2 inhibitor, reached 65 years of age (applicable to the HIRD only), death, end of study period.

Note: Figure template is from Schneeweiss et al [115].

9.3.1 Eligibility Criteria

Individuals from each data source meeting all the inclusion criteria and none of the exclusion criteria were eligible for inclusion in the study cohorts (Figure 1).

9.3.1.1 Inclusion Criteria

Inclusion criteria, evaluated at each potential index date for dapagliflozin or comparator AD, are listed as follows for each data source:

- CPRD: currently registered in an up-to-standard participating general medical practice for at least 180 days before and including the index date; aged 40 years or older
- PHARMO: up-to-standard data in the PHARMO Data Network for at least 180 days before and including the index date; aged 40 years or older
- HIRD: complete pharmacy and medical coverage in a health insurance plan for at least 180 days before and including the index date with no enrolment gaps greater than 30 days; aged 40-64 years
- Medicare: enrolled in fee-for-service insurance in Parts A, B, and D for at least 180 days before and including the index date with no enrolment gaps greater than 30 days; aged 65 years or older; not enrolled because of disability or end-stage renal disease; resident of a US state or the District of Columbia; not enrolled in managed care coverage.

9.3.1.2 Exclusion Criteria

Specified medical conditions and procedures listed below as exclusion criteria were assessed at any time in the available history before and including the index date, unless otherwise specified. For all cancer outcomes, patients fulfilling any of the following criteria were excluded:

- Met the operational definition of type 1 diabetes mellitus (T1DM) for the respective data source—ie, had at least one diagnosis of T1DM in CPRD, at least one diagnosis code for T1DM or single insulin dispensing or aged < 30 years at first insulin dispensing recorded on or before the potential index date in PHARMO, or at least two claims on separate dates with a diagnosis code for T1DM in the HIRD and Medicare—on or before the index date.
- Had a recorded prescription or dispensing for an SGLT2 inhibitor medication other than dapagliflozin on or before the index date.
- Had a recorded prescription or dispensing of dapagliflozin before the index date.
- Had any diagnosis of invasive cancer (other than non-melanoma skin cancer, defined as basal and squamous cell skin cancers) recorded at any time before or on the index date.

Additional exclusions specific to the primary cancer outcomes are listed below.

For the breast cancer outcome only:

- Had a procedural code recorded within the 180 days before and including the index date for a breast biopsy.

For the bladder cancer outcome only:

- Had a diagnosis of in situ bladder cancer recorded before or on the index date.
- Met the operational electronic definition of haematuria—ie, had a code for haematuria within the 180 days before and including the index date—for the respective data source (haematuria was defined based on ICD-9-CM or ICD-10-CM, Read or ICD-10 WHO diagnosis codes for haematuria and/or procedural codes for haematuria specific to each data source).
- Had a procedural code recorded within the 180 days before and including the index date for any of the following procedures:
  - Cystoscopy/bladder biopsy
  - Urine cytology

9.3.2 Study Follow-up

For all data sources, follow-up for each patient began on the day after the index date (ie, follow-up did not include the index date) and continued until the earliest occurrence of one of the following censoring events (Figure 1):

- Date of the first diagnosis of *any* of the cancer endpoints of interest defined in Section 9.4.1 (except for a sensitivity analysis in which follow-up continued until diagnosis of the specific cancer endpoint of interest [see Section 9.9.5.3])
- Date of death
- End of available patient-specific data in each data source (ie, end of enrolment in a health plan or study database)
  - For CPRD, the earlier of the following: date of transfer out of a practice (if applicable) or date of last data collection by the practice
  - For PHARMO, transfer out of a pharmacy
  - For the HIRD, the earlier of the following: date of end of health plan eligibility or end of study period
  - For Medicare, the earlier of the following: date of disenrolment from Part A, B, or D insurance or the date of start of managed care coverage
  - End of study period (for the final [120-month] analysis, the end date of available data/observation period [ie, the date after which no further data were collected] as specified in Section 9.2)
- First date that a non-dapagliflozin SGLT2 inhibitor was initiated (ie, date of prescription in CPRD or date of dispensing in PHARMO, the HIRD, and Medicare) (except for a sensitivity analysis in which this censoring criterion was removed, allowing follow-up to continue after initiation of a non-dapagliflozin SGLT2 inhibitor in either the dapagliflozin or the comparator AD group [see Section 9.9.5.3.2]). Follow-up was not censored if any other ADs were prescribed or dispensed in addition to dapagliflozin or the comparator AD after the index date. As described in the amended SAP, version 3.0 dated 22 April 2020 (see list of stand-alone documents in Appendix A), starting at the third interim (72-month) analysis, the proportion of individuals that were censored due to initiation of a non-dapagliflozin SGLT2 inhibitor in the dapagliflozin group and the comparator AD group was monitored to assess for potential differential censoring.
- Follow-up time in the comparator AD group was censored if dapagliflozin was prescribed or dispensed (except for a sensitivity analysis in which these individuals continued to contribute follow-up time to *both* the comparator AD group and the dapagliflozin group after initiating dapagliflozin [see Section 9.9.5.3.3]); in the main analysis, the subsequent follow-up was attributed to dapagliflozin exposure (see Section 9.4.2.3).
- In the HIRD, patients were censored on the day before their 65th birthday.
- Follow-up ended at the time of the first occurrence of any invasive cancer during follow-up (irrespective of whether the cancer was one of the study cancer endpoints).

Dates of censoring were applied on the date for which a censoring event, diagnosis, or prescription was identified, and the date of the censoring event was considered in the patient's time-at-risk period.

9.4 Variables

9.4.1 Outcome Variables

The following study outcomes were evaluated in each study cohort within each of the data sources separately.

The primary outcomes are as follows:

1 Female invasive breast cancer
2 Sex-combined and sex-specific invasive and in situ bladder cancer

The secondary outcomes are as follows:

1 Composite of selected invasive cancers (prostate, colon/rectum, lung, stomach, NHL, and melanoma of skin), amongst males
2 Composite of selected invasive cancers (colon/rectum, lung, corpus uteri, ovary, stomach, NHL, and melanoma of skin), amongst females
3 Health care resource utilisation measures (including outpatient visit frequencies and use of breast and bladder cancer screening and diagnostic tests)

9.4.1.1 Electronic Algorithms for Identification of Primary Outcomes

9.4.1.1.1 Invasive Breast Cancer (Amongst Females)

Provisional cases of invasive breast cancer (amongst females) were identified by electronic algorithms tailored to each data source, which are documented in the data analytic plan and shown in Appendix D, Table 32. Code lists for the specific diagnoses for the breast cancer outcome (Read codes, ICD-O-3 codes, ICD-9-CM codes, ICD-10-CM codes, and ICD-10 codes) are appended in the study protocol available in the HMA-EMA Catalogue of RWD Studies [9].

9.4.1.1.2 Invasive and In Situ Bladder Cancer (Amongst Females and Males)

Provisional cases of invasive or in situ bladder cancer were identified by electronic algorithms tailored to each data source, which are shown in Appendix D, Table 32. The code lists for the specific diagnoses for the bladder cancer outcome (Read codes, ICD-O-3, ICD-9-CM codes, ICD-10-CM codes, or ICD-10 codes) are included in the study protocol [9].

In PHARMO, linkage between patients in the PHARMO Data Network and the Pathology Registry [98] for identifying all cancer outcomes was implemented up to the fourth (96-

month) interim data cut. For the final (120-month) data cut, linkage to the NCR, which comprises information on all newly diagnosed cases of cancer in the Netherlands, was implemented for the identification of cancer cases. The PHARMO–NCR linked data were available through 31 December 2022 (end of study period for the PHARMO data); therefore, linkage to the Pathology Registry to supplement cancer outcome identification during the study period was not needed for the final data cut.

9.4.1.2 Validation of Primary Outcomes

Outcome validation of provisional female breast cancer cases and sex-combined bladder cancer cases was conducted during the second (48-month) interim analysis for CPRD, the HIRD, and Medicare and during the fourth (96-month) interim analysis for the HIRD and Medicare. Details on the final validation methods and results are documented in the fourth interim (96-month) study report dated 27 September 2023 and summarised in this current report in Section 10.5.1, Table 23.

During the final (120-month) analysis, data from eligible patients identified in the PHARMO data were linked to the NCR data to identify cancer cases. The NCR collects and reports cases based on tumour site and morphology and implements various data quality and control procedures [75]; therefore, additional case validation was not required for cancer cases in PHARMO.

9.4.1.3 Secondary Outcomes

9.4.1.3.1 Composite Invasive Cancer Endpoints

- Incident cancer of the prostate, colon/rectum, lung, or stomach; NHL; and melanoma of skin (amongst males)
- Incident cancer of the colon/rectum, lung, corpus uteri, ovary, or stomach; NHL; and melanoma of skin (amongst females)

Selection of the secondary cancer outcomes was based on the 10 leading cancers by incidence in the EU in 2012, as reported by Ferlay et al [51]. Provisional cases of the sex-specific composite invasive cancer endpoints were identified by electronic algorithms tailored to each data source and documented in Appendix D.

9.4.1.3.2 HCRU Measures

- Outpatient, hospital, emergency department, gynaecologist, urologist, and other specialist visits; in the HIRD, all outpatient encounters including laboratory tests and other encounters
- Breast biopsies amongst females
- Mammograms amongst females

- Cystoscopies
  - Bladder biopsies
  - Urine cytologies

The selection of codes for the HCRU outcomes were tailored to the data available in each data source.

9.4.2 Exposure Variables

For all cancer outcomes, the start of the exposure time was determined by the date of the index prescription or dispensing and started on the day after that prescription or dispensing date (see Figure 1). The handling of missing values for variables used to define exposure is described in Section 9.9.4.

9.4.2.1 Primary Exposure

The primary exposure of interest is dapagliflozin use newly initiated during the study period by eligible patients with or without concomitant use of insulin or any other AD except for other SGLT2 inhibitors. The identification of study ADs relied on Gemscript codes in CPRD, Anatomical Therapeutic Chemical (ATC) codes in PHARMO, National Drug Codes (NDCs) and Generic Product Identifier (GPI) codes in the HIRD, and NDCs in Medicare. The data source–specific lists of dapagliflozin codes are presented in Appendix C.

9.4.2.2 Comparator Antidiabetic Drug Exposure

Comparator AD exposure was defined as new use of an eligible AD, including monotherapy or combination therapy with alpha-glucosidase inhibitors, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, meglitinides, and other blood glucose–lowering drugs (pramlintide, bromocriptine). Sodium-glucose cotransporter 2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy were not considered as eligible comparators. Metformin qualified as a comparator AD only if coprescribed with a non-metformin or non-insulin eligible comparator AD; a sulfonylurea qualified as a comparator AD only if coprescribed with a non-sulfonylurea or non-insulin eligible comparator AD; free-dose metformin and sulfonylurea coprescribed with no other eligible AD was eligible as a comparator AD. Comparator AD exposures were selected on the basis of individual drug substances (generic type) within a class, not by drug class. Comparator ADs were analysed as a group, not by individual medications. Antidiabetic drugs listed in Appendix B, Table 28, were eligible for inclusion in the comparison group.

9.4.2.3 Index Exposure and Time at Risk

The index medication was defined based on records of prescriptions or dispensings of the study medication (dapagliflozin or comparator AD) that qualified the patient to enter the study. To identify the index medication, the first use of dapagliflozin or of each potential comparator AD within the data source was found in the patient's entire available history,

along with all periods of use of metformin and sulfonylureas. The time frame was then restricted to the study period.

For each patient, the start of exposure time at risk for each index medication was determined by the index prescription or dispensing. The exposure time at risk started on the day after the index date for that medication (see Figure 1) and was defined differently for the dapagliflozin cohort and the comparator AD cohort, as follows:

- The time at risk for individuals in the dapagliflozin cohort extended to the end of follow-up. That is, if an individual initiated dapagliflozin (index use), he or she remained at risk throughout follow-up even if the patient switched to another AD or if a comparator AD was initiated after the index date.
- The time at risk of individuals initiating a comparator AD (index use) was censored from the comparator AD group if and when they subsequently received dapagliflozin (except in a sensitivity analysis where this censoring was removed and patients contributed person-time to both the comparator AD and dapagliflozin groups [see Section 9.9.5.3.3]). These individuals then entered the dapagliflozin cohort (at the dapagliflozin index prescription or dispensing).
  - The time at risk accumulated by these comparator AD patients before entry into the dapagliflozin cohort was included in the comparator AD cohort for the analysis of cancer incidence rates.
- For patients in the comparator AD cohort, time at risk was not censored if eligible ADs other than dapagliflozin were prescribed or dispensed in addition to the comparator AD or if a patient switched to another AD (not dapagliflozin) after the index date.
- At no time during the study period for the main analysis did any patient contribute time at risk to both the dapagliflozin-exposed and comparator-exposed cohorts simultaneously, and any cancer event was counted only once (in the exposure category in which the patient was accumulating time at risk at the time of the cancer event). The only exception to this was a sensitivity analysis in which patients in the comparator AD group subsequently initiated dapagliflozin and contributed person-time to both the comparator AD and dapagliflozin groups simultaneously (see Section 9.9.5.3.3).

9.4.2.4 Cumulative Exposure

Mutually exclusive categories of cumulative exposure were created for the dapagliflozin and comparator AD exposure groups based on treatment duration. Cumulative exposure was calculated as the duration of continuous treatment received by each individual during follow-up, overall (not stratified by insulin use at the index date), by insulin use at the index date and, for bladder cancer only, by pioglitazone use at the index date. The following calculations of treatment duration were performed:

- The duration of each prescription was based on the number of days' supply.
- Overlapping times of days' supply between consecutive prescriptions or dispensings were concatenated with the overlapping time counting only once.

- Patients were considered to be receiving continuous treatment if the gap between the end of one prescription and the next prescription was 30 days or less (ie, the duration of the continuous treatment was from treatment initiation until day 31 after the end of the last prescription). If days' supply was not available, it was imputed within each data source from the mode of the known values of days' supply for the same drug and strength within each calendar year.

The duration of continuous treatment was categorised into the following mutually exclusive categories: less than 1 year, from 1 year up to 2 years (1 to < 2), between 2 and 5 years (2 to < 5), and 5 years or more (≥ 5). Different categories of treatment duration could be defined in the different data sources depending on its distribution. Patients contributed time of follow-up from the date of the first prescription/dispensing until the end of follow-up, as defined in Section 9.3.2.

Similarly, within each data source and separately for each cancer outcome cohort, the dapagliflozin cumulative dose during follow-up was also described overall (not stratified by insulin use at the index date) and by insulin use at the index date and, for bladder cancer only, by pioglitazone use at the index date. The cumulative dose was measured in milligrams (mg), and the following calculations were performed:

- Cumulative dose was estimated by summing the number of days' supply and multiplying by daily dose.
- Overlapping times of days' supply between consecutive prescriptions or dispensings were concatenated, with the overlapping time counting only once.
- Gaps in time between the end of drug supply of one prescription and the start of drug supply of the next prescription were not counted for calculation of cumulative dose. Cumulative doses for each prescription were added together regardless of gaps in time between prescriptions or dispensings.
- If days' supply was not available, this value was imputed within each data source from the mode of the known values of days' supply for the same drug and strength within each calendar year.
- If daily dose was not available, this value was imputed within each data source from the mode of the known values of daily dose for the same drug and strength within each calendar year.

9.4.2.5 Antidiabetic Drugs Used Concomitantly During the Index Exposure

9.4.2.5.1 INSULIN

An indicator variable for current insulin use at each index date (0 = no use; 1 = use) was created for all study participants. The variable included any type of insulin. The indicator variable was set to "1" if the date of the prescription or dispensing of insulin was the same date as the index date or if the date of the prescription or dispensing of insulin occurred within

the 90 days before the index date AND (1) the calculated duration of the insulin prescription or dispensing episode (prescription or dispensing plus days' supply) continued for at least 1 day past the index date or (2) if another prescription or dispensing for insulin occurred within 90 days after the index date. The insulin use indicator was used as a stratification variable in the analyses of all cancer outcomes.

9.4.2.5.2 PIOGLITAZONE

An indicator variable for current pioglitazone use at the index medication date (0 = no use; 1 = use) was created for all study participants. The indicator variable was set to "1" if the date of the prescription or dispensing of pioglitazone was the same date as the start date of the index AD prescription or dispensing or if the date of the prescription or dispensing of pioglitazone occurred within 90 days before the index date AND (1) the calculated duration of the pioglitazone prescription or dispensing episode continued for at least 1 day past the index date or (2) if another prescription/dispensing for pioglitazone occurred within 90 days after the index date.

The pioglitazone use indicator variable was used as a stratification variable in the analyses of the incidence of bladder cancer.

9.4.2.5.3 ADD-ON AND SWITCHING INDICATORS

The index medication could be initiated as monotherapy, added to another AD, switched from another AD to the index medication, or initiated as index combined therapy (more than one drug was initiated on the index date). Dichotomous indicator variables (0 = no, 1 = yes) were created to classify the index medication into categories. The indicator of use was assigned based on individual drugs, not on drug class. For creation of the indicator variables, three intervals of time were considered: interval 1 was the 90 days before (and not including) the index date, interval 2 was the study drug index date; and interval 3 was the 90 days after (and not including) the index date. The following dichotomous indicator variables for index medication exposure were created (for propensity score modelling, the categories described in this section were recategorised to generate fewer indicator variables to aid in model convergence; details are described in Section 9.8).

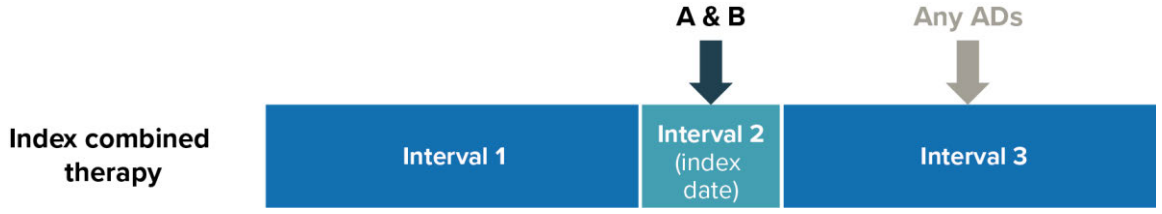
Index monotherapy with no prior treatment. As illustrated in Figure 2, only a single index drug substance was prescribed or dispensed on the date of the index medication (interval 2), and there was no prescription or dispensing for an AD or insulin in interval 1. Note that interval 3 could be less than 90 days and the definition of index monotherapy would still apply.

Figure 2 Index Monotherapy With No Prior Treatment

Notes: Interval 1 is the 90 days before (and not including) the index date; interval 2 is the study drug index date; interval 3 is the 90 days after (and not including) the index date. The letter A represents a specific AD. The grey arrow indicates that the prescription or dispensing may or may not occur.

Index combined therapy with no prior treatment. As illustrated in Figure 3, multiple drug substances were prescribed or dispensed at interval 2, and there was no prescription or dispensing for any AD or insulin in interval 1. Note that interval 3 could be less than 90 days and the definition of index combined therapy would still apply.

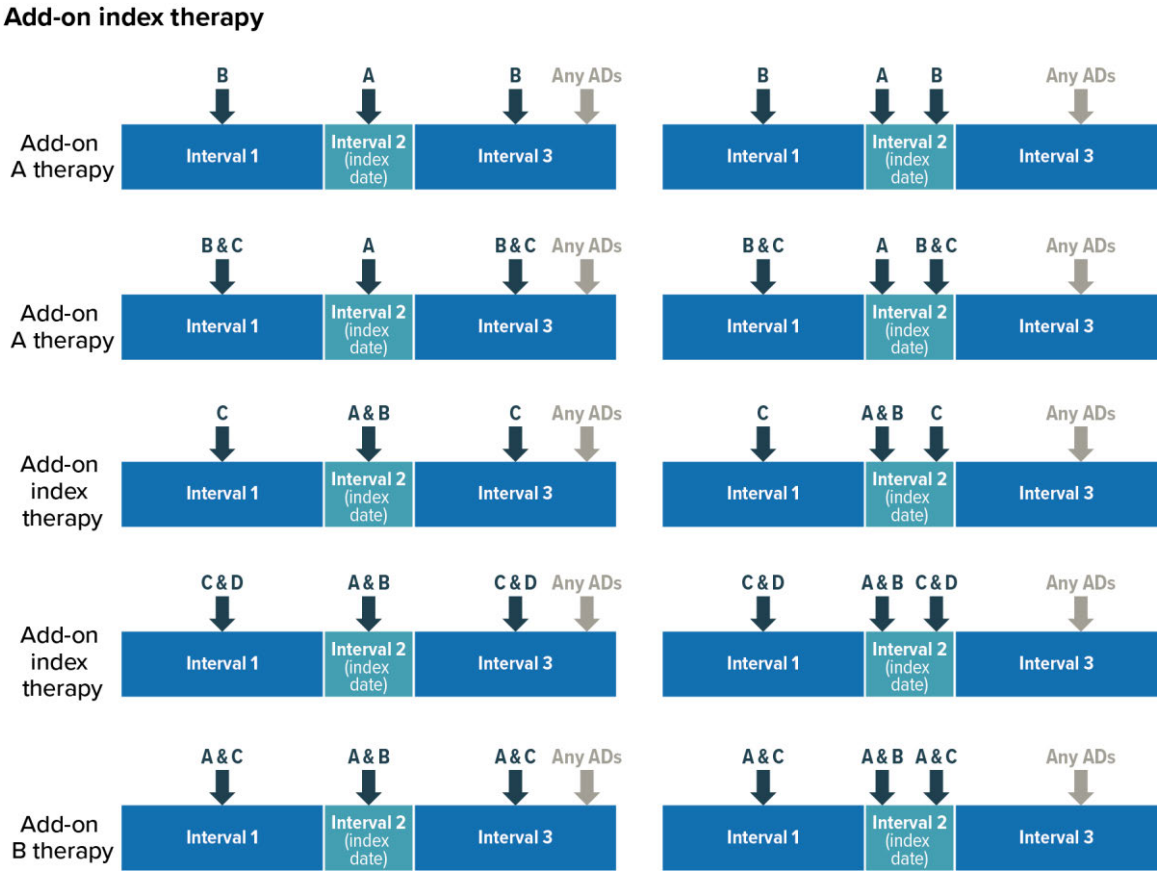
Figure 3 Index Combined Therapy With No Prior Treatment



Notes: Interval 1 is the 90 days before (and not including) the index date; interval 2 is the study drug index date; interval 3 is the 90 days after (and not including) the index date. The letters A and B represent specific but different ADs. The grey arrow indicates that the prescription or dispensing may or may not occur.

Add-on index therapy. As illustrated in Figure 4, an AD or insulin other than the index medication was prescribed or dispensed during interval 1, and then a subsequent prescription or dispensing for the same drug substance was identified during interval 2 or 3. If multiple drug substances were identified during interval 1, then all these drug substances would need a new prescription or dispensing during interval 2 or 3 to fit this category.

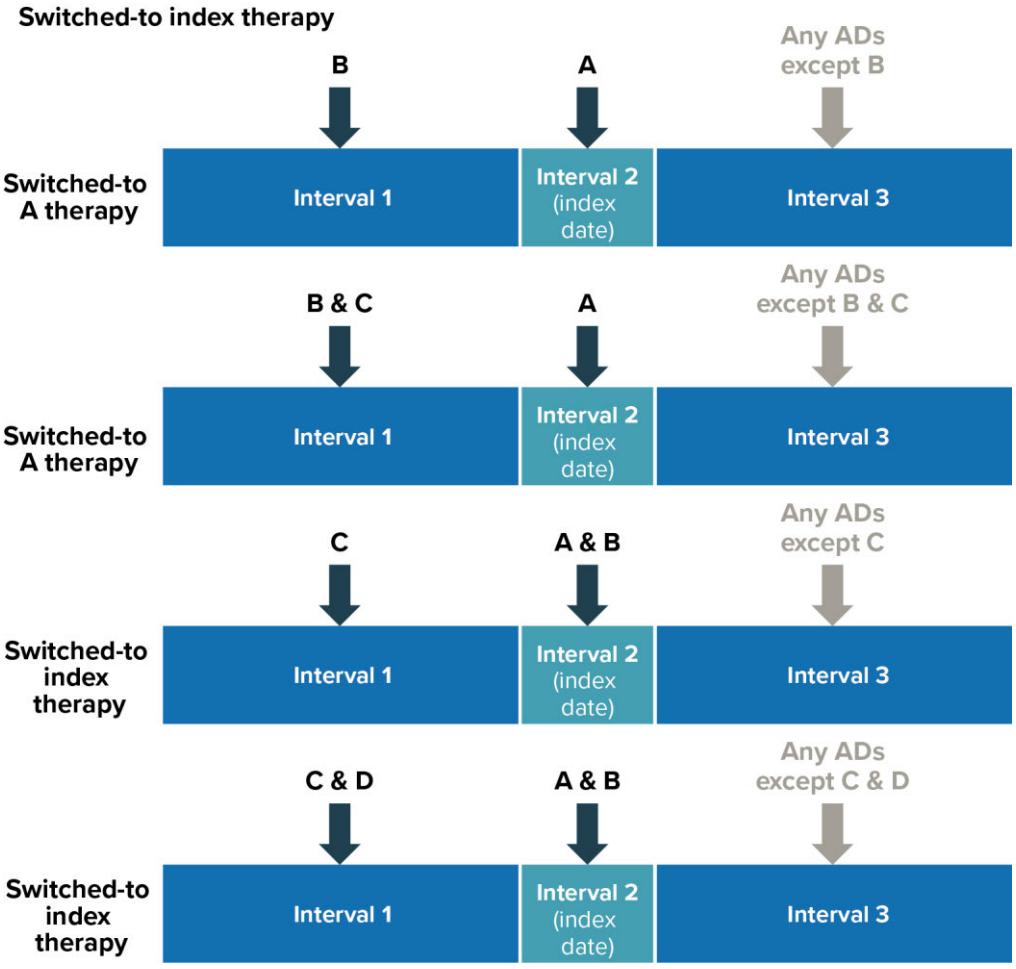
Figure 4 Add-on Index Therapy



Notes: Interval 1 is the 90 days before (and not including) the index date; interval 2 is the study drug index date; interval 3 is the 90 days after (and not including) the index date. The letters A, B, C, and D represent specific but different ADs. The grey arrows indicate that the prescription or dispensing may or may not occur.

Switched-to index therapy. As illustrated in Figure 5, a drug substance(s) other than the index AD was prescribed or dispensed during interval 1 and had no subsequent prescriptions or dispensings during intervals 2 and 3. If multiple drug substances were identified in interval 1, no additional prescriptions or dispensings for any of these substances could occur in intervals 2 and 3.

Figure 5 Switched-to Index Therapy

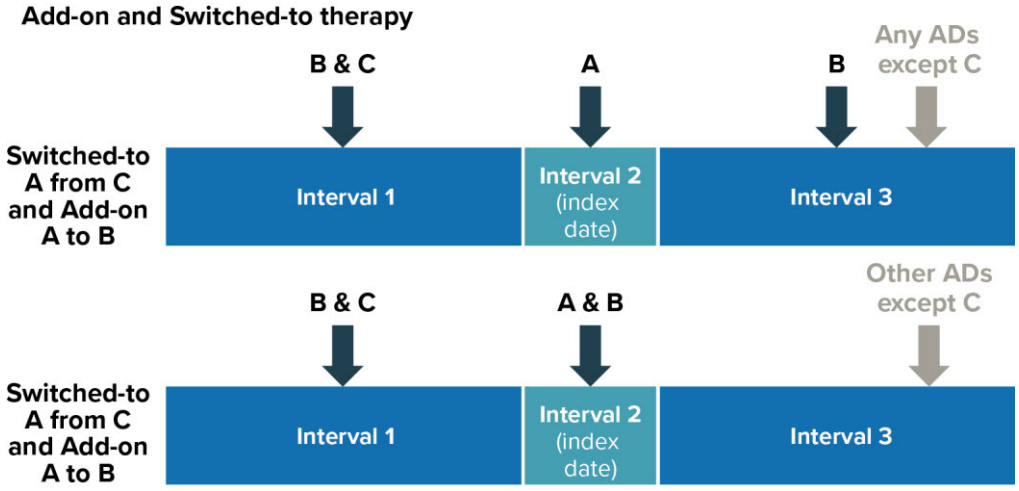


Notes: Interval 1 is the 90 days before (and not including) the index date; interval 2 is the study drug index date; interval 3 is the 90 days after (and not including) the index date. The letters A, B, C, and D represent specific but different ADs. The grey arrows indicate that the prescription or dispensing may or may not occur.

Add-on and switched-to index therapy. As illustrated in Figure 6, a patient had multiple drug substances with a reported prescription or dispensing in interval 1 and the following two criteria were met:

- At least one drug substance had a prescription or dispensing during interval 2 or interval 3.
- At least one drug substance had no prescription or dispensing during interval 2 and interval 3.

Figure 6 Add-on and Switched-to Index Therapy



Notes: Interval 1 is the 90 days before (and not including) the index date; interval 2 is the study drug index date; interval 3 is the 90 days after (and not including) the index date. The letters A, B, and C represent specific but different ADs. The grey arrows indicate that the prescription or dispensing may or may not occur.

Non-evaluable index treatment. A drug substance(s) was prescribed or dispensed during interval 1, and the patient had less than 90 days of follow-up (interval 3); therefore, it could not be determined if treatment had been added on or switched to or both.

9.4.2.5.4 Prior Use of Antidiabetic Drugs

To assess exposure to ADs before the index date, eligible ADs were grouped by the corresponding AD classes: insulins, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides. The number of different eligible AD classes prescribed or dispensed before the index date was assessed within the following time windows before the index date based on all available data: 12 months, more than 12 months through 24 months, and more than 24 months.

Separate variables for each of these time windows were created as proxies for the stage of treatment at baseline and were categorised based on the frequency distributions. Combination treatment with two different classes of ADs was counted in both AD classes. The value of

these variables was set to missing if the lookback period for a given patient did not extend into the corresponding time window.

9.4.3 Potential Confounding Variables

Variables that were evaluated as potential confounding variables at baseline—for use in propensity score models—were medical conditions related to diabetes severity, cancer risk factors, other medical comorbidities, selected medications, demographic and lifestyle factors as available in each data source, and HCRU variables (see Appendix E, Table 33, Table 34, and Table 35 for a complete list of these variables and specifications in the data sources). Race/ethnicity data were available for Medicare beginning with the second interim (48-month) analysis (the first analysis that included Medicare data) and were included in the final (120-month) analysis from CPRD and the HIRD in addition to Medicare; information on race/ethnicity is not reported in PHARMO data. Note that HCRU variables, assessed during follow-up, are part of the secondary study outcomes (see Section 9.4.1.3). As recommended by Brunelli et al [20], covariates were assessed using all available lookback data in each data source up to and including the index date, unless otherwise indicated (see Figure 1 and Appendix E, Table 33, Table 34, and Table 35). The earliest data for the lookback period differed across data sources (see Appendix F, Table 36 for a summary of the characteristics of the data sources). It is worth noting that in Medicare, no data are available before age 65 years. The minimum lookback period was 180 days including the index date. Medications were evaluated in the 180 days up to and including the index date. Lifestyle variables were assessed using all available data in each data source up to and including the index date, and HCRU variables were assessed in the 180 days before but not including the index date.

Medical conditions were identified in CPRD GOLD by using Read codes from diagnoses recorded in the electronic medical records by GPs and specialists; in PHARMO using ICD-9 and ICD-10 codes in the hospital data and, in the subset sample with GP data available, also using ICPC codes for some conditions eg, heart failure, chronic kidney disease, and T2DM; and in the HIRD and Medicare using ICD-9-CM diagnosis codes (ICD-10-CM codes in data recorded after October 2015). Some medical conditions were identified in the HIRD and Medicare data sources using algorithms comprising diagnosis, procedure (CPT or HCPCS), and/or medication (NDC) codes (see Appendix G, Table 37). Some medical conditions assessed at baseline were outcome specific, as shown in Appendix E, Table 33. To account for the additional approved indications for dapagliflozin since the start of the study, for the final analysis, heart failure and chronic kidney disease were added as medical condition covariates in the descriptive analyses and for consideration in the propensity score models. In light of the additional approved indications for dapagliflozin in recent years, T2DM was added as a medical condition variable for the composite cancer outcome cohorts to describe the proportion of individuals with a recorded T2DM diagnosis before or on the index date. The T2DM variable was not considered for inclusion in the propensity score models (see Appendix E, Table 33) as it was anticipated that nearly all patients would have a recorded

diagnosis of T2DM and that diagnosis of T2DM would be similarly distributed across the dapagliflozin and comparator AD groups. Additional post hoc analyses were conducted in the HIRD to assess this assumption and the impact that the violation of this assumption may have had on effect estimates (see Section 9.9.6 for details). Covariate medications were identified using Gemscript product codes from the electronic prescription information in CPRD GOLD data; ATC codes in PHARMO data; and GPI codes, NDCs, or HCPCS codes in the HIRD and Medicare. Medications were grouped into the categories shown in Appendix E, Table 34.

Lifestyle variables (BMI, smoking history, and alcohol use) were available in CPRD GOLD and for approximately 40% of patients with GP data available in PHARMO. The categories of some lifestyle variables are detailed in Section 9.8. For each variable, the most recent value on or before the index date was used, and the entire medical history was used as the lookback period (0, Table 35). Lifestyle variables were not available in the HIRD or Medicare (although diagnosis codes for obesity are recorded in the HIRD, obesity was not considered for inclusion as a covariate in the HIRD because the diagnosis codes for obesity are underreported in claims and the extent of missingness is unknown) [4; 118]. Socioeconomic status was recorded in CPRD, PHARMO, and Medicare. Other variables such as HbA1c test results were available only in CPRD and PHARMO, whilst the number of HbA1c tests in the 180 days before or on the index date were captured only in the HIRD and Medicare.

Baseline HCRU variables used as covariates were the number of outpatient encounters to a GP or outpatient clinic, number of hospitalisations, number of emergency department visits, and number of specialty care visits recorded. Categories of each variable are shown in 0, Table 35.

9.4.4 Time-Varying Variables

Changes in diabetes severity and in the intensity of antidiabetic treatments were assessed at baseline and at fixed 3-month (90-day) time intervals during follow-up.

9.4.4.1 Diabetes Severity

Changes in the severity of diabetes were assessed by deriving a diabetes severity score based on the Diabetes Complications Severity Index (DCSI), which has been validated using ICD-9-CM codes [24] and more recently updated and transformed to ICD-10-CM codes [57]. The DCSI consists of seven categories of diabetes complications with scores 0, 1, or 2, resulting in an overall score ranging from 0 to 13. In CPRD, PHARMO, and the HIRD, identification of the medical conditions in each DCSI category was based on a single diagnosis code. In PHARMO, identification of the medical conditions was based on hospital discharge diagnoses. In Medicare, medical conditions in the DCSI categories were identified through algorithms (ie, at least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code). Appendix H describes the DCSI categories and scores for retinopathy (Table 38), nephropathy (Table 39), and neuropathy (Table 40).

9.4.4.2 Antidiabetic Treatment Changes

Changes in the stage of treatment over time were assessed by the number of different classes of ADs that were simultaneously prescribed, defined by the number of AD classes (Table 4). The AD classes were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, meglitinides, and dapagliflozin (as the only SGLT2 inhibitor).

Table 4 Categories of Antidiabetic Drug Use During Follow-up

Treatment category	Definition
None	0 drugs (ie, treatment deintensified to 0 drugs)
First category	1 drug class
Second category	2 drug classes
Third category	≥ 3 drug classes or insulin in combination with any drug class
Fourth category	Insulin only

9.5 Data Sources and Measurement

The study was conducted according to the common protocol in the populations covered by the selected data sources. The main characteristics of the data sources and a summary of the collection and measurement of variables of interest are presented in Appendix F, Table 36.

9.5.1 Clinical Practice Research Datalink

The CPRD GOLD database contains diagnostic and prescribing information recorded by GPs as part of their routine clinical practice in the UK. As of January 2021, the data source contained information on approximately 19.5 million patients, 3.0 million of which were active (alive and currently registered), representing 4.5% of the UK population [27]. These data are linkable, at least partially, with other health care data sets (eg, hospitalisation records and national mortality data) via the patient's National Health Service number, sex, date of birth, and postal code. Detailed information on prescriptions written by GPs, including prescribed dosage, is automatically recorded in the data source. Read codes are used for diagnoses. Additional diagnostic and treatment information can be found in letters from specialists and hospitals and other sources.

The linkage of CPRD GOLD data to the HES database enables access to hospitalisation data including disease and procedural coding data in England. Linkage of CPRD GOLD data with other patient-level data sets, including the HES database, is available for a subset of English practices that have consented to participate in the linkage scheme. These linkages cover approximately 75% of the contributing CPRD practices in England, or roughly 58% of contributing CPRD practices in the UK [62]. CPRD GOLD linkage with HES is not available for Scotland, Wales, nor Northern Ireland.

9.5.2 PHARMO Data Network

The PHARMO Data Network is a population-based data source that combines anonymous electronic health care data from different primary and secondary health care settings in the Netherlands [75]. For this study, the outpatient pharmacy data, hospital data–*Hospital admissions*, clinical laboratory data, general practitioner data, Pathology Registry, and the NCR were used. The outpatient pharmacy data, which comprise GP- or specialist-prescribed health care products dispensed by community pharmacies, was used to select patients initiating dapagliflozin or newly initiating an eligible comparator given that ADs are predominantly prescribed in outpatient settings (eg, by GPs). Of note, for the data cut used in the fourth interim (96-month) analysis only, data from hospital-based outpatient pharmacies were also used to identify study patients, which resulted in duplicate records of AD dispensings for some patients due to limitations in identifying all instances of duplication from the lack of a common unique patient identifier across the community pharmacy data and hospital-based outpatient pharmacy data; hospital-based outpatient pharmacy data were not used in the final (120-month) data cut. The linkage of PHARMO data to other databases, such as the Pathology Registry and the NCR is conducted on a patient level through probabilistic linkage based on validated algorithms [75]. Detailed information on the methodology and validation of the probabilistic record linkage method used can be found elsewhere [60; 131]. The longitudinal nature of the PHARMO Data Network system enables follow-up of more than 10 million persons of a well-defined population in the Netherlands for an average of 12 years [75]. Currently, the PHARMO Data Network covers over 7 million active persons out of 17 million inhabitants of the Netherlands. The data collection period, catchment area, and overlap between data sources in the network differ. All electronic patient records in the PHARMO Data Network include information on age, sex, socioeconomic status, and mortality. In studies that aim to assess the occurrence of cancer outcomes amongst patients selected based on use of medications, such as the current study, it is pivotal to establish a catchment area that captures both medication use and cancer diagnoses. To achieve this, a valid linkage between NCR and the PHARMO Data Network requires patients to have complete and accurate patient-identifying information that is available for approximately 50% of the population within the PHARMO Data Network, thereby enabling a valid linkage for half of the PHARMO Data Network. The quality of the medical record linkage process with regional cancer registry data was evaluated in a random sample and resulted in a sensitivity of 98.3% (95% CI, 97.7%-98.7%) and a specificity of 99.5% (95% CI, 99.4%-99.7%) [131]. The data are collected, processed, linked, and anonymised by STIZON, an ISO/IEC 27001–certified and NEN 7510–certified foundation, compliant with the GDPR [117]. STIZON acts as a trusted third party between the multiple data sources and users of the anonymised data, who can request proportional study-specific data sets in accordance with the GDPR.

9.5.3 Healthcare Integrated Research Database

The HIRD contains a broad, clinically rich, and geographically diverse spectrum of longitudinal medical and pharmacy claims data from more than 91 million commercially insured and Medicare Advantage health plan members across the US in the Elevance Health system and is directly accessed by Carelon Research (formerly HealthCore), a fully owned but independently operating subsidiary of Elevance Health [13]. Member enrolment, medical care (professional and facility claims), outpatient prescription drug use, outpatient laboratory test result data, and HCRU may be tracked for health plan members in the database dating back to January 2006. The HIRD has the ability to link its claims data to complementary data sources, including inpatient and outpatient medical records, national vital statistics records, cancer and vaccine registries (state-by-state), disease and device registries, member and provider surveys, and point-of-care clinical data. The current study used data restricted to commercially insured beneficiaries under the age of 65 years.

9.5.4 Medicare

Medicare is a federally sponsored health insurance programme in the US that, in calendar year 2021, offered health coverage to 63.9 million people, including 55.9 million people aged 65 years or older and 8.0 million non-elderly people with a permanent disability [26]. Most adults become eligible for Medicare when they reach 65 years of age, although younger adults can qualify if they are permanently disabled or have end-stage renal disease. Of the total enrolees, 36.4 million have original (fee-for-service) Medicare enrolment. In 2021, Medicare beneficiaries included close to 100% of the US population aged 65 years or older [26; 128]. From 2013 through 2016, 21.4% of adults aged 65 years or older had been diagnosed with diabetes [23], and most would have had Medicare coverage. Therefore, Medicare data are particularly useful for the current study. Medicare consists of Part A, which is hospital insurance; Part B, which covers physician services and outpatient care; and Part D, which is outpatient prescription drug coverage. Parts B and D are optional, and enrolees must pay a monthly premium for this coverage. Part D coverage has been available since 2006 and is purchased by beneficiaries through private insurance companies approved by Medicare. As of 2021, 76% of Medicare beneficiaries were enrolled in Part D [26]. The current study used data restricted to fee-for-service Medicare beneficiaries with Part A, Part B, and Part D coverage.

9.6 Bias

An exposure propensity score approach was used to control for confounding by observed covariates because the number of outcomes was expected to be small in relation to the large number of possible confounding factors [11]. The propensity score for each patient is the predicted probability of being assigned to a particular treatment conditional on a set of observed covariates [18; 28; 101]. According to simulation studies, variables that are unrelated to the exposure but are related to the outcome increase the precision of the estimated exposure effect without increasing bias and should always be included in the estimation of

propensity scores [22]. In contrast, inclusion of variables that are related to the exposure but not to the outcome can decrease precision of the estimated effect of exposure without decreasing bias.

Propensity scores were estimated for each eligible patient at the index date after the relevant cohort exclusions were applied for each outcome cohort in each data source. For patients who switched during follow-up from a comparator AD to dapagliflozin, separate propensity scores were calculated for the comparator AD index date and for the dapagliflozin index date.

Construction of propensity score models is described in Section 9.9.3.2. Covariates that were considered are described in Section 9.4.3 and listed in Appendix E. Trimming of propensity score values below the 2.5th percentile value of the comparator AD–exposed distribution of propensity scores and values greater than the 97.5th of the dapagliflozin-exposed distribution of propensity scores removed the least exchangeable individuals and lessened the potential for confounding by retaining those with greatest equipoise for treatment (see Section 9.9.3.2.2). All comparative analyses were conducted in the propensity score–trimmed cohorts (hereafter, the "trimmed cohorts" or "trimmed population").

Misclassification of the outcomes may have occurred due to several reasons, including if primary cancer diagnoses recorded as part of the initial diagnostic workup included "rule out" diagnoses. Identification of outcomes through record linkage with hospital records (CPRD) or cancer and pathology registries (PHARMO) minimised the potential for outcome misclassification, although in PHARMO, misclassification may occur due to the probabilistic linkage to cancer registry data. In the HIRD and Medicare, the identification of provisional cancer cases required the presence of two claims associated with one of the specific cancer diagnosis codes on separate dates, within 60 days of each other. Validation of the algorithms used to identify provisional cases of the female breast cancer and bladder cancer outcomes was performed in CPRD at the second interim analysis and in the HIRD and Medicare at the second and fourth interim analyses; the final PPVs (95% CIs) were presented in the fourth interim report. At the final analysis, for the female breast cancer cohorts and sex-combined bladder cancer cohorts in CPRD, the HIRD, and Medicare, the estimated PPVs and quantitative bias analysis were used to assess whether the interpretation of the primary study results would change under the hypothetical worst-case scenario of potential differential outcome misclassification (Section 9.9.8).

9.7 Study Size

The sample size at study completion was dependent on the market uptake of dapagliflozin in the UK, the Netherlands, and the US, and was not known before the final analysis. It was estimated in the study protocol that, at the end of the 10-year study period, there would be a total of approximately 850,000 person-years of follow-up amongst new users of dapagliflozin across all data sources. Because uptake of dapagliflozin was lower than anticipated during the

early years of study implementation, starting with the third interim (72-month) analysis, to provide a more accurate study size and precision estimate anticipated for the final (120-month) data analysis, sample size projections and statistical precision estimates were updated based on the observed uptake of dapagliflozin use and trends in the capture of new users within each data source. The updated study size projections and statistical precision estimates were included in the third (72-month) and fourth (96-month) interim reports.

Data source–specific assumptions based on the actual observed dapagliflozin exposure and person-years accumulated up to the 96-month data cut in the female breast cancer and sex-combined bladder cancer cohorts after propensity score trimming and included in the fourth interim (96-month) report are shown in Table 5.

Table 5 Study Size Assumptions From the Interim Report 4 (96-Month) Analysis Used to Estimate Study Size Projections and Precision for the Final (120-Month) Analysis, by Data Source

Data source	Average annual change in dapagliflozin use, breast cancer cohort	Average annual change in dapagliflozin use, bladder cancer cohort	Drop-out due to loss to follow-up/death or length of follow-up	Excluded by propensity score trimming
CPRD	−4%	−2%	15%	25%
PHARMO	−1%	3%	5%	20%
HIRD	17%	18%	1.5-year follow-up	15%
Medicare	26%	26%	10%	15%

Estimates of the study precision focused on calculating the minimum detectable value of the upper limit of the 95% CI for the IRR of exposure to dapagliflozin versus comparator AD for female breast cancer and for sex-combined bladder cancer cohorts (Table 6). The calculations were derived from the projected study person-years of dapagliflozin exposure for the female breast cancer and sex-combined bladder cancer cohorts, the observed ratio of comparator AD new users to dapagliflozin new users, and the weighted average incidence rates per 10,000 for the female breast and sex-combined bladder cancer outcomes, assuming an IRR of 1 and 80% power. This approach avoids the drawbacks of study precision estimates calculated based on statistical power, which rely heavily on statistical significance testing [110].

Table 6 Projected Dapagliflozin Person-time Exposure and Statistical Precision Estimates at Study End (Based on the Fourth Interim [96-Month] Analysis): Female Breast Cancer and Sex-Combined Bladder Cancer Cohorts, Overall and Stratified by Insulin Use at the Index Date, by Data Source; Propensity Score–Trimmed Analysis Samples

Data source	Insulin use at the index date	Female breast cancer cohorts		Sex-combined bladder cancer cohorts	
		Projected dapagliflozin person-years[a]	Upper limit of IRR 95% CI[b]	Projected dapagliflozin person-years[a]	Upper limit of IRR 95% CI[b]
CPRD	Insulin	~5,000	2.4	~12,460	3.6
	No insulin	~28,300	1.4	~70,600	1.7
	Overall	~33,300	1.4	~83,060	1.6
PHARMO	Insulin	~1,300	5.0	~3,010	11.1
	No insulin	~7,300	2.0	~17,060	2.8
	Overall	~8,600	1.9	~20,070	2.6
HIRD	Insulin	~5,950	2.1	~14,400	5.5
	No insulin	~33,850	1.4	~81,570	2.1
	Overall	~39,800	1.3	~95,960	1.9
Medicare	Insulin	~7,830	1.8	~15,090	1.6
	No insulin	~ 44,390	1.3	~85,510	1.2
	Overall	~52,220	1.2	~100,600	1.2

[a] The total *projected* person-years of dapagliflozin exposure at study end, the final (120-month) data cut, based on the observed person-years of dapagliflozin exposure at the fourth interim (96-month) data cut and the assumptions presented in Table 5.

[b] Minimum detectable value of the upper 95% confidence limit of the IRR with 80% power, assuming a true IRR of 1.0.

9.8 Data Transformation

9.8.1 Data Management

The analyses of the selected cohorts were conducted by RTI Health Solutions and the research partners (Carelon Research in the US and PHARMO Institute in the Netherlands) using SAS software for Windows or SAS for Linux, version 9.4 or higher (SAS Institute, Inc., Cary, North Carolina, US). All analyses on CPRD data, PHARMO data, and HIRD data were run using a validated SAS server; all programs, data, and outputs were stored and executed on the validated SAS server. Analyses on Medicare data were conducted within the CMS Virtual Research Data Center, a restricted virtual research environment in which data management and statistical analysis activities can be conducted whilst accessing Medicare individual-level data under a project-specific data use agreement approved by the CMS privacy board. Data management for each data source was conducted in accordance with institutional procedures

and/or standard operating procedures. Routine procedures included checking electronic files, maintaining security and data confidentiality, following the statistical and epidemiological analysis plan, and performing quality-control checks of all programs (details on quality-control requirements are described in Section 9.10).

Security processes were in place to ensure the safety of all systems and data. Every effort was made to ensure that data were kept secure so that they could not be accessed by anyone except select study staff. Appropriate data storage and archiving procedures, with periodic backup of files, were followed.

9.8.2 Groupings of Quantitative Data

The following transformations were made in the analysis data:

- Age was categorised according to 10-year age groups except for the groups aged 40 to 54 years and 85 years or older; age was modelled as a continuous variable in the propensity score models (see 0, Table 35).
- The number of HCRU visits in the 180 days before the index date were grouped into data source–specific categories depending on the distribution of this variable for each data source (see 0, Table 35).
- Categories of BMI, alcohol use, HbA1c, and smoking status were not available in Medicare data nor in the full study population of the HIRD and therefore were not included. In CPRD data, categories were based on groupings specified in Read codes; in PHARMO data, for the subset of patients with GP data (about 40%), categories were based on groupings and codes specified by the Working Group on Coordinating Information Automation in the Netherlands [32], as follows:
  - Body mass index was either based on the most recent BMI value in the medical record or calculated as kg/m^2 from the height and weight values recorded closest to the index date within up to 5 years before the index date. Body mass index was then classified as follows: less than 20 (underweight); 20 to < 25 (normal weight); 25 to < 30 (overweight); 30 to < 40 (obese); 40 or higher (severely obese); and unknown/missing (see 0, Table 35).
  - Alcohol use was classified as follows: non-drinker; low-moderate intake (1 to 6 units/week); heavy or very heavy intake (≥ 7 units/week); drinker, unknown quantity; and unknown/missing (see 0, Table 35).
  - Smoking status was classified as follows: current smoker; former smoker; non-smoker; and unknown/missing (see 0, Table 35).
  - HbA1c was classified as follows: good control, < 7.0% (< 53 mmol/mol); borderline control, 7.0% to 10% (53-86 mmol/mol); poor control, > 10% (> 86 mmol/mol); and unknown (see 0, Table 33).
- The index therapy indicator variables were recoded for inclusion into propensity score models by recategorising “add-on” therapy as “add-on” or “add-on and switched-to” therapy, and “switched-to” therapy as “switched-to” or “add-on and switched-to” therapy (index therapy definitions can be found in Section 9.4.2.5.3). This recategorisation of the

original index therapy indicator variable was required to improve convergence issues encountered in some models.

- Calendar year (time period) of the index date was categorised in the final analysis, including in propensity score models, to account for variability in health care and prescribing practices during the coronavirus disease 2019 (COVID-19) era (see 0, Table 35).
  - For index dates occurring before 2020 (ie, from the start of the study period through 2019), the time period of the index date was categorised by calendar year (single category for each individual calendar year) for index dates occurring in the first year of the study period (except for CPRD, which combined index dates in 2012 and 2013 into one category since the study period covered only the last 2 months in 2012) through 2019.
  - For index dates occurring in 2020 or after, the January 2020 to February 2020 time period was categorised as a single category. Starting in March 2020 (ie, beginning of the "COVID-19 era"), single categories were generated for each 6-month period. During propensity score modelling, some categories in the COVID-19 era were collapsed, if needed, due to small numbers of events or to improve model convergence issues.

9.8.3 Handling of Small Cell Counts

The study investigators adhered to the reporting rules for small cell counts under the data protection policies applicable to each data source. CPRD requires suppression of any cells with frequency values of 1 to 4 or cells that allow a value of 1 to 4 to be derived from other reported cells. Although PHARMO usually does not allow reporting results for sample sizes of 1 to 5 patients, cell counts of 1 to 5 can be reported. For the final analysis, updated Carelon Research privacy rules (HIRD data) do not allow displaying cells with values of 1 to 4 or cells that allow a value of 1 to 4 to be derived from other reported cells. Medicare privacy rules do not allow displaying cells with frequency values of 1 to 10 or any cell that allows a frequency of 1 to 10 to be derived from other reported cells. Also, in Medicare, the reporting of minimum or maximum values, mode, and associated percentiles is not permitted. For all data sources, individual-level data were accessible to the investigators for conducting the analyses, irrespective of the sample size within certain strata.

9.9 Statistical Methods

Analyses were conducted separately for each data source according to the study protocol amended 06 June 2023. The planned analyses were detailed in the SAP, version 3.0, dated 22 April 2020, including the sensitivity analysis plan, dated 16 May 2024, and the pooled analysis plan, dated 20 November 2024, both appendences to the SAP (see list of stand-alone documents in Appendix A); updates are described in Section 9.9.9 of this report.

9.9.1 General Considerations

In accordance with the recommendations of the American Statistical Association [134], the International Committee for Medical Journal Editors [68], and expert opinion on the misuse of significance testing [5; 59; 89; 112], reliance on statistical significance to interpret study results was avoided. Instead of a dichotomous interpretation based on significance testing, this study relied on a quantitative interpretation that considers the magnitude, precision, and possible bias in the estimates derived and reported. This is a more appropriate approach than one that ascribes to chance any result that does not meet conventional criteria for statistical significance.

9.9.2 Main Summary Measures

For the descriptive analyses conducted to compare baseline covariates between dapagliflozin new users and comparator AD new users for each cancer outcome (see details in Section 9.9.3.1), categorical variables were summarised by frequencies and percentages, and continuous variables were summarised by means and standard deviations (SDs), medians, interquartile ranges (first quartile to third quartile), and minimum and maximum values. Descriptive statistics were presented in the overall cohorts and in the cohorts stratified by the presence or absence of concomitant insulin use at the index date (ie, insulin users and insulin non-users). Similarly, for sex-combined bladder cancer only, descriptive statistics were also stratified by pioglitazone use (ie, pioglitazone users and pioglitazone non-users). Descriptive statistics were calculated in each data source before and after propensity score trimming.

To assess the balance of measured baseline covariates between dapagliflozin and comparator AD index dates in the full samples (before propensity score trimming) and in the propensity score–trimmed samples, we used the absolute standardised difference (StDiff). The StDiff for continuous variables is the difference in the mean of the variable for dapagliflozin and comparator AD groups (all comparator ADs included in one group) divided by the pooled SD. For categorical variables, the StDiff is the absolute value of the difference in the prevalence of each level of the variable in the dapagliflozin and comparator AD groups divided by the pooled SD [10]. According to Austin [10], values of the StDiff of 0.2, 0.5, and 0.8 roughly correspond to small, medium, and large differences, respectively, in the level of the covariate between the treatment and comparator groups. Values of the StDiff are included in the results tables for each level of each covariate.

Using the propensity score–trimmed analysis samples (see details on propensity score trimming in Section 9.9.3.2), crude incidence rates and exact 95% CIs were calculated in each index exposure group and were expressed per 10,000 person-years (Section 9.9.3.3.1). Both crude incidence rates and propensity score–adjusted incidence rates were calculated for the two categories of exposure, dapagliflozin and comparator AD, for the overall cohorts; for the cohorts stratified by insulin use at the index date; and, for bladder cancer, stratified by sex and stratified by pioglitazone use at the index date. In CPRD and PHARMO, the only data sources

that contributed to this study patients aged < 65 years and patients aged ≥ 65 years, the propensity score–adjusted incidence rates were also calculated for subgroups of age at the index date (< 65 years and ≥ 65 years) (see Section 9.9.3.3.2). Crude IRRs and propensity score–adjusted IRRs, with corresponding 95% CIs, comparing each outcome in dapagliflozin new users and comparator AD new users, were also calculated (Section 9.9.3.4).

9.9.3 Main Statistical Methods

9.9.3.1 Descriptive Statistics

After applying the inclusion and exclusion criteria, descriptive analyses were conducted within each data source for each cancer outcome to (1) compare baseline characteristics of dapagliflozin new users and comparator AD new users, (2) compare HCRU measures amongst dapagliflozin new users and comparator AD new users in each cancer outcome cohort during the 180 days before the index date, and (3) characterise the index prescription of dapagliflozin amongst new users of dapagliflozin in the composite cancer cohorts by calendar year of the index date. In each data source, using the cohort of eligible patients for each cancer outcome (female breast cancer, female bladder cancer, male bladder cancer, sex-combined bladder cancer, female composite cancer, and male composite cancer), these analyses were conducted before and after propensity score trimming in (1) the overall cohorts; (2) the cohorts stratified by insulin use at the index date; and (3), for bladder cancer only, the cohorts stratified by pioglitazone use at the index date.

Additional descriptive analyses were conducted using the propensity score–trimmed analysis samples for each outcome to assess the distribution of key selected covariates overall and by insulin use at the index date; by propensity score stratum; and, for the bladder cancer cohort only, by pioglitazone use at the index date.

9.9.3.2 Propensity Score Analysis

Propensity score models were built to estimate the probability that each individual patient would receive dapagliflozin versus comparator AD. Within each data source, a propensity score model was built separately for each cohort (the overall cohort; the cohorts stratified by insulin use at the index date; and, for sex-combined bladder cancer, the cohorts stratified by pioglitazone use at the index date) for each cancer outcome (female breast cancer, male bladder cancer, female bladder cancer, sex-combined bladder cancer, male composite cancer, and female composite cancer). Table 7 lists each of the 20 cohorts for which a separate propensity score model was built.

Table 7 Independent Cohorts for Propensity Score Model Building (Including Covariate Selection and Propensity Score Estimation) and Propensity Score Trimming

Cohort No.[a]	Outcome cohort	Stratification	Additional covariates in base PS model[b]
1	Female breast cancer	Overall (unstratified)	Insulin use at the index date
2		Insulin use at the index date	None
3		No insulin use at the index date	None
4	Female bladder cancer	Overall (unstratified)	Insulin use at the index date; pioglitazone use at the index date
5		Insulin use at the index date	Pioglitazone use at the index date
6		No insulin use at the index date	Pioglitazone use at the index date
7	Male bladder cancer	Overall (unstratified)	Insulin use at the index date; pioglitazone use at the index date
8		Insulin use at the index date	Pioglitazone use at the index date
9		No insulin use at the index date	Pioglitazone use at the index date
10	Sex-combined bladder cancer	Overall (unstratified)	Sex; insulin use at the index date; pioglitazone use at the index date
11		Insulin use at the index date	Sex; pioglitazone use at the index date
12		No insulin use at the index date	Sex; pioglitazone use at the index date
13		Pioglitazone use at the index date	Sex; insulin at the index date
14		No pioglitazone use at the index date	Sex; insulin at the index date
15	Female composite cancer	Overall (unstratified)	Insulin use at the index date
16		Insulin use at the index date	None
17		No insulin use at the index date	None
18	Male composite cancer	Overall (unstratified)	Insulin use at the index date
19		Insulin use at the index date	None
20		No insulin use at the index date	None

[a] Cohort numbers are used in Table 9 in Section 9.9.5 to identify cohorts used for individual sensitivity analyses.

[b] The covariates forced into the base propensity score model for all of the cohorts were (1) age at the index date; (2) race/ethnicity (Medicare, the HIRD, and CPRD only); (3) time period of the index date; (4) duration of lookback period; (5) geographic region (PHARMO, the HIRD, and Medicare only); (6) primary care practice region (CPRD only); (7) indicator variable for add-on index therapy; (8) indicator variable for switched-to index therapy; (9) number of different AD classes prescribed/dispensed during various time windows before the index date; and (10) an indicator variable for whether GP data were available (PHARMO only).

No. = number; PS = propensity score.

9.9.3.2.1 SELECTION OF COVARIATES

For each data source, covariates included in the propensity score models were selected separately for each of the cohorts listed in Table 7 based on the following: (1) the covariate was known a priori from the literature or previous interim study results to be a strong

confounder of the association of dapagliflozin exposure with the specific cancer outcome, and/or (2) the magnitude of the change in the effect estimate (ie, HR) of the association of dapagliflozin exposure with the specific cancer outcome when adding the covariate as an independent variable. The details of the propensity score model building process are outlined below.

Covariates Included in the Base Propensity Score Model

First, for each propensity score model, within each data source, a base model was constructed by Cox proportional hazards regression using the outcome variable as the dependent variable and dapagliflozin as the exposure variable. Within the base model, a defined set of covariates were forced into the propensity score model and included the following, as available within each data source:

- Age at the index date (Section 9.8 and 0, Table 35)
- Race/ethnicity (Medicare, HIRD, and CPRD only) (0, Table 35)
- Time period of the index date (Section 9.8 and 0, Table 35)
- Duration of lookback period (0, Table 35)
- Geographic region (PHARMO, HIRD, and Medicare only) (0, Table 35)
- Primary care practice region (CPRD only) (0, Table 35)
- Indicator variable for add-on index therapy (Section 9.8)
- Indicator variable for switched-to index therapy (Section 9.8)
- Number of different AD classes prescribed/dispensed during various time windows before the index date (0, Table 34)
- Indicator variable for whether GP data were available for the patient (PHARMO only)
  - Since only a subset of patients in PHARMO have linkage to GP data (Section 9.5.2), for models in PHARMO data only, when assessing the inclusion of variables that might contain missing information attributed to the lack of linkage to GP data (eg, BMI, smoking status, alcohol use), the models contained an interaction term between the indicator variable for the covariate of interest and the inverse of the indicator variable for whether GP data were available (ie, 1 − indicator variable value for whether GP data were available) [125].

Additionally, within the overall cohorts (ie, not stratified by insulin use at the index date), an indicator variable for insulin use at the index date was included in the base model. In the sex-combined bladder cancer cohorts, sex was included in the base model, and, in the bladder cancer cohorts not stratified by pioglitazone use, an indicator variable for pioglitazone use at the index date was included in the base model. A summary of the additional variables included in the propensity score base model for each individual cohort is provided in Table 7. For each individual cohort, the base propensity score model provided an initial estimate of the HR for the association between dapagliflozin use and the outcome of interest.

Covariate Selection in Addition to Those in Base Propensity Score Model

For additional lifestyle, medical condition, medication, HCRU, and socioeconomic covariates not included in the base propensity score model, individual Cox proportional hazards regression models were built for each covariate of interest, containing the set of base variables in addition to the covariate of interest. A covariate not included in the base set of covariates was selected for inclusion in the final propensity score model for a given cohort if the inclusion of the covariate of interest in the Cox proportional hazards model resulted in a treatment-related HR that met at least one of two conditions when compared with the HR when including only the base set of covariates: (1) change in the absolute value of the HR estimate of more than 0.0045 or (2) change in the value of the HR estimate of more than 0.045%.

9.9.3.2.2 PROPENSITY SCORE ESTIMATION AND TRIMMING

Within each data source, once the covariates for inclusion were determined for each cohort-specific propensity score model, the propensity scores were calculated by fitting separate multivariate logistic regression models for each individual cohort (Table 7). For each propensity score model, the exposure as the dependent variable (0 = comparator AD new user and 1 = dapagliflozin new user) was regressed on the relevant base set of covariates for that model (Table 7) plus all other covariates that were identified as independent variables in the variable selection process.

For each cohort, the distribution of the propensity scores was plotted by treatment group to visually assess the magnitude of the overlap between the propensity score distributions for dapagliflozin new users and comparator AD new users (see Section 10.2.3.2); greater overlap suggests greater exchangeability between treatment groups. Next, for all cohorts, extreme values were trimmed from the distributions by excluding all patients below the 2.5th percentile value of the comparator AD–exposed distribution of propensity scores and all those greater than the 97.5th percentile of the dapagliflozin-exposed distribution of propensity scores. The proportion of patients excluded from the distributions and the frequency of key covariates for excluded patients were calculated.

9.9.3.2.3 PROPENSITY SCORE ASSESSMENT AND STRATIFICATION

For each cohort, the absolute StDiff values were calculated for all covariates and were compared before and after propensity score trimming in the specific cohort. The final cohort-specific analytic data set after propensity score trimming, ie, the "propensity score–trimmed analysis sample," was divided into strata defined by the percentiles of the propensity score distribution after propensity score trimming, ie, "propensity score strata." Each new user was assigned to one mutually exclusive propensity score stratum based on his/her individual propensity score.

Where possible, deciles of the propensity scores were used to define the propensity score strata, with use of fewer strata (no fewer than three), if needed, to accommodate a small number of outcome events. The number of propensity score strata was defined as the maximum (up to 10) for which the following two conditions were met: (1) the number of events in both exposure groups was at least equal to the number of propensity score strata and (2) at least one event occurred in each propensity score stratum (regardless of exposure group). Plots of the StDiff values were created to assess balance of key covariates across treatment groups within each propensity score stratum and examine the need for model refinement because of residual imbalance in covariates.

9.9.3.3 Incidence Analysis

Incidence analyses were conducted separately in each data source for each primary and secondary outcome using the propensity score–trimmed analysis samples (1) overall; (2) stratified by insulin use at the index date; and (3), for bladder cancer only, stratified by pioglitazone use at the index date. Incident events were identified using the provisional events identified by the electronic algorithms for each outcome. Time at risk started to accrue on the day following the first exposure to the study drugs that granted study entry (Section 9.4.2.3). In the main analysis, for each outcome cohort, the first cancer outcome event of interest occurring during the time at risk and before censoring was used to calculate the corresponding incidence rate.

9.9.3.3.1 Crude Incidence Rates

In each data source, using the propensity score–trimmed samples, the crude incidence rate of the cancer outcome of interest in each index exposure group was the number of incident cancer events divided by the total number of person-years at risk. In each independent cohort (Table 7), person-time was summed across all dapagliflozin new users and comparator AD new users in each exposure group (ie, amongst dapagliflozin new users and amongst comparator AD new users). Incidence rates were estimated per 10,000 person-years with exact 95% CIs calculated using the relationship between the Poisson distribution and chi-square distribution as described in Dobson et al [31].

Crude Incidence Rates by Propensity Score Stratum

Crude incidence rates were calculated separately for dapagliflozin-exposed person-time and person-time for comparator AD exposure. Unadjusted (crude) incidence rates were also estimated within each individual propensity score stratum using the stratum-specific counts of patients, person-time, and number of outcomes. These stratum-specific estimates were used to adjust the incidence rates as described in Section 9.9.3.3.2 and to compute IRRs adjusted for propensity scores (see Section 9.9.3.4.2).

Crude Incidence Rates by Cumulative Exposure/Treatment Duration

For each outcome in each data source—overall; stratified by insulin use at the index date; and, for bladder cancer only, by pioglitazone use at the index date—unadjusted incidence rates were estimated by categories of cumulative duration of dapagliflozin and comparator AD exposure during follow-up. Cumulative exposure was defined as the continuous treatment duration in each cohort (dapagliflozin and comparator AD). These measures were estimated by summing together the number of new cases of each study endpoint during the follow-up period in a given category of treatment duration and dividing it by the total person-time accumulated by individuals in the treatment duration category, where each individual contributed person-time to only one treatment category (see treatment duration definition in Section 9.4.2.4). No IRRs were estimated by cumulative exposure.

9.9.3.3.2 Propensity Score–Adjusted Incidence Rates

Within each data source, incidence rates for each cohort were standardised across the propensity score strata within each category defined by exposure group, sex (for the bladder cancer outcome only), and, when applicable, age (< 65 years and ≥ 65 years). For each exposure group, the stratum-specific incidence rate was calculated and standardised using the person-years in the dapagliflozin propensity score–trimmed cohort to estimate the standardised incidence rate (ie, the "propensity score–adjusted incidence rate") and variance.

The exact limits method, proposed by Dobson et al [31], was used to calculate the confidence limits for the propensity score–adjusted incidence rate.

9.9.3.4 Comparative Analysis

For each cancer outcome, and within each data source, IRR estimates were generated to compare the incidence rate in new users in the dapagliflozin exposure group with the incidence rate in new users in the comparator AD exposure group.

9.9.3.4.1 Crude Incidence Rate Ratios

Crude IRRs were calculated by dividing the crude (ie, not accounting for propensity score) incidence rates for each cancer outcome in dapagliflozin new users by the crude incidence rates in comparator AD new users. Corresponding 95% CIs were calculated using a Poisson distribution and test-based methods defined in Sahai and Khurshid [113].

9.9.3.4.2 Propensity Score–Adjusted Incidence Rate Ratios

To adjust for propensity score strata, IRRs with corresponding 95% CIs were generated using the Mantel-Haenszel approach outlined in Rothman K. J. et al [111] for each propensity score stratum. As appropriate in each data source, these methods were applied to generate the adjusted overall IRRs across the different strata defined by insulin use at the index date (yes/no or overall); age (< 65 years and ≥ 65 years); for the bladder cancer outcome only, by

sex and by pioglitazone use at the index date (yes/no); and propensity score decile or other stratum (eg, quintile, quartile, tertile).

9.9.3.4.3 Handling Cohorts With a Low Number of Events

Data source–specific policies for reporting low cell counts are described in Section 9.8.3.

The expected low or zero number of events in some cohorts prevented the calculation of meaningful IRRs. In these scenarios, the crude incidence rates were displayed but the IRRs (crude and adjusted) and the corresponding cells were suppressed in the tables. The IRRs were suppressed if *both* of the following conditions were true:

1 The number of events in at least one of the exposure groups (ie, dapagliflozin or comparator AD) was less than the number of defined propensity score strata. The rationale for this criterion was that if the number of events in each exposure group was equal to or higher than the number of defined propensity score strata, there was a higher probability that there was at least one event from each exposure group within each propensity score stratum.
2 There was not at least one event (regardless of exposure group) in each propensity score stratum.

9.9.3.5 Analyses of Time-Varying Variables

Changes in diabetes severity and changes in AD treatment were considered time-varying variables in this study. Under the assumption that the exposure groups are exchangeable at baseline (owing to the implementation of propensity scores), differences between the two exposure groups in changes in diabetes severity and changes in AD treatment were assumed to be linked to dapagliflozin or comparator AD exposure. Analyses focused on describing the effects of dapagliflozin on changes in diabetes severity and for treatment changes over time to characterise the effects of the drug on patients' evolution.

For each outcome, descriptive analyses were conducted using the propensity score–trimmed analysis samples overall (ie, not stratified by insulin use at the index date) and stratified by insulin use at the index date for all propensity score strata combined.

9.9.3.5.1 Diabetes Severity

The diabetes severity score was computed at baseline and during follow-up, at fixed 90-day time intervals (see Section 9.4.4). At each time point, the score was defined as the maximum severity score achieved during all available time up to and including those dates. The DCSI is mostly composed of chronic conditions that either stabilise or worsen over time, ie, once a patient has a record of a diagnosis for a medical condition from the list of the components of

the DCSI, this condition will persist. Therefore, the score would be expected to increase monotonically.

In each data source, the summary statistics in Table 8 were calculated for the outcome cohorts for female breast cancer, sex-combined bladder cancer, female composite cancer, and male composite cancer, in the overall cohorts and in the insulin use–stratified cohorts.

Table 8 Summary Statistics Calculated to Describe Diabetes Severity Over Follow-up in Each Data Source

Group summary statistic	Corresponding feature of box-and-whiskers plot
Maximum	Endpoint of upper whisker
Third quartile (75th percentile)	Upper edge of box
Median (50th percentile)	Line inside box
Mean	Symbol marker
First quartile (25th percentile)	Lower edge of box
Minimum	Endpoint of lower whisker

Note: The Medicare data use agreement terms do not allow for median values, minimum and maximum values, and percentiles to be calculated; therefore, in Medicare, only the mean, standard deviation, and interquartile range were calculated.

9.9.3.5.2 ANTIDIABETIC DRUG TREATMENT CHANGES

An indicator variable was derived for each of the AD classes defined in Section 9.4.4.2. To derive this variable for a drug class, an indicator variable was created for the current use of any AD within the class every 90 days during follow-up (details on this analysis are described in the SAP, version 3.0, dated 22 April 2020).

9.9.3.6 HCRU During Follow-up: Secondary Outcome

For the female breast cancer and the bladder cancer outcomes, descriptive analyses were conducted to compare HCRU measures (listed below) amongst new users of dapagliflozin and new users of comparator ADs within each data source, overall and stratified by insulin use at the index date. Measures of HCRU were estimated as rates over follow-up time and as the number of encounters or procedures for each year of follow-up, to assess temporal trends in medical surveillance. Data for the specified HCRU measures were not consistently captured across all data sources. Each of the following HCRU measures were categorised either as quartiles or as 0 vs ≥ 1, depending on the frequency distribution in each data source.

- Outpatient visits to GPs, gynaecologists (amongst female patients), urologists, and any other specialists; in the HIRD, *all* outpatient encounters including laboratory tests
- Inpatient hospitalisation (excluding hospital outpatient visits and emergency department visits that did not result in an inpatient stay)

- Mammograms (amongst females)
- Biopsies
  - Breast (amongst females)
  - Bladder
- Cystoscopies
- Urinary cytologies

In PHARMO, outpatient visits, gynaecologist visits, and visits to specialists were based on patients with GP data available (about 40% of patients). Data on mammograms are not presented for PHARMO because only data on inpatient mammograms were available. Neither cystoscopies nor urinary cytologies were available for PHARMO because cystoscopies are performed in ambulatory care (data were not available), and urinary cytologies are based only on hospital procedures.

9.9.3.7 Mediation Analyses

Mediation analyses were planned to be conducted if a positive total effect of dapagliflozin on the incidence of female breast cancer or bladder cancer was detected in the overall cohort or analyses stratified by insulin or by pioglitazone use (for bladder cancer only). Use of health care resources was considered a potential mediator. The aim was to decompose the total effect into the natural direct effect and the natural indirect effect of dapagliflozin on breast and bladder cancer incidence [129]. The criterion driving the conduct of this analysis was a lower bound of the 95% CI for the propensity score–adjusted IRR greater than 1.5, a threshold selected to focus these analyses on potential strong associations between dapagliflozin and the cancer outcome.

9.9.4 Missing Values

The approach to addressing missing values varied by variable and data source. In CPRD, when information on medication exposure (ie, dose frequency per day, duration of exposure, and daily dose of dapagliflozin) was not directly reported or could not be derived based on other variables, values were imputed as the mode for prescriptions of products with the same strength of dapagliflozin that were prescribed during the same calendar year. When determining concomitant insulin and pioglitazone use at the index date, a default value of 30 days was used for insulin prescriptions with missing days' supply. The frequency of missing data values for other variables, including BMI, smoking status, alcohol use, and HbA1c, is reported in the descriptive tables.

In PHARMO, information on BMI, smoking status, alcohol use, HbA1c and specialty care visits are captured for the subset of patients with linked GP data. Therefore, the frequency of "unknown" typically means "unknown due to no GP data available" or "unknown/not recorded by the GP," with this information reported for each of the variables. The percentage

of patients with GP data was reported. The GP subset is assumed to be a representative sample of the overall cohorts (not stratified by insulin use at the index date or by pioglitazone use at the index date) of all patients and is therefore considered to represent missing data occurring at random.

In CPRD, the HIRD, and Medicare, lack of a diagnosis code or drug prescription or dispensing record in each respective data source was assumed to indicate absence of the coded condition or drug prescription or dispensing. Laboratory values and variables such as smoking status, alcohol consumption levels, and BMI are unavailable in Medicare and have high levels of missing values in the HIRD; therefore, these variables were not included in the analyses for these data sources.

9.9.5 Sensitivity Analyses

The planned sensitivity analyses to be conducted as part of the final set of analyses were outlined in Section 2.8.7 of SAP version 3.0, dated 22 April 2020 (see list of stand-alone documents in Appendix A). A sensitivity analyses plan dated 16 May 2024 and appended to the SAP described the approach and methods for each of the sensitivity analyses conducted as part of the final analysis and documented any updates to the initially planned sensitivity analyses (eg, planned analyses that would not be conducted, additional analyses) since the SAP was written; amendments to the initially planned sensitivity analyses are summarised in Section 9.9.9, Table 10.

The sensitivity analyses were planned to provide further information on the robustness of the propensity score–adjusted IRRs estimated for the primary outcomes in the main analysis. Table 9 provides a summary of all the sensitivity analyses conducted for the final (120-month) analysis; prespecified conditions for conducting each sensitivity analysis, if applicable; and the cohorts included in each sensitivity analysis. Unless otherwise specified in the description of each sensitivity analysis in Table 9, all planned sensitivity analyses were conducted in the primary outcome cohorts (female breast cancer, female bladder cancer, male bladder cancer, sex-combined bladder cancer), both in the overall cohorts and in the cohorts stratified by insulin use at the index date. Sensitivity analyses were not conducted on the pioglitazone use–stratified bladder cancer cohorts. The sensitivity analyses were conducted using the propensity score–trimmed analysis samples, unless otherwise stated in Table 9. As most of the sensitivity analyses focused on the impact of censoring, a descriptive analysis was conducted before implementing each of the sensitivity analyses in order to describe the frequency of the occurrence of each of the censoring criteria (outlined in Section 9.3.2) in the primary cancer outcome cohorts.

Table 9 Summary of Sensitivity Analyses for Final (120-Month) Analysis

Description of sensitivity analysis	Condition for performing analysis	Cohorts considered for inclusion	New PS models (yes or no)
Time-to-event analyses using the Kaplan-Meier estimator	None	PS-trimmed analysis sample[a] Primary outcome cohorts[b], overall[c] and stratified by insulin use (cohorts 1-12 in Table 7)	No
Potential impact of unmeasured confounding	Performed on specific primary outcome cohorts for which a pooled IRR was estimated	Overall[c] pooled IRRs[d] Primary outcome cohorts[b], overall[c] only (cohorts 1, 4, 7, and 10 in Table 7)	Not applicable
Removing various censoring criteria			
First occurrence of each type of malignancy	Performed in a given primary outcome cohort if any users were censored due to having a cancer diagnosis (other than the cancer outcome of interest for that specific cohort)	PS-trimmed analysis sample[a] Primary outcome cohorts[b], overall[c] and stratified by insulin use (cohorts 1-12 in Table 7)	No
Removal of censoring of follow-up at non-dapagliflozin SGLT2 inhibitor initiation	None	PS-trimmed analysis sample[a] Primary outcome cohorts[b], overall[c] and stratified by insulin use (cohorts 1-12 in Table 7)	No
Removal of censoring of follow-up at dapagliflozin initiation for new users of comparator AD	Performed in primary outcome cohorts with an observed increased risk for the outcome of interest amongst new users of dapagliflozin in the main analysis[e]	PS-trimmed analysis sample[a] Primary outcome cohorts[b], overall[c] and stratified by insulin use (cohorts 1-12 in Table 7)	No
Evaluation of deaths (all-cause mortality)	PS-adjusted mortality rates and mortality rate ratios estimated in all primary outcome cohorts If an increased risk of mortality was observed in the dapagliflozin group in a given cohort, cancer risk was re-assessed using a competing risk approach in the respective cohort	PS-trimmed analysis sample[a] Primary outcome cohorts[b], overall[c] and stratified by insulin use (cohorts 1-12 in Table 7)	Yes, new PS models were generated accounting for mortality as the outcome of interest

[a] PS-trimmed analysis sample = cohort after PS trimming was applied in the main analysis.

[b] Primary outcome cohorts consisted of female breast cancer, female bladder cancer, male bladder cancer, and sex-combined bladder cancer.

[c] Overall cohort = cohort not stratified by insulin use at the index date.

[d] The pooled IRR is the IRR estimated across all data sources (see Section 9.9.7).

[e] Main analysis = analyses conducted as part of the main study (ie, all analyses other than the sensitivity analyses)

PS = propensity score.

9.9.5.1 Time-to-Event Analysis

Temporal trends in cancer risk amongst new users of dapagliflozin and new users of a comparator AD were assessed by estimating the risk of cancer over follow-up using the KM estimator. The propensity score calculated in the main analysis was used to generate stabilised inverse probability of treatment weights (IPTW) [61], which were included in the KM estimator [138]. The censoring criteria were the same as used for the main analysis. The distribution of stabilised weights was checked, specifically the mean of stabilised weights, to ensure they were close to 1 [12]. A visual display of cancer risk over follow-up for each exposure group was ascertained by plotting the 1 – KM estimates (ie, cumulative incidence).

9.9.5.2 Potential Impact of Unmeasured Confounding

Using the method described by Lash et al [76], quantitative bias analysis was used to evaluate the effect of potential unmeasured confounding (eg, because of unmeasured smoking for bladder cancer, body mass index for female breast cancer, or other unmeasured variables) on the association between dapagliflozin use and each of the primary cancer outcomes using the observed IRRs pooled across the data sources (Section 9.9.7). This sensitivity analysis was conducted for each of the primary outcomes for which a pooled IRR was estimated (Section 9.9.7). Three bias parameters were defined for this analysis: (1) the expected association between a hypothetical unmeasured confounder and the outcome, (2) the prevalence of the hypothetical unmeasured confounder amongst those exposed to dapagliflozin, and (3) the prevalence of the hypothetical unmeasured confounder amongst those without dapagliflozin exposure. These three bias parameters were entered into a formula by Lash et al [76] to calculate the overall pooled association between dapagliflozin and the outcome of interest expected after adjustment for the hypothetical unmeasured confounder.

The IRRs were estimated under multiple variations of two different scenarios: (1) ranging the association of the unmeasured confounder with the outcome from a RR of 1.5 to 4.5 and (2) varying the difference in prevalence of the unmeasured confounder in each treatment group from −100% (ie, present in every patient in the comparator AD group but not present in any patient in the dapagliflozin group) to 100% (ie, present in every patient in the dapagliflozin group but not present in any patient in the comparator AD group). The resulting IRRs for the overall pooled analysis were plotted under these varying assumptions of unmeasured confounding compared with the observed IRR estimate from the pooled analysis.

9.9.5.3 Removal of Various Censoring Criteria

9.9.5.3.1 First Occurrence of Each Type of Malignancy

In this analysis, the censoring criterion of diagnosis of any study cancer and any invasive cancer other than the cancer outcome of interest was removed. Patients were followed for the first occurrence of *each* specific primary cancer endpoint regardless of a prior occurrence of another cancer after the index date; this allowed for patients to have more than one

malignancy and to evaluate any potential impact of censoring at the incidence of a cancer diagnosis that was not the cancer outcome under study on the IRRs. This analysis was done in a stepwise approach. First, for each propensity score–trimmed analysis sample, the number of new users in the main analysis who were censored due to a cancer diagnosis that was not the cancer diagnosis of interest was determined. Next, if there were no users censored in the main analysis for this reason, then the sensitivity analysis was not conducted in that cohort (as the time at risk would be unchanged from the main analysis). If at least one patient was censored due to a cancer other than the cancer of interest, the sensitivity analysis removing the censoring criterion was conducted in that specific cohort. All other censoring criteria were retained, and propensity score–adjusted incidence rates and IRRs were calculated using the same methods as in the main analysis. If diagnoses for breast cancer and bladder cancer were recorded for a patient on the same date, both cancers were counted in each respective cancer outcome cohort.

9.9.5.3.2 Removal of Censoring of Follow-up at Initiation of a Non-Dapagliflozin SGLT2 Inhibitor

This analysis removed the censoring criterion of initiating a non-dapagliflozin SGLT2 inhibitor in either the dapagliflozin or the comparator AD group, whilst retaining all other censoring criteria that were used in the main analysis. Propensity score–adjusted incidence rates and IRRs were estimated for each cohort using the same methods used in the main analysis.

9.9.5.3.3 Removal of Censoring of Follow-up at Dapagliflozin Initiation for New Users of Comparator ADs

This analysis removed the censoring criterion of dapagliflozin initiation within the comparator AD group, whilst retaining all other censoring criteria used in the main analysis. If a patient's index medication was a comparator AD but later switched to dapagliflozin, that person then continued to contribute follow-up time to the comparator AD group and started to contribute follow-up time to the dapagliflozin group after initiating dapagliflozin, ie, contributed follow-up time to both the comparator AD group and the dapagliflozin group simultaneously, until the earliest occurrence of the another censoring criterion or the cancer outcome event of interest. This analysis was planned to be conducted only for cohorts in which an increased risk was observed amongst new users of dapagliflozin in the main analysis. An "increased risk" for a given cohort was defined as the lower bound of the 95% confidence limit of the propensity score–adjusted IRR being higher than 1.5 in the main analysis for the cancer outcome of interest. Propensity score–adjusted incidence rates and IRRs were estimated using the same methods described in the main analysis.

9.9.5.4 Evaluation of Deaths (All-Cause Mortality)

The goal of this analysis was to explore the effect of potential differences in the risk of death (all-cause mortality) between new users in the dapagliflozin group and those in the comparator AD group. To estimate all-cause mortality, this analysis counted death as the event of interest, removed the censoring criterion of the occurrence of any cancer diagnosis, and did not censor at the occurrence of the cancer outcome of interest; all other censoring criteria used in the main analysis were retained in this analysis.

In CPRD, deaths were identified using information on date of death derived first from available vital records (UK Office for National Statistics), then from in-hospital deaths (HES), and, third, from GP records (GOLD); CPRD GOLD captures > 98% of deaths [54]. In PHARMO, deaths were identified using death dates from the outpatient pharmacy data (linked through the Dutch Central Bureau of Genealogy) and GP data. In the HIRD, mortality was defined using a composite algorithm that draws from several sources including the Social Security Death Master File, online obituaries, and other sources, and has been shown to have > 85% sensitivity and specificity compared with the gold standard of the National Death Index [70]. In Medicare, deaths were identified using the date of death variable, which is ascertained by CMS using Medicare claims, online date of death edits submitted by family members, and benefit information used to administer the Medicare programme collected from agencies such as the Social Security Administration; 99% of death dates in the Medicare database are validated [109].

To estimate propensity score–adjusted mortality rates amongst new users of dapagliflozin and new users of comparator AD and to estimate propensity score–adjusted mortality rate ratios, the first step was to generate new propensity score models and implement propensity score trimming in each of the cohorts using the methods outlined in Section 9.9.3.2. These new propensity score models accounted for death as the outcome (as opposed to the cancer of interest), and variables were included based on their association with the outcome of death. Instead of stratifying propensity scores as was done for the main analysis, confounding adjustment was done using stabilised IPTW [61]. Adjusted mortality rates and mortality rate ratios and 95% CIs were estimated using a Poisson regression model including the weights described above, with robust estimation of the variance (to account for the weights) [145]. The model had the outcome as the dependent variable, the natural logarithm of the person-years at risk as the offset, and cohort as the exposure of interest (dapagliflozin vs comparator AD).

It was planned that if an increased mortality risk was observed in a primary outcome cohort (ie, the lower bound of the 95% CI of the propensity score–adjusted mortality rate ratio for a given cohort was higher than 1.5), then an additional competing risk analysis to re-assess cancer risk in the respective cohort, ie, accounting for death as a competing risk of cancer, would be considered as described in the sensitivity analyses plan [143].

9.9.6 Post Hoc Sensitivity Analysis (HIRD)

Given the recent approval of additional non-T2DM indications for dapagliflozin and new medications within the comparator AD group, an analysis was undertaken to examine the prevalence of a recorded T2DM diagnosis on or before the index date in both exposure groups. In this final (120-month) analysis, it was observed that, in the female composite cancer cohort in the HIRD, there was a lower prevalence of having a recorded T2DM diagnosis before or on the index date amongst those in the comparator AD group (63%) than in the dapagliflozin group (85%) (Appendix J, FemaleCompositeCa Table 2 HIRD). This finding persisted in the HIRD in the overall propensity score–trimmed samples (which did not adjust for T2DM). When exploring further in the female composite cancer cohort in the HIRD, codes[2] indicating an obesity diagnosis in the 6 months before or on the index date were more common in new users of comparator AD (33.4%) than in new users of dapagliflozin (24.0%). Detailed descriptive results of the prevalence of a recorded T2DM diagnosis and obesity over the study period in the HIRD are described in Section 10.5.2.6. Because new users of GLP-1 receptor agonists make up a portion of the comparator AD group, and GLP-1 receptor agonists were recently approved for chronic weight management in the US in individuals with obesity or who are overweight [95], it was hypothesised that the imbalance in having a recorded T2DM diagnosis between the exposure groups may indicate potential uncontrolled confounding by obesity or T2DM, ie, a higher proportion of comparator AD new users may have been taking the medication for a weight loss indication, as opposed to taking the medication for T2DM. Thus, users in the comparator AD group may be less likely to have T2DM and instead use ADs for weight management than users in the dapagliflozin group (where patients may be more likely to be taking the medication for T2DM). This observation was primarily seen in the HIRD and not in Medicare, likely because, during the study period, US federal law prohibited Medicare coverage of medications for weight management indications, making it less likely for the Medicare database to include study drugs prescribed solely for the purpose of weight management [73].

This potential confounding by T2DM and obesity was a particular concern for the female breast cancer outcome given that obesity is a risk factor for breast cancer, particularly amongst postmenopausal women [6; 29; 103; 108]. Therefore, a post hoc sensitivity analysis was undertaken in the HIRD female breast cancer cohort in which the propensity score was re-calculated using a base model that included the T2DM diagnosis variable to improve the balance of this variable across the exposure groups (T2DM was not considered as a potential covariate in the propensity score models in the main analysis, as described in Section 9.4.3). Obesity was not included as a covariate in this sensitivity analysis nor in the main analysis (Section 9.4.3) because diagnosis codes for obesity are underreported in claims [4; 118]. The

[2] Codes indicating an obesity diagnosis included ICD-9-CM codes (V85.3*, V85.4*, 278.00, 278.01, 278.03) and ICD-10-CM codes (Z68.3*, Z68.4*, E66.0*, E66.1, E66.2, E66.8*, E66.9).

remaining steps for calculating the propensity score, trimming, and assessment of covariate balance were repeated as described in Section 9.9.3.2. Using this new propensity score–trimmed analysis sample, propensity score–adjusted incidence rates and IRRs were calculated to examine differences in female breast cancer incidence between new users of dapagliflozin or of comparator AD, using methods described in Sections 9.9.3.3.2 and 9.9.3.4.2. Findings were descriptively compared with those obtained in the main analysis for female breast cancer in the HIRD data source to assess whether any changes in effect estimates were observed with the inclusion of the T2DM variable in the propensity score model.

9.9.7 Pooled Analysis

For each primary outcome (female breast cancer, male bladder cancer, female bladder cancer, and sex-combined bladder cancer), propensity score–adjusted IRR estimates were planned to be pooled across the data sources according to the pooled analysis plan, version 1.0, 20 November 2024, an appendix to the SAP. Pooling of the propensity score–adjusted IRRs was considered separately for each primary outcome, in the overall cohort and in the cohorts stratified by insulin use at the index date. Only those data sources in which a propensity score–adjusted IRR was estimated were considered eligible for inclusion in the pooled analysis for a given outcome and cohort.

For cohorts in which calculating a pooled IRR was determined to be appropriate, within each qualifying data source, outcome event counts and person-time were aggregated by exposure category and propensity score stratum. Using Mantel-Haenszel methods [111], the stratum-specific estimates for all data sources were then used to estimate a pooled adjusted IRR.

This study used two approaches to assess statistical heterogeneity across the data source–specific propensity score–adjusted IRRs, and thus the appropriateness of pooling estimates across the data sources for a given outcome and cohort: (1) visual inspection of the overlap across the 95% CIs of the data source–specific propensity score–adjusted IRRs and (2) the I^2 index, which describes the percentage of total variability between the individual estimates that is due to heterogeneity (rather than chance) [30; 63]. Specifically, the I^2 index was used to assess statistical heterogeneity between the propensity score stratum-specific IRR estimates across all data sources (ie, amount of "between-stratum variation"). The more heterogeneity observed between the propensity score stratum-specific IRR estimates, the less credible the pooled IRR. In this study, if there was visually good overlap across the data source–specific propensity score–adjusted IRR CIs and the calculated I^2 index point estimate was below 50%, then the pooled Mantel-Haenszel–adjusted IRR was reported; otherwise, only the data source–specific propensity score–adjusted IRR estimates were reported for a given outcome and cohort, as described in Section 9.9.3.4.

9.9.8 Assessment of the Potential Impact of Differential Outcome Misclassification

For the data sources where validation of primary cancer outcomes was performed and PPVs were estimated during validation (ie, CPRD, the HIRD, and Medicare), a quantitative bias analysis was conducted to evaluate the potential impact of differential outcome misclassification between the dapagliflozin and comparator AD groups on the propensity score–adjusted IRR results for female breast cancer and sex-combined bladder cancer in each data source.

In a simulation analysis, for each outcome of interest, a "corrected" adjusted IRR was estimated under a hypothetical worst-case scenario of differential outcome misclassification [19]. The hypothetical worst-case scenario assumed perfect outcome classification in the dapagliflozin group (hypothetical $PPV_{dapagliflozin}$ = 100%) with a corresponding degree of outcome misclassification existing within the comparator AD group based on the observed PPV (under this hypothetical scenario, the corresponding $PPV_{comparator}$ is lower than the observed PPV estimated between the two exposure groups, ranging from 79.2% to 98.8% in the female breast cancer cohorts and from 86.7% to 97.7% in the sex-combined bladder cancer cohorts). The corrected adjusted IRR point estimate was then assessed to determine whether the interpretation of the primary study results may change under the worst-case scenario of differential misclassification.

For each outcome of interest (female breast cancer and sex-combined bladder cancer), inputs to the analysis were the data source–specific observed overall adjusted IRRs from the main analysis, a spectrum of PPV values for the outcome of interest associated with dapagliflozin (ranging from 0 to 100%), and a spectrum of PPV values for the outcome of interest associated with the comparator AD group.

9.9.9 Amendments to the Statistical Analysis Plan

Amendments to prior SAP versions 1.0 to 3.0 are outlined in previous interim reports. Additional analysis updates to SAP version 3.0 since the fourth interim (96-Month) report implemented in the final (120-month) analysis are outlined in Table 10.

Table 10 Analysis Updates Implemented in the Final (120-Month) Analysis

Section of SAP	Update	Reason
Section 2.3.3, Inclusion and Exclusion Criteria	Bexagliflozin and the combination drug of empagliflozin, linagliptin, and metformin were included as SGLT2 inhibitor medications in the final analysis.	New SGLT2 inhibitor medications that became available in the study countries between the fourth interim analysis through October 2023 were accounted for in the final analysis.

Section of SAP	Update	Reason
Section 2.6.1, Demographic and Lifestyle Variables	Race/ethnicity was included as a covariate only for Medicare in previous interim analyses. For the final analysis, race/ethnicity was also described as a covariate for CPRD and the HIRD.	Information on race/ethnicity became available in the CPRD and HIRD data sources in recent years; therefore, race/ethnicity was added as a covariate in these data sources for the final analysis.
Section 2.6.1, Covariates, Demographic and Lifestyle Variables	The variable "calendar year of index date" was modified such that all time before 2020 was modelled as categories, each with a discreet year; calendar time starting in 2020 was modelled more granularly such that January thru February 2020 was one category and, beginning in March 2020 (start of the COVID-19 era), categories comprised 6-month time increments to the end of the study period. The covariate was re-labelled "time period of the index date."	The covariate was modelled in the final analysis to account for variability in healthcare practice (eg, frequency of healthcare visits, medication prescribing patterns) during the COVID-19 era.
Section 2.6.2, Covariates, Medical Conditions	Heart failure and chronic kidney disease were added to the medical conditions covariates list to include in the descriptive analyses and to consider as potential covariates in the propensity score models.	Dapagliflozin was approved in the US and the EU for the treatment of heart failure in 2020 and approved for the treatment of chronic kidney disease in 2021. The study periods for all data sources overlap with the years these new indications were approved. The use of dapagliflozin for these new indications was accounted for in the final analysis by describing the frequency of patients with a recorded diagnosis of these conditions before or on the index date and considering these 2 medical conditions as potential covariates in the propensity score models.
Section 2.6.2, Covariates, Medical Conditions	T2DM was added to the medical covariates list to include in the descriptive analyses for the female composite cancer cohort and male composite cancer cohort.	Given that dapagliflozin was approved for indications other than T2DM in 2020 and 2021, the frequency of patients with a recorded T2DM diagnosis before or on the index date was added to the analysis for descriptive purposes.
Section 2.7, Study size and statistical precision	To refine the approach for calculating study size, the projected number of new users of dapagliflozin and precision estimates (ie, the upper limit of the IRR 95% CI) were updated based on the number of new users of dapagliflozin in each data source observed at the interim report 4 (96-month) analysis.	To reflect lower-than-expected dapagliflozin uptake and to more accurately provide study size estimates that would provide sufficient precision at the time of the final (120-month) data analysis.

Section of SAP	Update	Reason
Section 2.8.2, Propensity Score Development	The covariates included in the base propensity score models were updated as follows: the variable "sex" was included in the base propensity score model for all sex-combined bladder cancer cohorts; the indicator variable for "pioglitazone use on the index date" was included in the base propensity score model for all bladder cancer cohorts except for those in the pioglitazone use–stratified analyses; the "race/ethnicity" variable was included in all base propensity score models in CPRD and the HIRD.	The covariates included in the base models were updated to improve confounder adjustment; the same covariates were included in the base models across the data sources to ensure consistency across the data sources.
Section 2.8.2, Propensity Score Development	In any cohort, when the number of covariates was high relative to the number of dapagliflozin users during propensity score modelling, fewer covariates were selected into the propensity score model of the respective cohort. In this situation, covariates with the highest impact on the HR estimate of dapagliflozin exposure were prioritised to be included in the propensity score model.	To address potential propensity score model convergence issues in cohorts with a small number of dapagliflozin initiators (eg, insulin use cohorts, pioglitazone use cohorts).
Section 2.8.2, Propensity Score Development	For a given cohort, if substantial imbalance remained in the distribution of covariates across the exposure groups after propensity score modelling, trimming, and stratification, then the inclusion of higher order terms or interaction terms could be considered for the propensity score model and/or a larger proportion of the extreme values of the propensity score distribution could be considered for trimming.	To improve covariate balance and confounder control across the dapagliflozin and comparator AD groups.

Section of SAP	Update	Reason
Section 2.8.7, Sensitivity Analyses	The detailed approach and methods for the sensitivity analyses to be conducted in the final analysis were outlined in a separate document, the sensitivity analysis plan (version 1.0 dated 16 May 2024). The sensitivity analysis plan serves as an appendix to the SAP. The sensitivity analyses and methods conducted in the final analysis are summarised in Section 9.9.5 of the Final Report.	One planned sensitivity analysis has been conducted in previous interim analyses (ie, removing stratification by insulin use at the index date); all other sensitivity analyses described in the protocol and SAP were planned to be conducted only during the final analysis. The feasibility and approach for each sensitivity analysis was informed by the results from the fourth interim analyses.
Sensitivity Analysis Plan, version 1.0 (SAP appendix)	The planned sensitivity analyses assessing the temporal relationship of dapagliflozin exposure and cancer risk by (1) stratifying risk by time since first exposure (surveillance bias), (2) restricting the follow-up time by applying different time windows after cohort entry (lag time analysis), and (3) evaluating changes in cancer risk over time after the discontinuation of dapagliflozin (time since treatment discontinuation analysis), as described in SAP Section 2.8.7, were not performed in the final analysis.	The results from the fourth interim (96-month) analysis did not suggest an increased risk for any cancer outcome linked with dapagliflozin exposure. With failure to observe an increased risk of cancer associated with dapagliflozin use, the results of analyses of surveillance bias, lag time, and time since treatment discontinuation, as described in SAP Section 2.8.7, would not be meaningful as these analyses were intended to better understand a potential increased risk of cancer amongst dapagliflozin initiators.
Sensitivity Analysis Plan, version 1.0 (SAP appendix)	Time-to-event analysis using the KM estimator was implemented as an alternative approach for assessing temporal trends in cancer risk over follow-up across the exposure groups.	To describe the temporal trends in cancer risk over follow-up for each exposure group, cancer risk was estimated over time via the KM estimator, which, unlike the IRR, does not assume that risk is constant over follow-up. The resulting KM cumulative incidence plots also allow for a visual assessment of potential lag time or surveillance bias.
Sensitivity Analysis Plan, version 1.0 (SAP appendix)	The planned sensitivity analysis stratifying by insulin type (long-acting vs short-acting) at the index date amongst insulin users was not performed in the final analysis.	Due to the limited number of insulin users observed throughout the study across all data sources (at the fourth [96-month] analysis, only 10% to 25% of new users in each exposure group used insulin at the index date across all outcome cohorts), the insulin subgroup analyses were not feasible.

Section of SAP	Update	Reason
Sensitivity Analysis Plan, version 1.0 (SAP appendix)	Conditions were specified for performing several of the sensitivity analyses, including assessing the potential impact of unmeasured confounding; estimating IRRs using the first occurrence of each type of primary malignancy (regardless of whether another cancer was diagnosed earlier in follow-up); estimating IRRs upon removal of censoring of follow-up at the time dapagliflozin was initiated for patients in the comparator AD group.	The conditions set for performing specific sensitivity analyses were based on the feasibility of performing the specific sensitivity analysis or whether the results of the specific sensitivity analysis would be informative in the context of the results from the main analysis.
Sensitivity Analysis Plan, version 1.0 (SAP appendix)	For the sensitivity analysis evaluating deaths, if an increased risk of mortality was observed in the dapagliflozin group in a specific cohort, cancer risk would be re-assessed using a competing risk approach in the respective cohort.	Competing risk analysis would re-assess cancer risk accounting for death as a competing risk of the specific cancer outcome in the dapagliflozin group.

9.10 Quality Control

The standard operating procedures of each research centre were used to guide the conduct of the study. The minimum quality-control requirements for each step of the analysis are summarised in Table 11.

Table 11 Minimum Quality-Control Requirements for Analytic Programming

Activity	Quality-control procedure
Selection of cohorts: identification of patients (inclusion/exclusion criteria, exposure to study drugs)	Review of code and output by a second analyst
Selection of outcomes	Review of code and output by a second analyst
Definition of covariables	Review of code and output by a second analyst
Creation of the analysis data sets	Review of code and output by a second analyst
Analysis programs including descriptive and comparative analyses	Review of code by a second analyst
Generation of tables	Spot-check of output tables by epidemiologist and a second analyst/researcher

10. RESULTS

10.1 Participants

The cohort selection process for all outcome cohorts (ie, breast cancer, bladder cancer, female composite cancer, and male composite cancer), for the overall study population and by index exposure, is detailed in Figure 7 and Figure 8, as well as Appendix J, Table 0, for each data source.

During the study period, there were 501,398 potential new-use dates (dapagliflozin or an eligible comparator AD) in CPRD, 87,713 in PHARMO, 4,919,473 in the HIRD, and 8,708,661in Medicare. After applying the eligibility criteria common to all cancer outcomes and selecting the first comparator AD new-use date, the final number of eligible new users (dapagliflozin or comparator AD) was 102,743 in CPRD, 47,320 in PHARMO, 684,162 in the HIRD, and 1,249,555 in Medicare. The proportion of new users meeting the inclusion criteria but excluded for meeting the definition of T1DM was 2.9% in CPRD, 0.9% in PHARMO, 5.1% in the HIRD, and 7.7% in Medicare. The proportion excluded due to use of another SGLT2 inhibitor medication on or before the index date was 24.4% in CPRD, 3.9% in PHARMO, 28.4% in the HIRD, and 25.6% in Medicare.

The final number of eligible new users for each data source and cancer outcome cohort—for the dapagliflozin and comparator AD exposure groups before propensity score trimming (ie, the full samples)—is presented in Figure 7. Across all data sources, the full samples included a total of 80,707 dapagliflozin new users and 999,103 comparator AD new users for the female breast cancer cohort and 182,050 dapagliflozin new users (79,025 females and 103,025 males) and 1,839,218 comparator AD new users (974,493 females and 864,725 males) for the sex-combined bladder cancer cohort. Compared with the fourth interim 96-month analysis, in this final 120-month analysis, the number of new users in the female breast cancer cohort and the sex-combined bladder cancer cohort, respectively, decreased by 1,393 (−6%) and 1,511 (−3%) in PHARMO, and increased by 162,165 (85%) and 262,185 (64%) in the HIRD, and 141,689 (27%) and 263,774 (28%) in Medicare. In CPRD, for this final analysis, the number of new users in the female breast cancer cohort and the sex-combined bladder cancer cohort was the same as the number of new users in the fourth interim analysis (see Section 9.5.1).

Figure 7 **Selection of New Users of Dapagliflozin and Comparator ADs Into Cancer Outcome Cohorts (Full Samples), by Data Source**

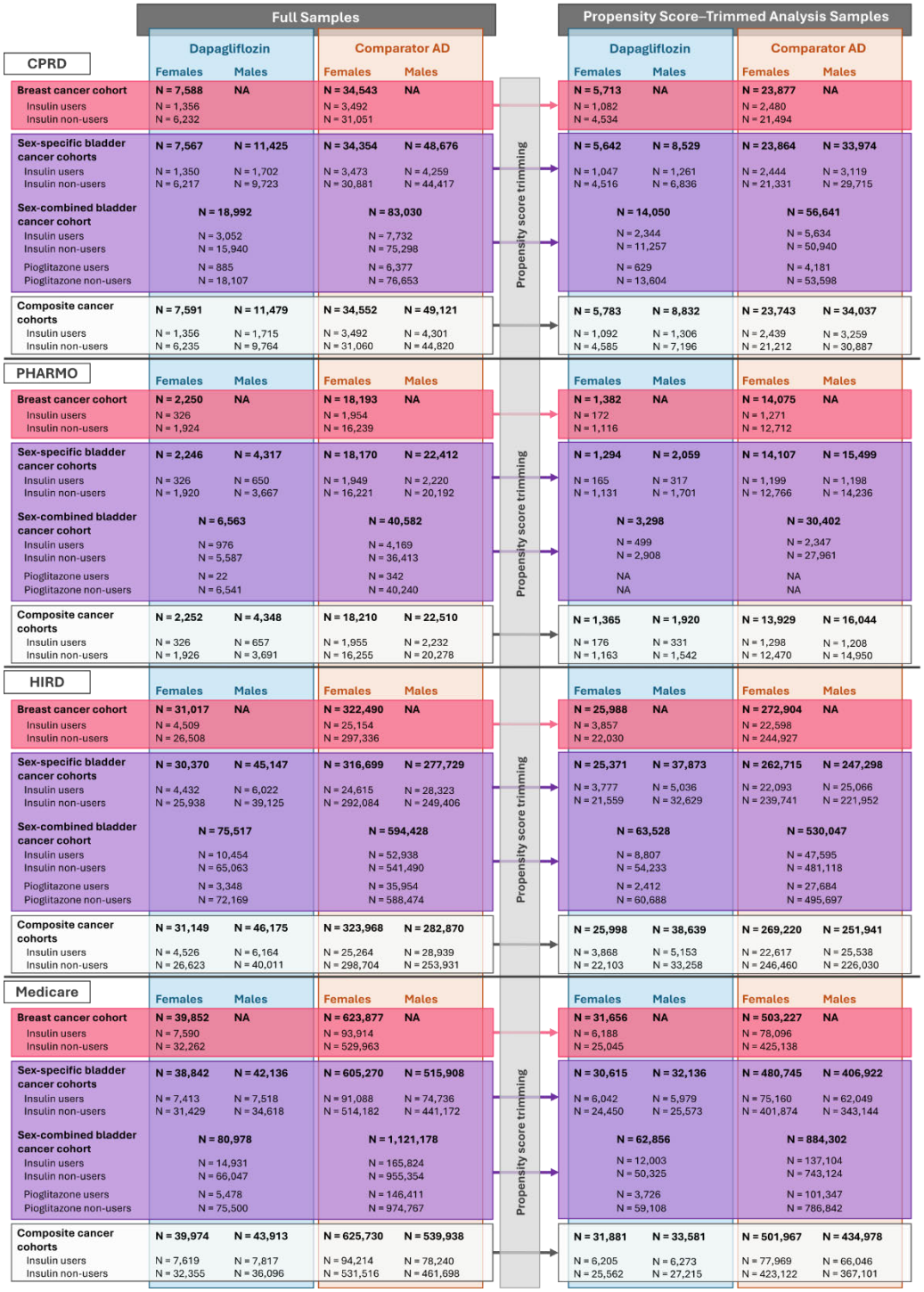
[a] For PHARMO, the potential index dates displayed are those that were eligible for linkage with the NCR (see Appendix J, Table 0 PHARMO).

Note: The eligibility criteria common to all cancer outcomes and the exclusion criteria specific to each cancer outcome are described in Section 9.3.1. The detailed results of the number of index dates and new users excluded for each exclusion criterion are presented in Appendix J, Table 0, for each data source.

NA = not applicable.

Figure 8 displays the final number of eligible new users for each data source and cancer outcome cohort—for the dapagliflozin and comparator AD exposure groups before propensity score trimming (ie, the full samples) and after propensity score trimming, for the overall cohorts and in the cohorts stratified by insulin use at the index date (and, for the sex-combined bladder cancer cohort, stratified by pioglitazone use at the index date). The final propensity score−trimmed samples, which were used to conduct the incidence analyses and the final comparative analyses for each outcome, are described in more detail (eg, proportions of the full samples that were "trimmed out") in Section 10.2.3.

Figure 8 **Number of New Users of Dapagliflozin and Comparator AD in All Cancer Outcome Cohorts, Full Samples (Before Propensity Score Trimming) and Propensity Score–Trimmed Analysis Samples, by Data Source**



Note: In the propensity score–trimmed analysis samples, the size of the overall sample may not equal the sum of the insulin use–stratified samples, as propensity score estimation and trimming were performed separately in each sample.

NA = not applicable.

10.2 Descriptive Data

10.2.1 Description of the Index Prescription of Dapagliflozin

The description of the characteristics of the index dapagliflozin prescription or dispensing for each year of each data source–specific period focused on the composite cancer outcome cohorts because these cohorts were larger than the female breast and sex-specific bladder cancer cohorts. *Results are summarised below, and complete results are presented in Appendix J, FemaleCompositeCa Table 6 and MaleCompositeCa Table 6, for each data source.*

Female Composite Cancer

Amongst new users of dapagliflozin in the female composite cancer cohort, the average number of months of dapagliflozin exposure accumulated at the final data cut was 35.6 months in CPRD, 12.7 months in PHARMO, 13.8 months in the HIRD, and 21.3 months in Medicare. In all data sources, the index medication type for new users of dapagliflozin was most commonly classified as add-on index therapy (range, 40.7% in PHARMO to 86.6% in CPRD), and this finding was consistent across all years of the study (except in PHARMO, where in the most recent study years, 2021 and 2022, dapagliflozin was most commonly initiated as index monotherapy). New users of dapagliflozin classified as index monotherapy ranged from 4.1% in CPRD to 28.5% in PHARMO, with relevant increases over time in PHARMO, the HIRD, and Medicare; the increase was more pronounced during the most recent study years (up to 42.5% in PHARMO).

When assessing the entire study period, the proportion of new users of dapagliflozin with concomitant insulin use at the index date was similar across the data sources (17.9% CPRD, 14.5% in PHARMO and the HIRD, and 19.1% in Medicare). The proportion of new users of dapagliflozin with concomitant insulin use decreased substantially over time in CPRD (from 27.2% of index dates in 2012-2013 to 4.6% of index dates in 2021), remained relatively stable over time in the HIRD and Medicare, and varied considerably over time in PHARMO.

Male Composite Cancer

Amongst new users of dapagliflozin in the male composite cancer cohort, the average number of months of dapagliflozin exposure accumulated at the final data cut was 34.9 months in CPRD, 12.0 months in PHARMO, 15.1 months in the HIRD, and 19.8 months in Medicare. In all four data sources, the index medication type for new users of dapagliflozin was most commonly add-on index therapy (range, 36.9% in PHARMO to 89.0% in CPRD), and this finding was consistent across all years of the study (except in PHARMO, where in the most recent study years, 2021 and 2022, dapagliflozin was most commonly initiated as index monotherapy). During the entire study period, new users of dapagliflozin classified as index monotherapy ranged from 3.5% in CPRD to 33.7% in PHARMO, with relevant increases over

time in PHARMO, the HIRD, and Medicare; increases were more pronounced during the most recent study years (up to 47.7% in PHARMO).

When assessing the entire study period, the proportion of new users of dapagliflozin with concomitant insulin use at the index date was similar across the data sources (14.9% in CPRD, 15.1% in PHARMO, 13.3% in the HIRD, and 17.8.% in Medicare). The variation patterns in the proportions of new users of dapagliflozin with concomitant insulin use over time were similar to those described for the female composite cancer cohorts.

10.2.2 Baseline Characteristics

The baseline characteristics of new users of dapagliflozin and of comparator ADs overall and by insulin use at the index date—before propensity score trimming, including the StDiff values of the comparison for each variable—are provided in outcome-specific tables for each data source in Appendix J, Tables 1 (baseline characteristics), 2 (specific baseline medical conditions), 3 (specific baseline medications), and 4 (HCRU during the 180 days before the index date). For the bladder cancer outcome only, baseline characteristics of new users of dapagliflozin and of comparator ADs stratified by pioglitazone use at the index date before propensity score trimming are provided for each data source in Appendix J, BladderCa Table 17 (baseline characteristics), BladderCa Table 18 (specific baseline medical conditions), BladderCa Table 19 (specific baseline medications), and BladderCa Table 20 (HCRU during the 180 days before the index date). This analysis addresses secondary objective 2 (see Section 7.2.2).

For all cancer outcome cohorts, the average duration of lookback time before the index date of dapagliflozin or comparator AD was longer in CPRD (about 12 years in both exposure groups) and PHARMO (14 to 15 years and 10 to 12 years, respectively) than in the HIRD (about 4 years in both exposure groups) and Medicare (4 to 5 years in both exposure groups)—see Appendix J, BreastCa Table 1, BladderCa Tables 1, 1M, and 1F, FemaleCompositeCa Table 1, and MaleCompositeCa Table 1, for each data source.

As described in Section 9.9.2, values of the StDiff of 0.2, 0.5, and 0.8 roughly correspond to small, medium, and large differences, respectively, in the level of the covariate between dapagliflozin and comparator AD groups; therefore, variables that had StDiff > 0.20 between the exposure groups are presented in bold font in the tables summarising baseline characteristics for each outcome cohort.

10.2.2.1 Female Breast Cancer

For the female breast cancer cohorts, results of selected baseline characteristics (measured before or on the index date [see Figure 1]) before propensity score trimming are summarised in Table 12, and the results of all assessed baseline characteristics for each data source are presented in Appendix J, BreastCa Tables 1, 2, 3 and 4.

Table 12 Selected Baseline Characteristics of the Female Breast Cancer Cohort Before Propensity Score Trimming (Full Sample), Overall and Stratified by Insulin Use at the Index Date, by Data Source

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	**Dapagliflozin**	**Comparator AD**
No. of new users[a], n (%)	Insulin	1,356 (17.9)	3,492 (10.1)	326 (14.5)	1,954 (10.7)	4,509 (14.5)	25,154 (7.8)	7,590 (19.0)	93,914 (15.1)
	No insulin	6,232 (82.1)	31,051 (89.9)	1,924 (85.5)	16,239 (89.3)	26,508 (85.5)	297,336 (92.2)	32,262 (81.0)	529,963 (84.9)
	Overall	7,588 (100.0)	34,543 (100.0)	2,250 (100.0)	18,193 (100.0)	31,017 (100.0)	322,490 (100.0)	39,852 (100.0)	623,877 (100.0)
Age (years), mean (SD)	Insulin	**57.6 (8.96)**	**61.7 (11.90)**	65.9 (10.45)	67.9 (12.63)	54.4 (6.57)	53.9 (6.72)	71.6 (5.68)	72.2 (6.81)
	No insulin	**58.6 (9.89)**	**64.7 (12.49)**	66.6 (11.37)	65.2 (12.29)	**54.0 (6.59)**	**52.2 (6.86)**	72.7 (6.24)	73.2 (7.05)
	Overall	**58.4 (9.73)**	**64.4 (12.46)**	66.5 (11.25)	65.5 (12.35)	**54.0 (6.59)**	**52.3 (6.87)**	72.5 (6.16)	73.1 (7.02)
Comorbidities, n (%)									
Heart failure	Insulin	26 (1.9)	201 (5.8)	**51 (15.6)**	**133 (6.8)**	416 (9.2)	1,442 (5.7)	1,716 (22.6)	20,601 (21.9)
	No insulin	164 (2.6)	1,232 (4.0)	**363 (18.9)**	**525 (3.2)**	**2,146 (8.1)**	**5,950 (2.0)**	6,523 (20.2)	77,019 (14.5)
	Overall	190 (2.5)	1,433 (4.1)	**414 (18.4)**	**658 (3.6)**	**2,562 (8.3)**	**7,392 (2.3)**	8,239 (20.7)	97,620 (15.6)
Chronic kidney disease	Insulin	**160 (11.8)**	**970 (27.8)**	45 (13.8)	147 (7.5)	566 (12.6)	2,155 (8.6)	2,416 (31.8)	29,782 (31.7)
	No insulin	**542 (8.7)**	**7,028 (22.6)**	176 (9.1)	715 (4.4)	**2,013 (7.6)**	**8,529 (2.9)**	7,679 (23.8)	105,725 (19.9)
	Overall	**702 (9.3)**	**7,998 (23.2)**	221 (9.8)	862 (4.7)	**2,579 (8.3)**	**10,684 (3.3)**	10,095 (25.3)	135,507 (21.7)
Indicators of diabetes severity, n (%)									
Diabetic nephropathy or renal insufficiency	Insulin	17 (1.3)	66 (1.9)	36 (11.0)	123 (6.3)	668 (14.8)	2,512 (10.0)	2,354 (31.0)	23,456 (25.0)
	No insulin	33 (0.5)	178 (0.6)	128 (6.7)	624 (3.8)	1,470 (5.5)	6,318 (2.1)	5,649 (17.5)	67,580 (12.8)
	Overall	50 (0.7)	244 (0.7)	164 (7.3)	747 (4.1)	2,138 (6.9)	8,830 (2.7)	8,003 (20.1)	91,036 (14.6)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	**Dapagliflozin**	**Comparator AD**
Retinopathy	Insulin	568 (41.9)	1,404 (40.2)	30 (9.2)	113 (5.8)	2,407 (53.4)	11,959 (47.5)	**5,569 (73.4)**	**55,577 (59.2)**
	No insulin	1,566 (25.1)	7,575 (24.4)	59 (3.1)	396 (2.4)	**7,771 (29.3)**	**43,748 (14.7)**	**15,751 (48.8)**	**184,373 (34.8)**
	Overall	2,134 (28.1)	8,979 (26.0)	89 (4.0)	509 (2.8)	**10,178 (32.8)**	**55,707 (17.3)**	**21,320 (53.5)**	**239,950 (38.5)**
Peripheral neuropathy	Insulin	108 (8.0)	210 (6.0)	17 (5.2)	87 (4.5)	116 (2.6)	559 (2.2)	541 (7.1)	7,728 (8.2)
	No insulin	142 (2.3)	851 (2.7)	77 (4.0)	470 (2.9)	422 (1.6)	3,709 (1.2)	1,914 (5.9)	34,569 (6.5)
	Overall	250 (3.3)	1,061 (3.1)	94 (4.2)	557 (3.1)	538 (1.7)	4,268 (1.3)	2,455 (6.2)	42,297 (6.8)
Peripheral vascular disease[b]	Insulin	52 (3.8)	197 (5.6)	13 (4.0)	39 (2.0)	2,321 (51.5)	11,561 (46.0)	**5,436 (71.6)**	**55,883 (59.5)**
	No insulin	119 (1.9)	1,245 (4.0)	75 (3.9)	281 (1.7)	**7,824 (29.5)**	**43,733 (14.7)**	**16,206 (50.2)**	**199,446 (37.6)**
	Overall	171 (2.3)	1,442 (4.2)	88 (3.9)	320 (1.8)	**10,145 (32.7)**	**55,294 (17.1)**	**21,642 (54.3)**	**255,329 (40.9)**
Coronary heart disease	Insulin	158 (11.7)	595 (17.0)	**132 (40.5)**	**356 (18.2)**	765 (17.0)	3,123 (12.4)	2,963 (39.0)	33,968 (36.2)
	No insulin	479 (7.7)	4,069 (13.1)	**695 (36.1)**	**2,218 (13.7)**	**3,351 (12.6)**	**18,005 (6.1)**	11,164 (34.6)	153,348 (28.9)
	Overall	637 (8.4)	4,664 (13.5)	**827 (36.8)**	**2,574 (14.1)**	**4,116 (13.3)**	**21,128 (6.6)**	14,127 (35.4)	187,316 (30.0)
Cerebrovascular disease	Insulin	73 (5.4)	285 (8.2)	36 (11.0)	145 (7.4)	243 (5.4)	1,127 (4.5)	1,444 (19.0)	18,549 (19.8)
	No insulin	264 (4.2)	2,580 (8.3)	171 (8.9)	715 (4.4)	985 (3.7)	6,824 (2.3)	5,750 (17.8)	87,881 (16.6)
	Overall	337 (4.4)	2,865 (8.3)	207 (9.2)	860 (4.7)	1,228 (4.0)	7,951 (2.5)	7,194 (18.1)	106,430 (17.1)
Amputation	Insulin	7 (0.5)	37 (1.1)	0 (0.0)	5 (< 0.5)	60 (1.3)	328 (1.3)	131 (1.7)	1,687 (1.8)
	No insulin	19 (0.3)	163 (0.5)	6 (< 0.5)	26 (< 0.5)	80 (0.3)	524 (0.2)	193 (0.6)	3,636 (0.7)
	Overall	26 (0.3)	200 (0.6)	6 (< 0.5)	31 (< 0.5)	140 (0.5)	852 (0.3)	324 (0.8)	5,323 (0.9)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	**Dapagliflozin**	**Comparator AD**
Insulin use at the index date[c], n (%)	Insulin	NA	NA	NA	NA	NA	NA	NA	NA
	No insulin	NA	NA	NA	NA	NA	NA	NA	NA
	Overall	**1,356 (17.9)**	**3,492 (10.1)**	326 (14.5)	1,954 (10.7)	**4,509 (14.5)**	**25,154 (7.8)**	7,590 (19.0)	93,914 (15.1)
Index medication type, n (%)									
Index monotherapy	Insulin	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	No insulin	312 (5.0)	1,848 (6.0)	**641 (33.3)**	**950 (5.9)**	**4,917 (18.5)**	**112,598 (37.9)**	6,831 (21.2)	81,938 (15.5)
	Overall	312 (4.1)	1,848 (5.3)	**641 (28.5)**	**950 (5.2)**	**4,917 (15.9)**	**112,598 (34.9)**	6,831 (17.1)	81,938 (13.1)
Add-on index therapy	Insulin	952 (70.2)	2,134 (61.1)	166 (50.9)	1,042 (53.3)	3,484 (77.3)	19,673 (78.2)	4,822 (63.5)	57,062 (60.8)
	No insulin	3,585 (57.5)	20,502 (66.0)	**750 (39.0)**	**9,240 (56.9)**	**15,393 (58.1)**	**125,993 (42.4)**	14,738 (45.7)	246,492 (46.5)
	Overall	4,537 (59.8)	22,636 (65.5)	**916 (40.7)**	**10,282 (56.5)**	**18,877 (60.9)**	**145,666 (45.2)**	19,560 (49.1)	303,554 (48.7)
Switched-to index therapy	Insulin	11 (0.8)	106 (3.0)	18 (5.5)	191 (9.8)	6 (0.1)	50 (0.2)	360 (4.7)	6,572 (7.0)
	No insulin	396 (6.4)	3,705 (11.9)	**108 (5.6)**	**2,476 (15.2)**	792 (3.0)	7,240 (2.4)	3,532 (10.9)	94,329 (17.8)
	Overall	**407 (5.4)**	**3,811 (11.0)**	**126 (5.6)**	**2,667 (14.7)**	798 (2.6)	7,290 (2.3)	3,892 (9.8)	100,901 (16.2)
Combined index therapy with no prior treatment	Insulin	**1-4 (0.1-0.3)**	**129 (3.7)**	**4 (1.2)**	**108 (5.5)**	86 (1.9)	910 (3.6)	141 (1.9)	4,020 (4.3)
	No insulin	94-97 (1.5-1.6)	1,438 (4.6)	**81 (4.2)**	**1,910 (11.8)**	1,822 (6.9)	19,029 (6.4)	1,740 (5.4)	48,720 (9.2)
	Overall	98 (1.3)	1,567 (4.5)	**85 (3.8)**	**2,018 (11.1)**	1,908 (6.2)	19,939 (6.2)	1,881 (4.7)	52,740 (8.5)
Add-on and switched-to index therapy	Insulin	364 (26.8)	1,020 (29.2)	85 (26.1)	542 (27.7)	499 (11.1)	1,682 (6.7)	1,817 (23.9)	23,778 (25.3)
	No insulin	**1,667 (26.7)**	**2,621 (8.4)**	223 (11.6)	1,157 (7.1)	**1,709 (6.4)**	**5,329 (1.8)**	3,953 (12.3)	48,614 (9.2)
	Overall	**2,031 (26.8)**	**3,641 (10.5)**	308 (13.7)	1,699 (9.3)	**2,208 (7.1)**	**7,011 (2.2)**	5,770 (14.5)	72,392 (11.6)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	**Dapagliflozin**	**Comparator AD**
Non-evaluable[d]	Insulin	25-28 (1.8-2.1)	103 (2.9)	**53 (16.3)**	**71 (3.6)**	434 (9.6)	2,839 (11.3)	450 (5.9)	2,482 (2.6)
	No insulin	175-178 (2.8-2.9)	937 (3.0)	121 (6.3)	506 (3.1)	1,875 (7.1)	27,147 (9.1)	1,468 (4.6)	9,870 (1.9)
	Overall	203 (2.7)	1,040 (3.0)	174 (7.7)	577 (3.2)	2,309 (7.4)	29,986 (9.3)	1,918 (4.8)	12,352 (2.0)
No. of AD classes[e] used within 12 months before the index date[f], n (%)									
0	Insulin	1-4 (0.1-0.3)	67-70 (1.9-2.0)	1 (< 0.5)	44 (2.3)	50 (1.1)	682 (2.7)	32 (0.4)	1,735 (1.8)
	No insulin	198-201 (3.2)	2,135-2,138 (6.9)	**593 (30.8)**	**1,450 (8.9)**	**5,578 (21.0)**	**128,912 (43.4)**	5,068 (15.7)	70,942 (13.4)
	Overall	202 (2.7)	2,205 (6.4)	**594 (26.4)**	**1,494 (8.2)**	**5,628 (18.1)**	**129,594 (40.2)**	5,100 (12.8)	72,677 (11.6)
1-2	Insulin	**602 (44.4)**	**2,207 (63.2)**	**200 (61.3)**	**1,429 (73.1)**	**2,264 (50.2)**	**17,502 (69.6)**	**3,754 (49.5)**	**62,447 (66.5)**
	No insulin	**4,341 (69.7)**	**27,254 (87.8)**	**1,028 (53.4)**	**13,486 (83.0)**	**17,114 (64.6)**	**160,582 (54.0)**	**20,162 (62.5)**	**420,479 (79.3)**
	Overall	**4,943 (65.1)**	**29,461 (85.3)**	**1,228 (54.6)**	**14,915 (82.0)**	19,378 (62.5)	178,084 (55.2)	**23,916 (60.0)**	**482,926 (77.4)**
≥ 3	Insulin	**750-753 (55.3-55.5)**	**1,215-1,218 (34.8-34.9)**	**122 (37.4)**	**320 (16.4)**	**2,195 (48.7)**	**6,970 (27.7)**	**3,804 (50.1)**	**29,732 (31.7)**
	No insulin	**1,690-1,693 (27.1-27.2)**	**1,659-1,662 (5.3-5.4)**	**263 (13.7)**	**295 (1.8)**	**3,816 (14.4)**	**7,842 (2.6)**	**7,032 (21.8)**	**38,542 (7.3)**
	Overall	**2,443 (32.2)**	**2,877 (8.3)**	**385 (17.1)**	**615 (3.4)**	**6,011 (19.4)**	**14,812 (4.6)**	**10,836 (27.2)**	**68,274 (10.9)**

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	**Dapagliflozin**	**Comparator AD**
No. of outpatient visits[g] in the 180 days before cohort entry, n (%)									
0-1	Insulin	270 (19.9)	723 (20.7)	52 (38.2)	309 (38.4)	457 (10.1)	3,246 (12.9)	816 (10.8)	12,036 (12.8)
	No insulin	1,441 (23.1)	7,150 (23.0)	391 (43.8)	3,099 (42.8)	4,004 (15.1)	50,668 (17.0)	3,940 (12.2)	76,904 (14.5)
	Overall	1,711 (22.6)	7,873 (22.8)	443 (43.1)	3,408 (42.4)	4,461 (14.4)	53,914 (16.7)	4,756 (11.9)	88,940 (14.3)
2-3	Insulin	228 (16.8)	520 (14.9)	24 (17.6)	110 (13.7)	772 (17.1)	4,341 (17.3)	1,192 (15.7)	15,890 (16.9)
	No insulin	1,306 (21.0)	5,677 (18.3)	121 (13.6)	1,062 (14.7)	5,521 (20.8)	62,647 (21.1)	5,758 (17.8)	105,509 (19.9)
	Overall	1,534 (20.2)	6,197 (17.9)	145 (14.1)	1,172 (14.6)	6,293 (20.3)	66,988 (20.8)	6,950 (17.4)	121,399 (19.5)
≥ 4	Insulin	858 (63.3)	2,249 (64.4)	60 (44.1)	386 (48.0)	3,280 (72.7)	17,567 (69.8)	5,582 (73.5)	65,988 (70.3)
	No insulin	3,485 (55.9)	18,224 (58.7)	380 (42.6)	3,072 (42.5)	16,983 (64.1)	184,021 (61.9)	22,564 (69.9)	347,550 (65.6)
	Overall	4,343 (57.2)	20,473 (59.3)	440 (42.8)	3,458 (43.0)	20,263 (65.3)	201,588 (62.5)	28,146 (70.6)	413,538 (66.3)
Hospitalisation in the 180 days before cohort entry, yes, n (%)	Insulin	206 (15.2)	706 (20.2)	36 (11.0)	317 (16.2)	530 (11.8)	3,065 (12.2)	989 (13.0)	17,106 (18.2)
	No insulin	595 (9.5)	4,379 (14.1)	**311 (16.2)**	**1,257 (7.7)**	1,776 (6.7)	11,334 (3.8)	3,159 (9.8)	63,266 (11.9)
	Overall	801 (10.6)	5,085 (14.7)	**347 (15.4)**	**1,574 (8.7)**	2,306 (7.4)	14,399 (4.5)	4,148 (10.4)	80,372 (12.9)

[a] The number of new users for each exposure group comprises those identified in each data source that met the eligibility criteria within each of the insulin use–stratified cohorts and overall cohorts. Unless otherwise specified, the values in the rows for "No. of new users" represent the denominators for each cohort and exposure group that were used in calculating the proportions of baseline characteristics presented throughout the remainder of the table.

[b] Includes peripheral artery disease.

[c] Frequencies and proportions of patients with insulin use at the index date are reported only in the overall cohorts. The "insulin use at the index date" variable was used to define the stratified insulin user and insulin non-user cohorts; therefore, this variable is reported as "NA" for the insulin use–stratified cohorts.

[d] Patients who did not have sufficient follow-up time to assess the 90-day add-on/switch requirement.

[e] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

[f] For PHARMO, users with insufficient lookback time in their medical record to assess this variable were categorised as "not applicable," as indicated in Appendix J, BreastCa Table 3 PHARMO, and include the following: insulin users cohort, n (%): dapagliflozin, **3 (0.9)**; comparator AD, **161 (8.2)**; insulin non-users cohort, n (%): dapagliflozin, **40 (2.1)**; comparator AD, **1,008 (6.2)**; overall cohort, n (%): dapagliflozin, **43 (1.9)**; comparator AD, **1,169 (6.4)**.

[g] Outpatient visits included GP and outpatient hospital visits. For PHARMO, the denominator for percentage calculations included only patients with available linked GP data (insulin users cohort: dapagliflozin, n = 136; comparator AD, n = 805; insulin non-users cohort: dapagliflozin, n = 892; comparator AD, n = 7,233; overall cohort: dapagliflozin, n = 1,028; comparator AD, n = 8,038).

Note 1: For CPRD data, any cell with a count of 1-4, or any cell that allows a count of 1-4 to be derived from other reported cells or information, cannot be reported. To comply with this reporting requirement, values that could be used to derive small count sizes in other cells are presented as ranges in this table.

Note 2: Values in **bold** in the table and the footnotes indicate a StDiff > 0.20 between new users of dapagliflozin and new users of comparator AD.

NA = not applicable; No. = number; StDiff = absolute standardised difference.

Most numbers and proportions for baseline variables for the overall and insulin use–stratified cohorts described in this section are presented in Table 12. Baseline variables with StDiff values > 0.20 are noted in the text below (and are bolded in Table 12) to highlight differences across the index exposure groups. Any proportions or StDiff values for baseline variables that are noted in the text below but are not presented in Table 12 include a reference to the relevant tables in Appendix J where the results can be found.

In the overall female breast cancer cohorts before propensity score trimming, dapagliflozin new users were younger than comparator AD new users in CPRD (StDiff, 0.54) and the opposite was observed in the HIRD (StDiff, 0.25). In PHARMO and Medicare, the mean age was similar across dapagliflozin new users and comparator AD new users. In CPRD, the HIRD, and Medicare, most dapagliflozin new users and comparator AD new users, respectively, with available data on race/ethnicity were White (57% and 59% in CPRD, 53% and 58% in the HIRD, and 73% and 76% in Medicare) (Appendix J, BreastCa Table 1, for CPRD, HIRD, and Medicare).

The variables heart failure and chronic kidney disease were added in the final analysis to describe and account for the additional approved indications of dapagliflozin. In the overall cohorts, the prevalence of heart failure was similar across both exposure groups in CPRD and Medicare, whilst in PHARMO and the HIRD, heart failure was more prevalent in dapagliflozin new users than comparator new users (StDiff, 0.49 in PHARMO and 0.27 in the HIRD). For chronic kidney disease, the prevalence was higher in comparator AD new users than in dapagliflozin new users in CPRD (StDiff, 0.38), whilst the opposite was observed in the HIRD, with a higher prevalence in dapagliflozin new users than in comparator AD new users (StDiff, 0.21). In PHARMO and Medicare, the prevalence of chronic kidney disease was similar across the exposure groups.

Amongst the indicators of diabetes severity, in the overall cohorts, retinopathy was the most common in CPRD, coronary heart disease was the most common in PHARMO, and both retinopathy and peripheral vascular disease were the most common (with similar proportions) in the HIRD and Medicare. In PHARMO, indicators of diabetes severity were obtained only from hospital diagnoses. Retinopathy and peripheral vascular diseases, respectively, were more prevalent in dapagliflozin new users than in comparator AD new users in the HIRD (StDiff, 0.36 and 0.37) and Medicare (StDiff, 0.31 and 0.27). The prevalence of coronary heart disease was higher in dapagliflozin new users than in comparator AD new users in PHARMO (StDiff, 0.54) and in the HIRD (StDiff, 0.23).

In CPRD, dapagliflozin new users were more likely to be severely obese (BMI ≥ 40) than comparator AD new users (25.4% vs 16.6%; StDiff, 0.22) (Appendix J, BreastCa Table 1 CPRD); whereas, in PHARMO (based on the subset of 40% of patients with GP data available), the proportions for each BMI category were similar across the exposure groups

(Appendix J, BreastCa Table 1 PHARMO). In CPRD and PHARMO, the proportions of new users who were current smokers or former smokers and the proportions with low-moderate alcohol intake or heavy or very heavy alcohol intake were similar amongst dapagliflozin new users and comparator AD new users (Appendix J, BreastCa Table 1 CPRD and BreastCa Table 1 PHARMO).

Dapagliflozin and comparator AD index medications were most commonly initiated as an add-on therapy across all data sources, with higher proportions of comparator AD new users than dapagliflozin new users in PHARMO (StDiff, 0.32), higher proportions of dapagliflozin new users than comparator AD new users in the HIRD (StDiff, 0.32), and similar proportions in both exposure groups in CPRD and Medicare. Dapagliflozin and comparator AD index medications were also frequently initiated as monotherapy in the HIRD and Medicare, with higher proportions of new users initiating the index medication as monotherapy in the comparator AD group than in the dapagliflozin group in the HIRD (StDiff, 0.45) and similar proportions across the exposure groups in Medicare; in PHARMO, only dapagliflozin index medications were frequently initiated as monotherapy, with few comparator AD new users initiating the index therapy as monotherapy (StDiff, 0.65). Compared with comparator AD, dapagliflozin was more frequently initiated as add-on and switched-to therapy in the CPRD (StDiff, 0.43).

Differences were observed in all data sources between new users of dapagliflozin and comparator AD new users in the variables measuring the number of AD classes used in time periods before the index date, with StDiff values up to 0.62 in some categories. Dapagliflozin new users were more likely than comparator AD new users to have had at least three AD classes before the index date in all data sources (Table 12 displays the time period of 12 months before the index date for each data source; Appendix J, BreastCa Table 3, shows all measured time periods before the index date for each data source).

In the overall cohorts in CPRD, the HIRD, and Medicare, approximately 60% or more of new users of dapagliflozin and of comparator ADs had at least four outpatient visits, with percentages lower in PHARMO for both exposure groups (approximately 43%), which is based on a subset of patients with available GP data. Across all data sources, the proportion of new users of dapagliflozin and of comparator ADs that had no hospitalisations in the 180 days before the index date ranged from 85% to 96%; the proportion of new users with at least one hospitalisation in the 180 days before the index date was similar or differed slightly between the exposure groups across all data sources; in PHARMO a higher proportion of comparator AD new users than dapagliflozin new users had no hospitalisation in the 180 days before the index date (StDiff, 0.21) (Appendix J, BreastCa Table 4 PHARMO).

Insulin Use at the Index Date

A higher proportion of dapagliflozin new users than comparator AD new users had concomitant insulin use at the index date in CPRD (StDiff, 0.23) and the HIRD (StDiff, 0.22).

The prevalence of heart failure at the index date was similar for both exposure groups amongst insulin users and insulin non-users in CPRD and Medicare, whilst in PHARMO heart failure was more prevalent in dapagliflozin new users than in the comparator AD new users amongst insulin users (StDiff, 0.28) and insulin non-users (StDiff, 0.51). In the HIRD, amongst insulin users, heart failure had a similar prevalence in both exposure groups, and amongst insulin non-users, heart failure was more frequent in the dapagliflozin group than in the comparator AD group (StDiff, 0.28). The prevalence of chronic kidney disease at the index date was higher in the comparator AD group than in the dapagliflozin group in CPRD, both amongst insulin users and amongst insulin non-users (StDiff, 0.41 and 0.39, respectively), with similar prevalence in both exposure groups across insulin users and insulin non-users in PHARMO and Medicare. In the HIRD, the prevalence of chronic kidney disease was similar in both exposure groups amongst insulin users and higher in the dapagliflozin group than in the comparator AD group amongst insulin non-users (StDiff, 0.21).

Across all data sources, except for some indicators of diabetes severity in PHARMO, in both dapagliflozin and comparator AD exposure groups, the prevalence of diabetes severity indicators was higher amongst insulin users than amongst insulin non-users at the index date. The indicators of diabetes severity with the highest baseline prevalences in both exposure groups were retinopathy in CPRD, the HIRD, and Medicare and coronary heart disease in PHARMO; these conditions were more common in insulin users than in insulin non-users. Also in both exposure groups, peripheral vascular disease and coronary heart disease were more prevalent at baseline in insulin users than in insulin non-users across all data sources. However, in PHARMO, the baseline prevalence of peripheral vascular disease was comparable amongst insulin users and insulin non-users in the dapagliflozin new user group. In CPRD, in the insulin non-users cohort, dapagliflozin new users were more likely to be severely obese than comparator AD new users (StDiff, 0.24) (see Appendix J, BreastCa Table 1 CPRD)

10.2.2.2 Bladder Cancer

For the sex-combined bladder cancer cohorts, selected baseline characteristics (measured before or on the index date [see Figure 1]) before propensity score trimming are summarised in Table 13, and the results of all assessed baseline characteristics for each data source are presented in Appendix J, BladderCa Table 1, BladderCa Table 2, BladderCa Table 3, and BladderCa Table 4. The baseline characteristics for the bladder cancer cohorts stratified by pioglitazone use at the index date are presented in Appendix J, BladderCa Table 17, BladderCa Table 18, BladderCa Table 19, and BladderCa Table 20, for each data source.

Table 13 Selected Baseline Characteristics of the Sex-Combined Bladder Cancer Cohort Before Propensity Score Trimming (Full Sample), Overall and Stratified by Insulin Use at the Index Date, by Data Source

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of new users[a], n (%)	Insulin	3,052 (16.1)	7,732 (9.3)	976 (14.9)	4,169 (10.3)	10,454 (13.8)	52,938 (8.9)	14,931 (18.4)	165,824 (14.8)
	No insulin	15,940 (83.9)	75,298 (90.7)	5,587 (85.1)	36,413 (89.7)	65,063 (86.2)	541,490 (91.1)	66,047 (81.6)	955,354 (85.2)
	Overall	18,992 (100.0)	83,030 (100.0)	6,563 (100.0)	40,582 (100.0)	75,517 (100.0)	594,428 (100.0)	80,978 (100.0)	1,121,178 (100.0)
Age (years), mean (SD)	Insulin	**58.4 (8.89)**	**61.8 (11.26)**	66.1 (10.13)	66.4 (11.99)	54.4 (6.51)	53.9 (6.69)	71.5 (5.46)	71.8 (6.36)
	No insulin	**58.9 (9.55)**	**63.6 (11.93)**	66.0 (10.93)	64.1 (11.68)	54.0 (6.57)	52.6 (6.83)	72.4 (5.96)	72.7 (6.61)
	Overall	**58.8 (9.45)**	**63.4 (11.88)**	66.0 (10.82)	64.3 (11.73)	54.0 (6.57)	52.7 (6.83)	72.3 (5.88)	72.5 (6.58)
Sex									
Female	Insulin	1,350 (44.2)	3,473 (44.9)	**326 (33.4)**	**1,949 (46.7)**	4,432 (42.4)	24,615 (46.5)	7,413 (49.6)	91,088 (54.9)
	No insulin	6,217 (39.0)	30,881 (41.0)	**1,920 (34.4)**	**16,221 (44.5)**	**25,938 (39.9)**	**292,084 (53.9)**	31,429 (47.6)	514,182 (53.8)
	Overall	7,567 (39.8)	34,354 (41.4)	**2,246 (34.2)**	**18,170 (44.8)**	**30,370 (40.2)**	**316,699 (53.3)**	38,842 (48.0)	605,270 (54.0)
Male	Insulin	1,702 (55.8)	4,259 (55.1)	**650 (66.6)**	**2,220 (53.3)**	6,022 (57.6)	28,323 (53.5)	7,518 (50.4)	74,736 (45.1)
	No insulin	9,723 (61.0)	44,417 (59.0)	**3,667 (65.6)**	**20,192 (55.5)**	**39,125 (60.1)**	**249,406 (46.1)**	34,618 (52.4)	441,172 (46.2)
	Overall	11,425 (60.2)	48,676 (58.6)	**4,317 (65.8)**	**22,412 (55.2)**	**45,147 (59.8)**	**277,729 (46.7)**	42,136 (52.0)	515,908 (46.0)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Comorbidities, n (%)									
Heart failure	Insulin	114 (3.7)	551 (7.1)	**154 (15.8)**	**304 (7.3)**	1,076 (10.3)	3,443 (6.5)	3,648 (24.4)	35,422 (21.4)
	No insulin	579 (3.6)	3,446 (4.6)	**1,197 (21.4)**	**1,307 (3.6)**	**6,664 (10.2)**	**13,570 (2.5)**	**15,125 (22.9)**	**135,389 (14.2)**
	Overall	693 (3.6)	3,997 (4.8)	**1,351 (20.6)**	**1,611 (4.0)**	**7,740 (10.2)**	**17,013 (2.9)**	18,773 (23.2)	170,811 (15.2)
Chronic kidney disease	Insulin	**334 (10.9)**	**1,961 (25.4)**	**139 (14.2)**	**286 (6.9)**	1,386 (13.3)	4,966 (9.4)	5,160 (34.6)	53,271 (32.1)
	No insulin	**1,175 (7.4)**	**14,074 (18.7)**	514 (9.2)	1,529 (4.2)	**5,359 (8.2)**	**17,953 (3.3)**	16,603 (25.1)	191,111 (20.0)
	Overall	**1,509 (7.9)**	**16,035 (19.3)**	**653 (9.9)**	**1,815 (4.5)**	**6,745 (8.9)**	**22,919 (3.9)**	21,763 (26.9)	244,382 (21.8)
Indicators of diabetes severity, n (%)									
Diabetic nephropathy or renal insufficiency	Insulin	62 (2.0)	183 (2.4)	**110 (11.3)**	**226 (5.4)**	1,577 (15.1)	5,655 (10.7)	5,007 (33.5)	42,868 (25.9)
	No insulin	96 (0.6)	604 (0.8)	373 (6.7)	1,284 (3.5)	3,895 (6.0)	14,735 (2.7)	12,020 (18.2)	125,021 (13.1)
	Overall	158 (0.8)	787 (0.9)	483 (7.4)	1,510 (3.7)	5,472 (7.2)	20,390 (3.4)	17,027 (21.0)	167,889 (15.0)
Retinopathy	Insulin	1,392 (45.6)	3,291 (42.6)	85 (8.7)	237 (5.7)	5,496 (52.6)	24,675 (46.6)	**10,762 (72.1)**	**95,926 (57.8)**
	No insulin	4,372 (27.4)	19,309 (25.6)	157 (2.8)	845 (2.3)	**19,283 (29.6)**	**93,797 (17.3)**	**30,889 (46.8)**	**322,383 (33.7)**
	Overall	5,764 (30.3)	22,600 (27.2)	242 (3.7)	1,082 (2.7)	**24,779 (32.8)**	**118,472 (19.9)**	**41,651 (51.4)**	**418,309 (37.3)**
Peripheral neuropathy	Insulin	247 (8.1)	524 (6.8)	61 (6.3)	208 (5.0)	275 (2.6)	1,123 (2.1)	1,024 (6.9)	12,678 (7.6)
	No insulin	394 (2.5)	2,143 (2.8)	242 (4.3)	1,095 (3.0)	915 (1.4)	6,642 (1.2)	3,749 (5.7)	57,202 (6.0)
	Overall	641 (3.4)	2,667 (3.2)	303 (4.6)	1,303 (3.2)	1,190 (1.6)	7,765 (1.3)	4,773 (5.9)	69,880 (6.2)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Peripheral vascular disease[b]	Insulin	188 (6.2)	543 (7.0)	60 (6.1)	153 (3.7)	5,362 (51.3)	24,121 (45.6)	**10,654 (71.4)**	**97,899 (59.0)**
	No insulin	463 (2.9)	3,619 (4.8)	318 (5.7)	831 (2.3)	**19,678 (30.2)**	**95,260 (17.6)**	**32,813 (49.7)**	**355,385 (37.2)**
	Overall	651 (3.4)	4,162 (5.0)	378 (5.8)	984 (2.4)	**25,040 (33.2)**	**119,381 (20.1)**	**43,467 (53.7)**	**453,284 (40.4)**
Coronary heart disease	Insulin	524 (17.2)	1,753 (22.7)	**479 (49.1)**	**983 (23.6)**	2,116 (20.2)	8,037 (15.2)	6,780 (45.4)	66,859 (40.3)
	No insulin	2,094 (13.1)	13,014 (17.3)	**2,502 (44.8)**	**6,664 (18.3)**	**11,092 (17.0)**	**46,001 (8.5)**	27,657 (41.9)	319,150 (33.4)
	Overall	2,618 (13.8)	14,767 (17.8)	**2,981 (45.4)**	**7,647 (18.8)**	**13,208 (17.5)**	**54,038 (9.1)**	34,437 (42.5)	386,009 (34.4)
Cerebrovascular disease	Insulin	173 (5.7)	669 (8.7)	97 (9.9)	318 (7.6)	605 (5.8)	2,423 (4.6)	2,868 (19.2)	31,702 (19.1)
	No insulin	770 (4.8)	6,157 (8.2)	524 (9.4)	1,761 (4.8)	2,582 (4.0)	13,344 (2.5)	11,937 (18.1)	153,216 (16.0)
	Overall	943 (5.0)	6,826 (8.2)	621 (9.5)	2,079 (5.1)	3,187 (4.2)	15,767 (2.7)	14,805 (18.3)	184,918 (16.5)
Amputation	Insulin	43 (1.4)	146 (1.9)	4 (< 0.5)	28 (0.7)	227 (2.2)	1,049 (2.0)	368 (2.5)	4,468 (2.7)
	No insulin	100 (0.6)	705 (0.9)	18 (< 0.5)	92 (< 0.5)	342 (0.5)	1,942 (0.4)	677 (1.0)	9,622 (1.0)
	Overall	143 (0.8)	851 (1.0)	22 (< 0.5)	120 (< 0.5)	569 (0.8)	2,991 (0.5)	1,045 (1.3)	14,090 (1.3)
Insulin use at the index date[c], n (%)	Insulin	NA	NA	NA	NA	NA	NA	NA	NA
	No insulin	NA	NA	NA	NA	NA	NA	NA	NA
	Overall	3,052 (16.1)	7,732 (9.3)	976 (14.9)	4,169 (10.3)	10,454 (13.8)	52,938 (8.9)	14,931 (18.4)	165,824 (14.8)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Index medication type, n (%)									
Index monotherapy	Insulin	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	No insulin	712 (4.5)	3,608 (4.8)	**2,090 (37.4)**	**1,517 (4.2)**	**11,984 (18.4)**	**154,871 (28.6)**	**14,620 (22.1)**	**132,213 (13.8)**
	Overall	712 (3.7)	3,608 (4.3)	**2,090 (31.8)**	**1,517 (3.7)**	**11,984 (15.9)**	**154,871 (26.1)**	14,620 (18.1)	132,213 (11.8)
Add-on index therapy	Insulin	**2,158 (70.7)**	**4,667 (60.4)**	511 (52.4)	2,217 (53.2)	8,149 (78.0)	41,706 (78.8)	9,471 (63.4)	100,272 (60.5)
	No insulin	9,438 (59.2)	51,770 (68.8)	**2,002 (35.8)**	**21,703 (59.6)**	38,256 (58.8)	265,514 (49.0)	30,521 (46.2)	463,534 (48.5)
	Overall	11,596 (61.1)	56,437 (68.0)	**2,513 (38.3)**	**23,920 (58.9)**	46,405 (61.4)	307,220 (51.7)	39,992 (49.4)	563,806 (50.3)
Switched-to index therapy	Insulin	22 (0.7)	208 (2.7)	53 (5.4)	400 (9.6)	10 (0.1)	102 (0.2)	717 (4.8)	11,573 (7.0)
	No insulin	775 (4.9)	7,180 (9.5)	**275 (4.9)**	**4,995 (13.7)**	1,596 (2.5)	11,273 (2.1)	6,556 (9.9)	157,346 (16.5)
	Overall	797 (4.2)	7,388 (8.9)	**328 (5.0)**	**5,395 (13.3)**	1,606 (2.1)	11,375 (1.9)	7,273 (9.0)	168,919 (15.1)
Combined index therapy with no prior treatment	Insulin	**14 (0.5)**	**255 (3.3)**	**10 (1.0)**	**263 (6.3)**	225 (2.2)	2,188 (4.1)	305 (2.0)	7,207 (4.3)
	No insulin	255 (1.6)	3,830 (5.1)	**222 (4.0)**	**4,455 (12.2)**	4,811 (7.4)	49,070 (9.1)	3,451 (5.2)	95,084 (10.0)
	Overall	269 (1.4)	4,085 (4.9)	**232 (3.5)**	**4,718 (11.6)**	5,036 (6.7)	51,258 (8.6)	3,756 (4.6)	102,291 (9.1)
Add-on and switched-to index therapy	Insulin	781 (25.6)	2,366 (30.6)	257 (26.3)	1,132 (27.2)	1,057 (10.1)	3,303 (6.2)	3,469 (23.2)	42,456 (25.6)
	No insulin	**4,341 (27.2)**	**6,686 (8.9)**	608 (10.9)	2,661 (7.3)	3,848 (5.9)	11,373 (2.1)	7,969 (12.1)	89,770 (9.4)
	Overall	**5,122 (27.0)**	**9,052 (10.9)**	865 (13.2)	3,793 (9.3)	4,905 (6.5)	14,676 (2.5)	11,438 (14.1)	132,226 (11.8)

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Non-evaluable[d]	Insulin	77 (2.5)	236 (3.1)	**145 (14.9)**	**157 (3.8)**	1,013 (9.7)	5,639 (10.7)	969 (6.5)	4,316 (2.6)
	No insulin	419 (2.6)	2,224 (3.0)	390 (7.0)	1,082 (3.0)	4,568 (7.0)	49,389 (9.1)	2,930 (4.4)	17,407 (1.8)
	Overall	496 (2.6)	2,460 (3.0)	**535 (8.2)**	**1,239 (3.1)**	5,581 (7.4)	55,028 (9.3)	3,899 (4.8)	21,723 (1.9)
No. of AD classes[e] used within 12 months before the index date[f], n (%)									
0	Insulin	5 (0.2)	129 (1.7)	5 (0.5)	109 (2.6)	162 (1.5)	1,704 (3.2)	82 (0.5)	3,270 (2.0)
	No insulin	518 (3.2)	4,813 (6.4)	**1,936 (34.7)**	**2,912 (8.0)**	**14,300 (22.0)**	**195,585 (36.1)**	11,294 (17.1)	123,482 (12.9)
	Overall	523 (2.8)	4,942 (6.0)	**1,941 (29.6)**	**3,021 (7.4)**	**14,462 (19.2)**	**197,289 (33.2)**	11,376 (14.0)	126,752 (11.3)
1-2	Insulin	**1,310 (42.9)**	**4,619 (59.7)**	**570 (58.4)**	**3,034 (72.8)**	**5,109 (48.9)**	**36,431 (68.8)**	**7,302 (48.9)**	**107,344 (64.7)**
	No insulin	**10,692 (67.1)**	**65,821 (87.4)**	**2,910 (52.1)**	**30,834 (84.7)**	40,757 (62.6)	326,743 (60.3)	**40,204 (60.9)**	**756,686 (79.2)**
	Overall	**12,002 (63.2)**	**70,440 (84.8)**	**3,480 (53.0)**	**33,868 (83.5)**	45,866 (60.7)	363,174 (61.1)	**47,506 (58.7)**	**864,030 (77.1)**
≥ 3	Insulin	**1,737 (56.9)**	**2,984 (38.6)**	**394 (40.4)**	**692 (16.6)**	**5,183 (49.6)**	**14,803 (28.0)**	**7,547 (50.5)**	**55,210 (33.3)**
	No insulin	**4,730 (29.7)**	**4,664 (6.2)**	**655 (11.7)**	**654 (1.8)**	**10,006 (15.4)**	**19,162 (3.5)**	**14,549 (22.0)**	**75,186 (7.9)**
	Overall	**6,467 (34.1)**	**7,648 (9.2)**	**1,049 (16.0)**	**1,346 (3.3)**	**15,189 (20.1)**	**33,965 (5.7)**	**22,096 (27.3)**	**130,396 (11.6)**

Characteristic	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of outpatient visits[g] in the 180 days before cohort entry, n (%)									
0-1	Insulin	665 (21.8)	1,687 (21.9)	172 (42.4)	760 (42.5)	1,307 (12.5)	8,274 (15.6)	1,571 (10.5)	20,943 (12.6)
	No insulin	4,166 (26.1)	19,092 (25.4)	1,313 (48.9)	7,246 (44.3)	12,667 (19.5)	116,490 (21.5)	7,833 (11.9)	142,644 (14.9)
	Overall	4,831 (25.4)	20,779 (25.0)	1,485 (48.1)	8,006 (44.1)	13,974 (18.5)	124,764 (21.0)	9,404 (11.6)	163,587 (14.6)
2-3	Insulin	567 (18.6)	1,232 (15.9)	58 (14.3)	249 (13.9)	2,078 (19.9)	10,479 (19.8)	2,383 (16.0)	28,744 (17.3)
	No insulin	3,633 (22.8)	15,264 (20.3)	405 (15.1)	2,584 (15.8)	15,415 (23.7)	127,652 (23.6)	12,096 (18.3)	197,521 (20.7)
	Overall	4,200 (22.1)	16,496 (19.9)	463 (15.0)	2,833 (15.6)	17,493 (23.2)	138,131 (23.2)	14,479 (17.9)	226,265 (20.2)
≥ 4	Insulin	1,820 (59.6)	4,813 (62.2)	176 (43.3)	778 (43.5)	7,069 (67.6)	34,185 (64.6)	10,977 (73.5)	116,137 (70.0)
	No insulin	8,141 (51.1)	40,942 (54.4)	966 (36.0)	6,531 (39.9)	36,981 (56.8)	297,348 (54.9)	46,118 (69.8)	615,189 (64.4)
	Overall	9,961 (52.4)	45,755 (55.1)	1,142 (37.0)	7,309 (40.3)	44,050 (58.3)	331,533 (55.8)	57,095 (70.5)	731,326 (65.2)
Hospitalisation in the 180 days before cohort entry, yes, n (%)	Insulin	443 (14.5)	1,487 (19.2)	140 (14.3)	729 (17.5)	1,326 (12.7)	6,913 (13.1)	1,878 (12.6)	28,385 (17.1)
	No insulin	1,393 (8.7)	9,678 (12.9)	**999 (17.9)**	**3,009 (8.3)**	5,012 (7.7)	23,127 (4.3)	6,680 (10.1)	105,721 (11.1)
	Overall	1,836 (9.7)	11,165 (13.4)	**1,139 (17.4)**	**3,738 (9.2)**	6,338 (8.4)	30,040 (5.1)	8,558 (10.6)	134,106 (12.0)

[a] The number of new users for each exposure group comprises those identified in each data source that met the eligibility criteria within each of the insulin use–stratified cohorts and overall cohorts. Unless otherwise specified, the values in the rows for "No. of new users" represent the denominators for each cohort and exposure group that were used in calculating the proportions of baseline characteristics presented throughout the remainder of the table.

[b] Includes peripheral artery disease.

[c] Frequencies and proportions of patients with insulin use at the index date are reported only in the overall cohorts. The "insulin use at the index date" variable was used to define the stratified insulin user and insulin non-user cohorts; therefore, this variable is reported as "NA" for the insulin use–stratified cohorts.

[d] Patients who did not have sufficient follow-up time to assess the 90-day add-on/switch requirement.

[e] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

[r] For PHARMO, users with insufficient lookback time in their medical record to assess this variable were categorised as "not applicable," as indicated in Appendix J, BladderCa Table 3 PHARMO, and include the following: insulin users cohort, n (%): dapagliflozin, **7 (0.7)**; comparator AD, **334 (8.0)**; insulin non-users cohort, n (%): dapagliflozin, **86 (1.5)**; comparator AD, **2,013 (5.5)**; overall cohort, n (%): dapagliflozin, **93 (1.4)**; comparator AD, **2,347 (5.8)**.

[g] Outpatient visits included GP and outpatient hospital visits. For PHARMO, the denominator for percentage calculations included only patients with available linked GP data (insulin users cohort: dapagliflozin, n = 406; comparator AD, n = 1,787; insulin non-users cohort: dapagliflozin, n = 2,684; comparator AD, n = 16,361; overall cohort: dapagliflozin, n = 3,090; comparator AD, n = 18,148).

Note: Values in **bold** in the table and the footnotes indicate a StDiff > 0.20 between new users of dapagliflozin and new users of comparator AD.

NA = not applicable; No. = number; StDiff = absolute standardised difference.

Most numbers and proportions for baseline variables for the overall and insulin use–stratified cohorts described in this section are presented in Table 13. Baseline variables with StDiff values > 0.20 are noted in the text below (and are bolded in Table 13) to highlight differences across the index exposure groups. Any proportions or StDiff values for baseline variables that are noted in the text below but are not presented in Table 13 include a reference to the relevant tables in Appendix J where these results can be found.

In the overall sex-combined bladder cancer cohorts before propensity score trimming, dapagliflozin new users were younger than comparator AD new users in CPRD (StDiff, 0.43), whilst in PHARMO, the HIRD, and Medicare, the mean age was similar amongst dapagliflozin new users and comparator AD new users. In CPRD and PHARMO, across the overall and insulin use–stratified cohorts, the percentage of males was higher than females in both exposure groups, whereas in the HIRD and Medicare, the percentage of males was higher than females only amongst dapagliflozin new users, and higher proportions of females than males were observed in the comparator AD new users.

In the overall cohorts, the prevalences of heart failure and chronic kidney disease were higher in dapagliflozin new users than in comparator AD new users, respectively, in PHARMO (StDiff, 0.52 and 0.21) and the HIRD (StDiff, 0.30 and 0.21) for each medical condition. In CPRD and Medicare, the prevalence of heart failure was similar in both exposure groups, and the prevalence of chronic kidney disease was higher in the comparator AD new user group than in the dapagliflozin new user group in CPRD (StDiff, 0.34) and similar in both exposure groups in Medicare. Across both exposure groups, amongst the indicators of diabetes severity, retinopathy was the most common in CPRD, coronary heart disease was the most common in PHARMO, and both retinopathy and peripheral vascular disease were the most common (with similar proportions) in the HIRD and Medicare. In PHARMO, indicators of diabetes severity were obtained from hospital diagnoses only. Retinopathy and peripheral vascular disease were more prevalent amongst dapagliflozin new users than comparator AD users in the HIRD (StDiff, 0.30 for both conditions) and Medicare (StDiff, 0.29 and 0.27, respectively). Coronary heart disease was also more prevalent amongst dapagliflozin new users than comparator AD new users in PHARMO (StDiff, 0.59) and the HIRD (StDiff, 0.25).

In CPRD and PHARMO, the proportions of new users who were severely obese (BMI ≥ 40), were current or former smokers, or had low-moderate alcohol intake or heavy or very heavy alcohol intake were similar in amongst dapagliflozin new users and comparator AD new users (Appendix J, BreastCa Table 1 CPRD and BreastCa Table 1 PHARMO). In PHARMO, information on these lifestyle-related variables was based on a subset of patients with available GP data.

Across all data sources, dapagliflozin and comparator AD index medications were most commonly initiated as add-on therapy; in PHARMO, the proportion initiating a comparator

AD as add-on therapy was higher than the proportion initiating dapagliflozin as add-on therapy (StDiff, 0.42). Dapagliflozin and comparator AD index medications were also frequently initiated as index monotherapy in the HIRD and Medicare, with higher proportions amongst comparator AD new users than dapagliflozin new users in the HIRD (StDiff, 0.25) and similar proportions observed in both exposure groups in Medicare; in PHARMO, only dapagliflozin index medications were frequently initiated as monotherapy, with few comparator AD new users initiating the index therapy as monotherapy (StDiff, 0.79). Dapagliflozin was also more frequently initiated as add-on and switched-to therapy than comparator AD index medications in CPRD (StDiff, 0.42).

In the overall sex-combined cohorts, differences were observed in all data sources between new users of dapagliflozin and comparator AD new users in some categories of the variables measuring the number of AD classes used in time periods before the index date, with StDiff values up to 0.69. In all data sources, dapagliflozin new users were more likely than comparator AD new users to have used at least three AD classes before the index date (Table 13 displays the time period of 12 months before the index date for each data source; Appendix J, BladderCa Table 3, shows all measured time periods before the index date for each data source).

In the overall cohorts in CPRD, the HIRD, and Medicare, approximately 50% or more of new users of dapagliflozin and of comparator ADs had at least four outpatient visits, with percentages lower in PHARMO (approximately 40%) in both exposure groups, which is based on a subset of patients with available GP data. Across all data sources, the proportion of new users of dapagliflozin and of comparator ADs with no hospitalisations in the 180 days before the index date ranged from 83% to 95%. The proportion of new users with at least one hospitalisation in the 180 days before the index date was similar between the exposure groups in CPRD, the HIRD, and Medicare; in PHARMO, a higher proportion of comparator AD new users than dapagliflozin new users had no hospitalisation in the 180 days before the index date (StDiff, 0.24) (Appendix J, BladderCa Table 4 PHARMO).

Insulin Use at the Index Date

In all data sources, similar proportions of dapagliflozin new users and comparator AD new users had concomitant insulin use at the index date. In CPRD, the HIRD and Medicare, a similar baseline prevalence of heart failure was observed in the insulin users cohorts compared with the insulin non-users cohorts amongst the dapagliflozin new user group, whilst the prevalence of heart failure in the comparator AD new user group was higher in the insulin users cohorts than in the insulin non-users cohorts. In PHARMO, the HIRD, and Medicare, amongst insulin non-users, heart failure was more prevalent in the dapagliflozin new user group than in the comparator AD new user group (StDiff, 0.56 in PHARMO, 0.32 in the HIRD, and 0.23 in Medicare). In PHARMO, amongst insulin users, heart failure was more prevalent in the dapagliflozin new user group than in the comparator AD new user group

(StDiff, 0.27). In CPRD, the prevalence of chronic kidney disease amongst insulin users and amongst insulin non-users, respectively, was higher in the comparator AD new user group than in the dapagliflozin new user group (StDiff, 0.38 and 0.34), whilst in PHARMO, the prevalence was higher in the dapagliflozin new user group than in the comparator AD new user group among insulin users (StDiff, 0.24). In the HIRD, chronic kidney disease had a similar prevalence in both exposure groups amongst insulin users and a higher prevalence in the dapagliflozin new user group than in the comparator AD new user group amongst insulin non-users (StDiff, 0.21). In Medicare, the prevalence of chronic kidney disease was similar across both exposure groups within each of the insulin use–stratified cohorts.

Across data sources, in both exposure groups, a higher proportion of insulin users than insulin non-users had recorded diagnoses for indicators of diabetes severity. In CPRD, the HIRD, and Medicare, for both exposure groups, differences between insulin users and insulin non-users were most pronounced for the baseline prevalence of retinopathy (CPRD, the HIRD, and Medicare) and for the baseline prevalence of peripheral vascular disease (the HIRD and Medicare).

Amongst insulin users in CPRD, dapagliflozin was more frequently initiated as an add-on therapy than comparator AD index medications (StDiff, 0.22); and amongst insulin non-users in CPRD, dapagliflozin was more frequently initiated as an add-on and switched-to index therapy than comparator AD new use (StDiff, 0.49). In PHARMO, comparator AD index medications were more likely than dapagliflozin to be initiated as a combined index therapy with no prior treatment amongst insulin users (StDiff, 0.28) and insulin non-users (StDiff, 0.31), and as add-on-index therapy (StDiff, 0.49) and switched-to index therapy (StDiff, 0.31) amongst insulin non-users. In the HIRD, compared with dapagliflozin, comparator AD medications were more frequently initiated as monotherapy amongst insulin non-users (StDiff, 0.24); the opposite was observed in PHARMO (StDiff, 0.90) and Medicare (StDiff, 0.22).

In all data sources, higher proportions of insulin non-users than insulin users had no other AD classes used across the different time periods before the index date. Amongst insulin non-users, the proportions with no other AD classes before the index date were higher amongst comparator AD new users than amongst dapagliflozin new users, particularly in CPRD when assessing the number of other AD classes used in the 12 to 24 months and in the more than 24 months before the index date, respectively (StDiff range, 0.25 to 0.27) and in the HIRD across all assessed time periods before the index date (StDiff range, 0.22 to 0.32); whilst in PHARMO the opposite was observed, with higher proportions with no other AD classes before the index date amongst dapagliflozin new users than comparator AD new users when assessing other AD classes in the 12 months and in the 12 to 24 months before the index date (StDiff, 0.69 and 0.26, respectively) (Appendix J, BladderCa Table 3, for each data source).

In all data sources, for both exposure groups, insulin users were more likely than insulin non-users to have at least four outpatient visits in the 180 days before the index date. In all data sources except PHARMO, for both exposure groups, insulin users were also more likely than insulin non-users to have had a hospitalisation in the 180 days before the index date.

Females

Complete results for the female bladder cancer cohorts are displayed in Appendix J, BladderCa Table 1F (baseline demographic characteristics), BladderCa Table 2F (specific baseline medical conditions), BladderCa Table 3F (specific baseline medications), and BladderCa Table 4F (HCRU during the 180 days before the index date), for each data source.

Across all data sources, results of the baseline covariates in the female bladder cancer cohort were similar both in absolute magnitude and relative differences to those described for the female breast cancer cohort in Section 10.2.2.1. In this section, we describe only exceptions and notable differences that are relevant specifically to the female bladder cancer cohorts.

For the overall female bladder cancer cohort, the baseline prevalences of chronic or recurrent urinary tract infections, chronic or recurrent urinary cystitis, kidney stones, and bladder stones were similar amongst new users of dapagliflozin and new users of comparator ADs across all data sources. The baseline prevalence of kidney stones was highest in Medicare (9.7% in dapagliflozin new users and 8.1% in comparator AD new users), and the baseline prevalence of bladder stones was less than 0.6% across all data sources. The highest baseline prevalence of chronic or recurrent cystitis was in CPRD in both exposure groups (approximately 17% of patients).

In the HIRD, females were more likely to initiate a comparator AD than dapagliflozin as monotherapy (StDiff, 0.46) and were more likely to initiate dapagliflozin than a comparator AD as an add-on therapy (StDiff, 0.32). In addition, in the HIRD, female comparator AD new users were more likely than female dapagliflozin new users to report no prior use of other AD classes during the time periods before the index date (StDiff range, 0.33 to 0.51).

Males

Complete results for the male bladder cancer cohorts are displayed in Appendix J, BladderCa Table 1M (baseline demographic characteristics), BladderCa Table 2M (specific baseline medical conditions), BladderCa Table 3M (specific baseline medications), and BladderCa Table 4M (HCRU during the 180 days before the index date), for each data source.

In the overall male bladder cancer cohorts, dapagliflozin new users were younger than comparator AD new users in CPRD (StDiff, 0.36), whilst the opposite was observed in PHARMO (StDiff, 0.22). In the HIRD and Medicare, the mean age was similar in dapagliflozin and comparator AD new users. Across all data sources, the baseline prevalence

of urinary tract–related medical conditions in the male bladder cancer cohorts was similar for both exposure groups and was lower in males than in females.

In the bladder cancer cohorts, the baseline prevalence of heart failure and chronic kidney disease was in general higher amongst males than amongst females. In the male bladder cancer cohorts, the baseline prevalence of heart failure was similar in both exposure groups in CPRD and higher in dapagliflozin new users than in comparator AD new users in PHARMO (StDiff, 0.54), the HIRD (StDiff, 0.31), and Medicare (StDiff, 0.27). The baseline prevalence of chronic kidney disease was also higher in dapagliflozin new users than comparator AD new users in PHARMO (StDiff, 0.22), whilst the opposite was observed in CPRD, where the baseline prevalence was higher in comparator AD new users than in dapagliflozin new users (StDiff, 0.30). In the HIRD and Medicare, the baseline prevalence of chronic kidney disease was similar amongst both exposure groups.

Amongst the indicators of diabetes severity in males, the baseline prevalence of retinopathy was similar in both exposure groups in CPRD and was higher in dapagliflozin new users than in comparator AD users in the HIRD (StDiff, 0.22) and Medicare (StDiff, 0.28). Peripheral vascular disease had a higher baseline prevalence amongst dapagliflozin new users than comparator AD new users in the HIRD (StDiff, 0.22) and Medicare (StDiff, 0.27), as was coronary heart disease in the HIRD (StDiff, 0.23). In PHARMO, indicators of diabetes severity were obtained only from hospital diagnoses; the most common indicator of diabetes severity was coronary heart disease, which had a higher baseline prevalence amongst dapagliflozin new users than comparator AD new users (StDiff, 0.59).

In PHARMO, males were more likely to initiate dapagliflozin index medication than a comparator AD medication as an index monotherapy (Std Diff, 0.88) and less likely to initiate dapagliflozin index medication than comparator AD medication as an add-on therapy (StDiff, 0.49) or as a combined index therapy with no prior treatment (StDiff, 0.33).

Similar to females, differences between new users of dapagliflozin and new users of comparator ADs in the variables measuring the number of other AD classes used in time periods before the index date and the type of index therapy were observed in all data sources; StDiff values were up to 0.75 for some categories. New users of dapagliflozin were more likely than new users of comparator ADs to have had a higher number of other AD classes before the index date.

The results for HCRU during the 180 days before the index date in the male bladder cancer cohort were very similar to the HCRU results in the female cohorts for all data sources, as described in the female breast cancer cohort results in Section 10.2.2.1.

By Pioglitazone Use

Selected baseline characteristics for the sex-combined bladder cancer cohorts stratified by pioglitazone use at the index date are summarised in Table 14, and results of all assessed baseline characteristics (measured on or before the index date [see Figure 1]) are displayed in Appendix J, BladderCa Table 17 (baseline demographic characteristics), BladderCa Table 18 (specific baseline medical conditions), BladderCa Table 19 (specific baseline medications), and BladderCa Table 20 (HCRU during the 180 days before the index date), for each data source.

Table 14 Selected Baseline Characteristics of the Sex-Combined Bladder Cancer Cohort Before Propensity Score Trimming (Full Sample), by Pioglitazone Use at the Index Date and by Data Source

Characteristic	Pioglitazone use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of new users[a], n (%)	Pioglitazone	885 (4.7)	6,377 (7.7)	22 (0.3)	342 (0.8)	3,348 (4.4)	35,954 (6.0)	5,478 (6.8)	146,411 (13.1)
	No pioglitazone	18,107 (95.3)	76,653 (92.3)	6,541 (99.7)	40,240 (99.2)	72,169 (95.6)	558,474 (94.0)	75,500 (93.2)	974,767 (86.9)
Age (years), mean (SD)	Pioglitazone	**60.0 (8.77)**	**63.0 (11.10)**	**62.3 (9.39)**	**66.2 (10.94)**	54.7 (6.38)	54.3 (6.67)	72.0 (5.73)	71.9 (6.20)
	No pioglitazone	**58.7 (9.47)**	**63.5 (11.94)**	66.1 (10.82)	64.3 (11.74)	**54.0 (6.57)**	**52.6 (6.82)**	72.3 (5.89)	72.6 (6.63)
Sex									
Female	Pioglitazone	274 (31.0)	2,243 (35.2)	**7 (31.8)**	**143 (41.8)**	1,003 (30.0)	13,085 (36.4)	2,207 (40.3)	70,875 (48.4)
	No pioglitazone	7,293 (40.3)	32,111 (41.9)	**2,239 (34.2)**	**18,027 (44.8)**	**29,367 (40.7)**	**303,614 (54.4)**	36,635 (48.5)	534,395 (54.8)
Male	Pioglitazone	611 (69.0)	4,134 (64.8)	**15 (68.2)**	**199 (58.2)**	2,345 (70.0)	22,869 (63.6)	3,271 (59.7)	75,536 (51.6)
	No pioglitazone	10,814 (59.7)	44,542 (58.1)	**4,302 (65.8)**	**22,213 (55.2)**	**42,802 (59.3)**	**254,860 (45.6)**	38,865 (51.5)	440,372 (45.2)
Comorbidities, n (%)									
Heart failure	Pioglitazone	9 (1.0)	119 (1.9)	0 (0.0)	12 (3.5)	102 (3.0)	717 (2.0)	730 (13.3)	14,234 (9.7)
	No pioglitazone	684 (3.8)	3,878 (5.1)	**1,351 (20.7)**	**1,599 (4.0)**	**7,638 (10.6)**	**16,296 (2.9)**	18,043 (23.9)	156,577 (16.1)
Chronic kidney disease	Pioglitazone	**93 (10.5)**	**1,279 (20.1)**	1 (4.5)	20 (5.8)	237 (7.1)	1,695 (4.7)	1,472 (26.9)	31,051 (21.2)
	No pioglitazone	**1,416 (7.8)**	**14,756 (19.3)**	**652 (10.0)**	**1,795 (4.5)**	**6,508 (9.0)**	**21,224 (3.8)**	20,291 (26.9)	213,331 (21.9)

Characteristic	Pioglitazone use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Indicators of diabetes severity, n (%)									
Diabetic nephropathy or renal insufficiency	Pioglitazone	12 (1.4)	78 (1.2)	0 (0.0)	17 (5.0)	273 (8.2)	1,766 (4.9)	**1,314 (24.0)**	**22,093 (15.1)**
	No pioglitazone	146 (0.8)	709 (0.9)	483 (7.4)	1,493 (3.7)	5,199 (7.2)	18,624 (3.3)	15,713 (20.8)	145,796 (15.0)
Retinopathy	Pioglitazone	354 (40.0)	2,218 (34.8)	2 (9.1)	16 (4.7)	**1,283 (38.3)**	**8,462 (23.5)**	**3,156 (57.6)**	**54,021 (36.9)**
	No pioglitazone	5,410 (29.9)	20,382 (26.6)	240 (3.7)	1,066 (2.6)	**23,496 (32.6)**	**110,010 (19.7)**	**38,495 (51.0)**	**364,288 (37.4)**
Peripheral neuropathy	Pioglitazone	36 (4.1)	218 (3.4)	0 (0.0)	9 (2.6)	50 (1.5)	476 (1.3)	304 (5.5)	7,701 (5.3)
	No pioglitazone	605 (3.3)	2,449 (3.2)	303 (4.6)	1,294 (3.2)	1,140 (1.6)	7,289 (1.3)	4,469 (5.9)	62,179 (6.4)
Peripheral vascular disease[b]	Pioglitazone	25 (2.8)	284 (4.5)	0 (0.0)	8 (2.3)	**1,233 (36.8)**	**8,285 (23.0)**	**3,115 (56.9)**	**54,976 (37.5)**
	No pioglitazone	626 (3.5)	3,878 (5.1)	378 (5.8)	976 (2.4)	**23,807 (33.0)**	**111,096 (19.9)**	**40,352 (53.4)**	**398,308 (40.9)**
Coronary heart disease	Pioglitazone	87 (9.8)	838 (13.1)	3 (13.6)	52 (15.2)	349 (10.4)	3,002 (8.3)	1,837 (33.5)	41,581 (28.4)
	No pioglitazone	2,531 (14.0)	13,929 (18.2)	**2,978 (45.5)**	**7,595 (18.9)**	**12,859 (17.8)**	**51,036 (9.1)**	32,600 (43.2)	344,428 (35.3)
Cerebrovas-cular disease	Pioglitazone	39 (4.4)	368 (5.8)	0 (0.0)	16 (4.7)	100 (3.0)	903 (2.5)	875 (16.0)	19,467 (13.3)
	No pioglitazone	904 (5.0)	6,458 (8.4)	621 (9.5)	2,063 (5.1)	3,087 (4.3)	14,864 (2.7)	13,930 (18.5)	165,451 (17.0)
Amputation	Pioglitazone	1-4 (0.1-0.5)	52 (0.8)	0 (0.0)	0 (0.0)	26 (0.8)	187 (0.5)	52 (0.9)	1,399 (1.0)
	No pioglitazone	141 (0.8)	799 (1.0)	22 (< 0.5)	120 (< 0.5)	543 (0.8)	2,804 (0.5)	993 (1.3)	12,691 (1.3)
Insulin use at the index date, n (%)	Pioglitazone	131 (14.8)	628 (9.8)	**5 (22.7)**	**28 (8.2)**	524 (15.7)	3,846 (10.7)	808 (14.7)	15,903 (10.9)
	No pioglitazone	**2,921 (16.1)**	**7,104 (9.3)**	971 (14.8)	4,141 (10.3)	9,930 (13.8)	49,092 (8.8)	14,123 (18.7)	149,921 (15.4)

Characteristic	Pioglitazone use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Index medication type, n (%)									
Index monotherapy	Pioglitazone	0 (0.0)	121 (1.9)	0 (0.0)	38 (11.1)	**0 (0.0)**	**4,347 (12.1)**	**0 (0.0)**	**19,376 (13.2)**
	No pioglitazone	712 (3.9)	3,487 (4.5)	**2,090 (32.0)**	**1,479 (3.7)**	**11,984 (16.6)**	**150,524 (27.0)**	**14,620 (19.4)**	**112,837 (11.6)**
Add-on index therapy	Pioglitazone	**256 (28.9)**	**3,724 (58.4)**	6 (27.3)	86 (25.1)	**2,640 (78.9)**	**23,709 (65.9)**	2,773 (50.6)	72,078 (49.2)
	No pioglitazone	11,340 (62.6)	52,713 (68.8)	**2,507 (38.3)**	**23,834 (59.2)**	43,765 (60.6)	283,511 (50.8)	37,219 (49.3)	491,728 (50.4)
Switched-to index therapy	Pioglitazone	35 (4.0)	328 (5.1)	2 (9.1)	39 (11.4)	8 (0.2)	672 (1.9)	463 (8.5)	18,364 (12.5)
	No pioglitazone	762 (4.2)	7,060 (9.2)	**326 (5.0)**	**5,356 (13.3)**	1,598 (2.2)	10,703 (1.9)	6,810 (9.0)	150,555 (15.4)
Combined index therapy with no prior treatment	Pioglitazone	6 (0.7)	94 (1.5)	**1 (4.5)**	**55 (16.1)**	**101 (3.0)**	**3,129 (8.7)**	153 (2.8)	10,182 (7.0)
	No pioglitazone	**263 (1.5)**	**3,991 (5.2)**	**231 (3.5)**	**4,663 (11.6)**	4,935 (6.8)	48,129 (8.6)	3,603 (4.8)	92,109 (9.4)
Add-on and switched-to index therapy	Pioglitazone	**560 (63.3)**	**1,910 (30.0)**	**12 (54.5)**	**119 (34.8)**	**334 (10.0)**	**1,532 (4.3)**	**1,739 (31.7)**	**23,557 (16.1)**
	No pioglitazone	**4,562 (25.2)**	**7,142 (9.3)**	853 (13.0)	3,674 (9.1)	4,571 (6.3)	13,144 (2.4)	9,699 (12.8)	108,669 (11.1)
Non-evaluable[c]	Pioglitazone	28 (3.2)	200 (3.1)	1 (4.5)	5 (1.5)	265 (7.9)	2,565 (7.1)	**350 (6.4)**	**2,854 (1.9)**
	No pioglitazone	468 (2.6)	2,260 (2.9)	**534 (8.2)**	**1,234 (3.1)**	5,316 (7.4)	52,463 (9.4)	3,549 (4.7)	18,869 (1.9)

Characteristic	Pioglitazone use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of AD classes[d] used within 12 months before the index date[e], n (%)									
0	Pioglitazone	0 (0.0)	106 (1.7)	0 (0.0)	51 (14.9)	**87 (2.6)**	**6,212 (17.3)**	**40 (0.7)**	**15,922 (10.9)**
	No pioglitazone	523 (2.9)	4,836 (6.3)	**1,941 (29.7)**	**2,970 (7.4)**	**14,375 (19.9)**	**191,077 (34.2)**	11,336 (15.0)	110,830 (11.4)
1-2	Pioglitazone	**183 (20.7)**	**4,112 (64.5)**	**8 (36.4)**	**177 (51.8)**	**1,040 (31.1)**	**22,527 (62.7)**	**1,692 (30.9)**	**97,963 (66.9)**
	No pioglitazone	**11,819 (65.3)**	**66,328 (86.5)**	**3,472 (53.1)**	**33,691 (83.7)**	44,826 (62.1)	340,647 (61.0)	**45,814 (60.7)**	**766,067 (78.6)**
≥ 3	Pioglitazone	**702 (79.3)**	**2,159 (33.9)**	**14 (63.6)**	**80 (23.4)**	**2,221 (66.3)**	**7,215 (20.1)**	**3,746 (68.4)**	**32,526 (22.2)**
	No pioglitazone	**5,765 (31.8)**	**5,489 (7.2)**	**1,035 (15.8)**	**1,266 (3.1)**	**12,968 (18.0)**	**26,750 (4.8)**	**18,350 (24.3)**	**97,870 (10.0)**
No. of outpatient visits[f] in the 180 days before cohort entry, n (%)									
0-1	Pioglitazone	233 (26.3)	1,646 (25.8)	**5 (62.5)**	**75 (50.0)**	622 (18.6)	8,823 (24.5)	606 (11.1)	21,309 (14.6)
	No pioglitazone	4,598 (25.4)	19,133 (25.0)	1,480 (48.0)	7,931 (44.1)	13,352 (18.5)	115,941 (20.8)	8,798 (11.7)	142,278 (14.6)
2-3	Pioglitazone	179 (20.2)	1,331 (20.9)	0 (0.0)	25 (16.7)	858 (25.6)	9,452 (26.3)	1,016 (18.5)	32,459 (22.2)
	No pioglitazone	4,021 (22.2)	15,165 (19.8)	463 (15.0)	2,808 (15.6)	16,635 (23.1)	128,679 (23.0)	13,463 (17.8)	193,806 (19.9)
≥ 4	Pioglitazone	473 (53.4)	3,400 (53.3)	**3 (37.5)**	**50 (33.3)**	1,868 (55.8)	17,679 (49.2)	3,856 (70.4)	92,643 (63.3)
	No pioglitazone	9,488 (52.4)	42,355 (55.3)	1,139 (37.0)	7,259 (40.3)	42,182 (58.4)	313,854 (56.2)	53,239 (70.5)	638,683 (65.5)

Characteristic	Pioglitazone use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Hospitalisation in the 180 days before cohort entry, yes, n (%)	Pioglitazone	86 (9.7)	577 (9.0)	0 (0.0)	29 (8.5)	141 (4.2)	1,559 (4.3)	417 (7.6)	11,730 (8.0)
	No pioglitazone	1,750 (9.7)	10,588 (13.8)	**1,139 (17.4)**	**3,709 (9.2)**	6,197 (8.6)	28,481 (5.1)	8,141 (10.8)	122,376 (12.6)

[a] The number of new users for each exposure group comprises those identified in each data source that met the eligibility criteria within each of the pioglitazone use–stratified cohorts. Unless otherwise specified, the values in the rows for "No. of new users" represent the denominators for each cohort and exposure group that were used in calculating the proportions of baseline characteristics presented throughout the remainder of the table.

[b] Includes peripheral artery disease.

[c] Patients who did not have sufficient follow-up time to assess the 90-day add-on/switch requirement.

[d] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

[e] For PHARMO, users with insufficient lookback time in their medical record to assess this variable were categorised as "not applicable," as indicated in Appendix J, BladderCa Table 19 PHARMO and as follows: pioglitazone users cohort, n (%): dapagliflozin group, 0 (0.0); comparator AD group, 34 (9.9); pioglitazone non-users cohort, n (%): dapagliflozin group, **93 (1.4)**; comparator AD group, **2,313 (5.7)**.

[f] Outpatient visits included GP and outpatient hospital visits. For PHARMO, the denominator for percentage calculations included only patients with available linked GP data (pioglitazone users cohort: dapagliflozin group, n = 8; comparator AD group, n = 150; pioglitazone non-users cohort: dapagliflozin group, n = 3,082; comparator AD group, n = 17,998).

Note 1: For CPRD data, any cell with a count of 1-4, or any cell that allows a count of 1-4 to be derived from other reported cells or information, cannot be reported. To comply with these reporting requirements, values that could be used to derive small count sizes in other cells are presented as ranges in this table.

Note 2: Note: Values in **bold** in the table and the footnotes indicate a StDiff > 0.20 between new users of dapagliflozin and new users of comparator AD.

No. = number; StDiff = absolute standardised difference.

Most numbers and proportions for baseline variables for the pioglitazone use–stratified cohorts described in this section are presented in Table 14. Baseline variables with StDiff > 0.20 were noted in the text (and are bolded in Table 14) to highlight differences across the index exposure groups. Any proportions or StDiff values for baseline variables that are noted in the text below but are not presented in Table 14 include a reference to relevant tables in Appendix J where these results can be found.

For the pioglitazone users cohort, dapagliflozin new users were younger than comparator AD new users in CPRD (StDiff, 0.30) and PHARMO (StDiff, 0.38), whereas age was similar across exposure cohorts in the HIRD and Medicare. For pioglitazone non-users, dapagliflozin new users were also younger than comparator AD new users in CPRD (StDiff, 0.44), whilst the opposite was observed in HIRD (StDiff, 0.21). In all data sources and for both exposure groups, a higher proportion of males than females had pioglitazone use at the index date.

Differences in the baseline prevalence of heart failure amongst dapagliflozin new users and comparator AD new users in the pioglitazone non-user cohorts were similar to the differences observed in the overall sex-combined bladder cancer cohorts in all data sources (see Section 10.2.2.2). For chronic kidney disease, the baseline prevalence was higher in comparator AD new users than dapagliflozin new users for both the pioglitazone user (StDiff, 0.27) and pioglitazone non-user (StDiff, 0.34) cohorts in CPRD, whilst the opposite was observed amongst pioglitazone non-users in PHARMO (StDiff, 0.52) and in the HIRD (StDiff, 0.21); in Medicare, the baseline prevalence was similar across both pioglitazone use–stratified cohorts and exposure groups.

Amongst the most prevalent indicators of diabetes severity in each data source, in the pioglitazone user cohorts, dapagliflozin new users had a higher baseline prevalence of retinopathy than comparator AD new users in the HIRD (StDiff, 0.32) and Medicare (StDiff, 0.42), whereas in CPRD, the baseline prevalence of retinopathy was similar in both exposure groups. In the pioglitazone non-user cohorts, the baseline prevalence of retinopathy was higher amongst dapagliflozin new users than comparator AD new users in the HIRD (StDiff, 0.30) and Medicare (StDiff, 0.28). In the HIRD and Medicare, a highly prevalent diabetes severity indicator was peripheral vascular disease, and the baseline prevalence was higher amongst dapagliflozin new users than comparator AD new users both in both pioglitazone users and pioglitazone non-users (HIRD: StDiff, 0.30 in both pioglitazone use–stratified cohorts; Medicare: StDiff, 0.39 for pioglitazone users and 0.25 for pioglitazone non-users). The baseline prevalence of coronary heart disease in pioglitazone non-users was higher amongst dapagliflozin new users than comparator AD new users in PHARMO (StDiff, 0.60) and in the HIRD (StDiff, 0.26).

Amongst pioglitazone users, insulin use at the index date was higher in dapagliflozin new users than comparator AD new users in the HIRD (StDiff, 0.41), whilst there were similar

proportions of insulin users across the exposure groups amongst pioglitazone users in all other data sources. Amongst pioglitazone non-users, the proportion with insulin use at the index date was higher amongst dapagliflozin new users than comparator AD new users in CPRD (StDiff, 0.21), whilst there were similar proportions of insulin users across the exposure groups amongst pioglitazone non-users in all other data sources.

Amongst pioglitazone non-users, dapagliflozin and comparator AD use were both most commonly initiated as add-on therapy across all data sources, with a higher proportion in PHARMO that initiated comparator ADs as add-on therapy than those who initiated dapagliflozin as add-on therapy (StDiff, 0.43). Amongst pioglitazone users, dapagliflozin was most commonly initiated as both "add-on" and "switched-to" therapy in CPRD and PHARMO, with higher proportions of dapagliflozin new users than comparator AD new users (StDiff, 0.71 and 0.41, respectively, in each data source). Additionally, amongst pioglitazone users in PHARMO, comparator AD was more frequently initiated in combination with other ADs than dapagliflozin (StDiff, 0.39); amongst pioglitazone users in the HIRD and Medicare, dapagliflozin and comparator ADs were most commonly initiated as add-on therapy, with a higher proportion of dapagliflozin new users initiating as add-on therapy than comparator AD new users in the HIRD (StDiff, 0.29).

In all data sources, differences were observed between new users of dapagliflozin and new users of comparator AD by pioglitazone use at the index date in the variables measuring the number of AD classes used in time periods before the index date, with StDiff values up to 0.70 in some categories. Amongst both pioglitazone users and pioglitazone non-users in all data sources, dapagliflozin new users were more likely than comparator AD new users to have had at least three AD classes across all lookback time windows before the index date.

Regardless of pioglitazone use, 50% or more of dapagliflozin new users and of comparator AD new users had at least four outpatient visits in the 180 days before the index date in CPRD, the HIRD, and Medicare. In PHARMO, lower proportions were observed for pioglitazone users than for pioglitazone non-users amongst comparator AD new users. Across all data sources, regardless of pioglitazone use, the proportion of new users of dapagliflozin and of comparator ADs that had no hospitalisations in the 180 days before the index date ranged from 82.6% to 100%.

10.2.2.3 Female Composite Cancer

For the female composite cancer cohorts, complete results for baseline characteristics (measured before on the index date [see Figure 1]) for each data source are displayed in Appendix J, FemaleCompositeCa Table 1 (baseline demographic characteristics), FemaleCompositeCa Table 2 (specific baseline medical conditions), FemaleCompositeCa Table 3 (specific baseline medications), and FemaleCompositeCa Table 4 (HCRU during the 180 days before the index date).

For each data source, the female composite cancer cohort comprised mostly the same patients as those in the female breast cancer cohort. Therefore, the covariates for demographic characteristics, common medical conditions and diabetes severity indicators, medications, and HCRU in the female composite cancer cohorts were similar to those of patients in the female breast cancer cohorts described in Section 10.2.2.1.

The baseline medical conditions described in the female composite cancer cohorts that were not described for the primary cancer outcome cohorts are presented in Appendix J, FemaleCompositeCa Table 2, for each data source. In the overall female composite cancer cohorts, the prevalence of a recorded T2DM diagnosis before or on the index date was high for both exposure groups in CPRD (99%, both exposure groups) and Medicare (96%, dapagliflozin; 98%, comparator AD), with lower proportions in PHARMO (85%, dapagliflozin; 78%, comparator AD) and the HIRD (85%, dapagliflozin; 63%, comparator AD). The proportions with a T2DM diagnosis were similar across the exposure groups for all data sources, except for the HIRD with a StDiff of 0.51.

The baseline prevalences of comorbidities in female composite cancer cohorts assessed before or on the index date (ie, polycystic ovarian syndrome, colon polyps, Crohn's disease, ulcerative colitis, peptic ulcer disease, immunosuppressive diseases, *Helicobacter pylori* infection, and autoimmune diseases) were similar between dapagliflozin new users and comparator AD new users in each data source. Across the data sources, the baseline prevalences of these conditions were low, with the highest baseline prevalences observed for colon polyps in Medicare (24%) and the HIRD (14%) and autoimmune diseases in Medicare (8%).

For the covariate medications assessed in the 180 days before and on the index date in the female composite cancer cohorts that were not assessed in the primary cancer outcome cohorts (ie, unopposed oestrogen therapy, immunosuppressants, and inhaled corticosteroids), the proportions were similar across the exposure groups across all data sources (see Appendix J, FemaleCompositeCa Table 3, for each data source). The highest proportions with usage of these medications were observed (1) for immunosuppressants (including systemic steroids and excluding inhaled corticosteroids) in the HIRD (26%), Medicare (21%), CPRD (11%), and PHARMO (11%); and (2) for inhaled corticosteroids in CPRD (15%).

10.2.2.4 Male Composite Cancer

For the male composite cancer cohorts, complete results for baseline characteristics (measured before on the index date [see Figure 1]) for each data source are displayed in Appendix J, MaleCompositeCa Table 1 (baseline demographic characteristics), MaleCompositeCa Table 2 (specific baseline medical conditions), MaleCompositeCa Table 3 (specific baseline medications), and MaleCompositeCa Table 4 (HCRU during the 180 days before the index date).

In each data source, the male composite cancer cohort comprised mostly the same patients as those in the male bladder cancer cohort. Therefore, the covariates for demographic characteristics, common medical conditions and diabetes severity indicators, medications, and HCRU in the male composite cancer cohorts were similar to those of patients in the male bladder cancer cohorts described in Section 10.2.2.2.

The baseline medical conditions described in the male composite cancer cohorts that were not described for the primary cancer outcome cohorts are presented in Appendix J, MaleCompositeCa Table 2, for each data source. In the overall male composite cancer cohorts, the prevalence of a recorded T2DM diagnosis before or on the index date was high for both exposure groups in CPRD (99% in both exposure groups) and Medicare (94%, dapagliflozin; 98% comparator AD) with lower proportions in PHARMO (89%, dapagliflozin; 83% comparator AD) and the HIRD (84%, dapagliflozin; 83% comparator AD). The proportions with a T2DM diagnosis were similar across the exposure groups for all data sources, except for Medicare, for which the comparator AD group had a higher prevalence of T2DM than the dapagliflozin group (StDiff, 0.22).

The baseline prevalences of comorbidities in the male composite cancer cohorts assessed before or on the index date (ie, benign prostatic hyperplasia, colon polyps, Crohn's disease, ulcerative colitis, peptic ulcer disease, immunosuppressive diseases, *Helicobacter pylori* infection, and autoimmune diseases) were similar between dapagliflozin new users and comparator AD new users in each data source. Across the data sources, the baseline prevalences of most of these conditions were low, with the highest baseline prevalences observed for benign prostatic hyperplasia in Medicare (35%) and the HIRD (7%) and for colon polyps in Medicare (28%) and the HIRD (15%).

For the covariate medications assessed in the 180 days before and on the index date in the male composite cancer cohorts that were not assessed in the primary cancer outcome cohorts (ie, immunosuppressants and inhaled corticosteroids), the proportions were similar across the exposure groups across all data sources (see Appendix J, MaleCompositeCa Table 3, for each data source). The proportion of new users with usage of immunosuppressants (including systemic steroids and excluding inhaled corticosteroids) ranged from 6% (CPRD) to 17% (Medicare). The proportion of new users with usage of inhaled corticosteroids was 10% in CPRD, 2% in PHARMO, and < 2% in the HIRD and Medicare.

10.2.3 Propensity Score Results

10.2.3.1 Selection of Covariates for the Propensity Score Models

The predefined base set of covariates listed in Section 9.9.3.2.1 were included in the Cox model in all data sources. All other baseline covariates listed in Appendix J that were not already included in the base models were considered for entry into the propensity score models. *The covariates included in the propensity score models for each of the primary*

outcome cohorts, overall and stratified by insulin use at the index date, and also stratified by pioglitazone use at the index date for the sex-combined bladder cancer cohort only, are listed for each data source in Appendix I, Table 41 (female breast cancer), Table 42 (sex-combined bladder cancer), Table 43 (female bladder cancer), and Table 44 (male bladder cancer). For sex-combined bladder cancer, covariates included in the propensity score models for the cohorts stratified by pioglitazone use at the index date are listed for each data source except PHARMO in Appendix I, Table 45.

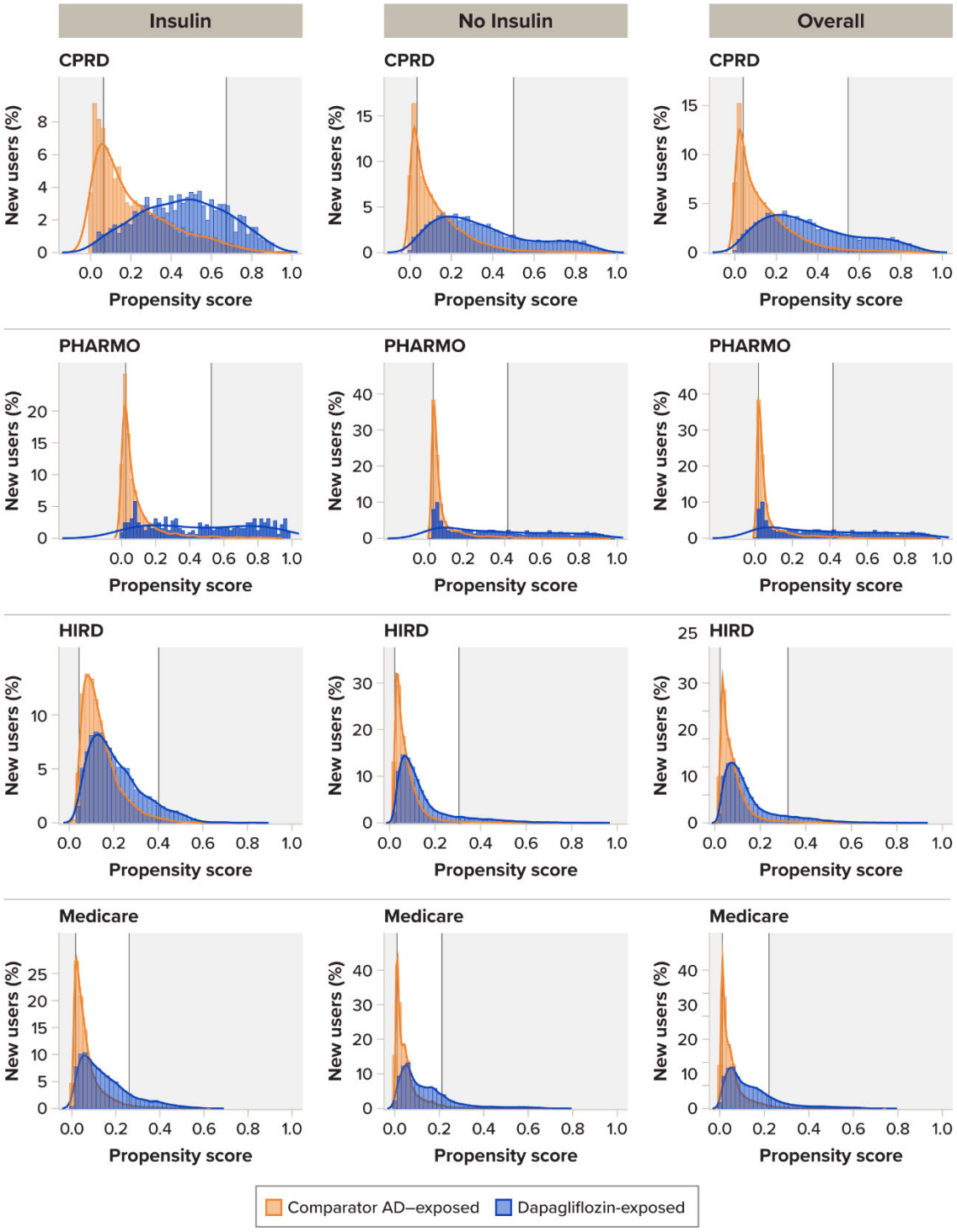
10.2.3.2 Propensity Score Distributions

For the primary cancer outcomes, Figure 9 and Figure 10 display propensity score distributions for each data source.

Female Breast Cancer. For the female breast cancer cohorts, overall and stratified by insulin use at the index date, there was adequate overlap ("common support") of the propensity score distributions for dapagliflozin new users and comparator AD new users in the HIRD and Medicare and somewhat more limited but adequate overlap in CPRD. In PHARMO, a right-skewed distribution was observed across the cohorts, overall and stratified by insulin use, with limited area of common support for dapagliflozin-exposed patients (Figure 9).

Bladder Cancer. For the sex-combined bladder cancer cohort and for both the female and male bladder cancer cohorts, overall and stratified by insulin use at the index date, there was adequate overlap of the propensity score distributions for dapagliflozin new users and comparator AD new users in the HIRD and Medicare and somewhat more limited but adequate overlap in CPRD (Figure 10). Furthermore, the distributions of propensity scores in the female bladder cancer cohorts were similar to those of the female breast cancer cohorts across CPRD, HIRD, and Medicare. For the PHARMO cohorts, overall and stratified by insulin use, there was limited area of common support in the propensity score distributions. For the sex-combined bladder cancer cohorts stratified by pioglitazone use at the index date, there was adequate overlap of the propensity score distributions for CPRD, the HIRD, and Medicare, similar to the cohort not stratified by pioglitazone use (Appendix J, BladderCa Figure 6). In PHARMO, propensity score estimation and trimming were not conducted in the cohorts stratified by pioglitazone use due to the small number of dapagliflozin new users and bladder cancer events amongst pioglitazone users.

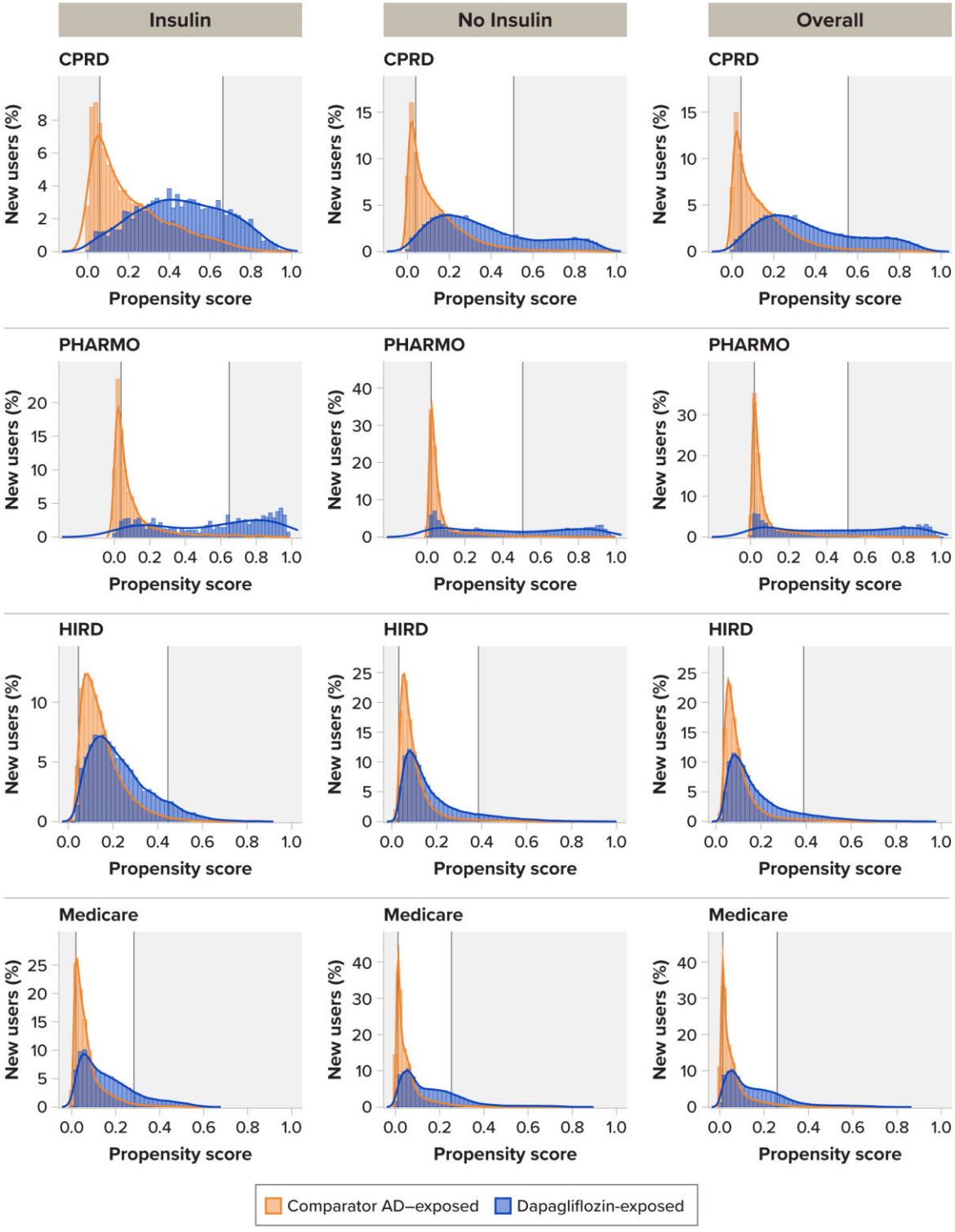
Figure 9 Distribution of Propensity Scores Amongst New Users of Dapagliflozin and of Comparator ADs in the Female Breast Cancer Cohort, Overall and Stratified by Insulin Use at the Index Date, by Data Source



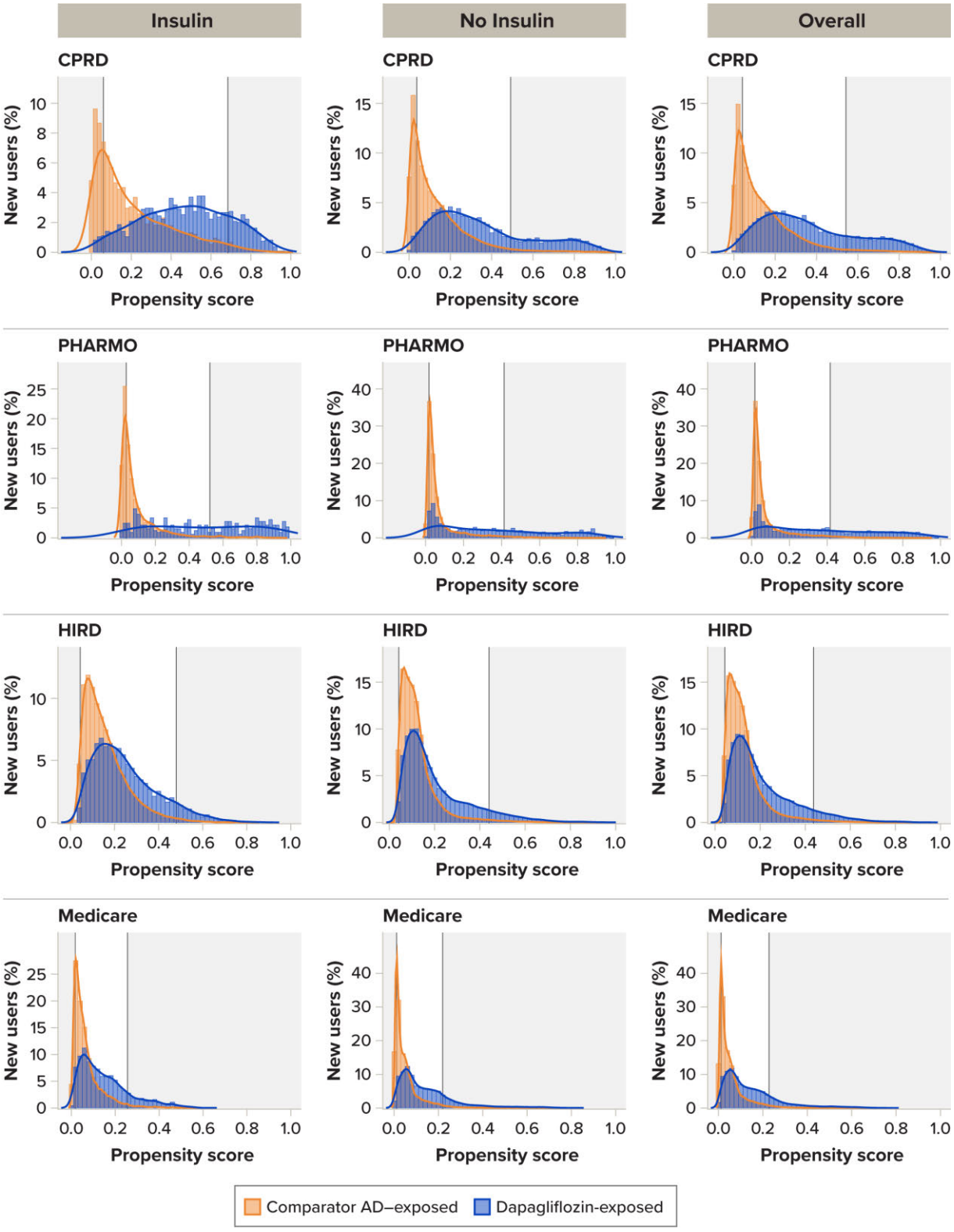
Note: The grey shading indicates propensity scores below the 2.5th percentile of the comparator AD group and above the 97.5th percentile of the dapagliflozin group that were trimmed out of each group.

Figure 10 **Distribution of Propensity Scores Amongst New Users of Dapagliflozin and of Comparator ADs in the Sex-Combined and Sex-Specific Bladder Cancer Cohorts, Overall and Stratified by Insulin Use at the Index Date, by Data Source**

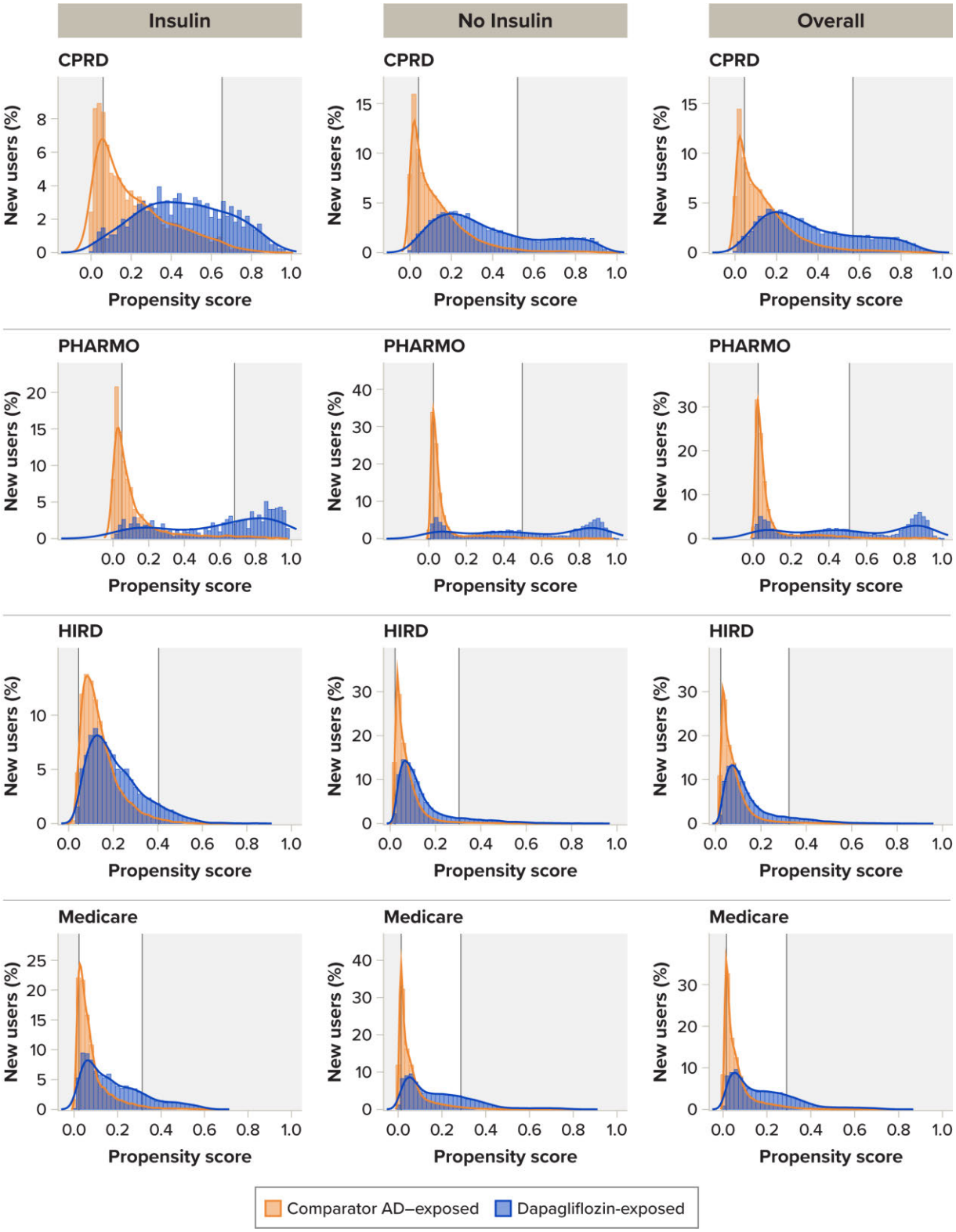
Sex-Combined Cohort



Female Cohort



Male Cohort



Note: The grey shading indicates propensity scores below the 2.5th percentile of the comparator AD group and above the 97.5th percentile of the dapagliflozin group that were trimmed out of each group.

Composite Cancer Cohorts. For both the female and male composite cancer cohorts, there was adequate overlap of the propensity score distributions (ie, there was adequate area of common support) for dapagliflozin new users and comparator AD new users across all data sources with the exception of PHARMO, where there was limited common support. See Appendix J, FemaleCompositeCa Figure 1 and MaleCompositeCa Figure 1, for each data source.

10.2.3.3 Propensity Score–Trimmed Analysis Samples

The final number of dapagliflozin new users and comparator AD new users included in the full (untrimmed) sample and the propensity score–trimmed analysis sample (ie, the number of new users in the full sample retained after propensity score trimming) for the female breast cancer and the sex-combined bladder cancer cohorts for all data sources is presented in Table 15. Across data sources, amongst new users in the overall female breast cancer cohorts, 61.4% to 83.8% of dapagliflozin new users and 69.1% to 84.6% of comparator AD new users were retained after propensity score trimming; amongst new users in the sex-combined bladder cancer cohorts, 50.3% to 84.1% of dapagliflozin new users and 68.2% to 89.2% of comparator AD new users were retained after propensity score trimming. The proportions retained in the cohorts stratified by insulin use at the index date differed slightly from those of the overall cohorts in all data sources.

Table 15 Number of Dapagliflozin and Comparator AD New Users Included in the Female Breast Cancer and the Bladder Cancer Cohorts, Full (Untrimmed) Samples and the Propensity Score–Trimmed Analysis Samples, Overall and Stratified by Insulin Use at the Index Date, by Data Source

Outcome cohort	Data source	Insulin use at the index date	Dapagliflozin new users		Comparator AD new users	
			Full (untrimmed) sample, n	PS-trimmed analysis sample, n (%)[a]	Full (untrimmed) sample, n	PS-trimmed analysis sample, n (%)[a]
Female breast cancer	CPRD	Insulin	1,356	1,082 (79.8%)	3,492	2,480 (71.0%)
		No insulin	6,232	4,534 (72.8%)	31,051	21,494 (69.2%)
		Overall	7,588	5,713 (75.3%)	34,543	23,877 (69.1%)
	PHARMO	Insulin	326	172 (52.8%)	1,954	1,271 (65.0%)
		No insulin	1,924	1,116 (58.0%)	16,239	12,712 (78.3%)
		Overall	2,250	1,382 (61.4%)	18,193	14,075 (77.4%)
	HIRD	Insulin	4,509	3,857 (85.5%)	25,154	22,598 (89.8%)
		No insulin	26,508	22,030 (83.1%)	297,336	244,927 (82.4%)
		Overall	31,017	25,988 (83.8%)	322,490	272,904 (84.6%)
	Medicare	Insulin	7,590	6,188 (81.5%)	93,914	78,096 (83.2%)
		No insulin	32,262	25,405 (78.7%)	529,963	425,138 (80.2%)
		Overall	39,852	31,656 (79.4%)	623,877	503,227 (80.7%)
Bladder cancer (females and males combined)	CPRD	Insulin	3,052	2,344 (76.8%)	7,732	5,634 (72.9%)
		No insulin	15,940	11,257 (70.6%)	75,298	50,940 (67.7%)
		Overall	18,992	14,050 (74.0%)	83,030	56,641 (68.2%)
	PHARMO	Insulin	976	499 (51.1%)	4,169	2,347 (56.3%)
		No insulin	5,587	2,908 (52.0%)	36,413	27,961 (76.8%)
		Overall	6,563	3,298 (50.3%)	40,582	30,402 (74.9%)
	HIRD	Insulin	10,454	8,807 (84.2%)	52,938	47,595 (89.9%)
		No insulin	65,063	54,233 (83.4%)	541,490	481,118 (88.9%)
		Overall	75,517	63,528 (84.1%)	594,428	530,047 (89.2%)
	Medicare	Insulin	14,931	12,003 (80.4%)	165,824	137,104 (82.7%)
		No insulin	66,047	50,325 (76.2%)	955,354	743,124 (77.8%)
		Overall	80,978	62,856 (77.6%)	1,121,178	884,302 (78.9%)

[a] The proportion of the untrimmed new users retained after propensity score trimming.

PS = propensity score.

10.2.3.4 Trimmed-Out Samples: Primary Outcome Cohorts

The "trimmed-out" sample (ie, the number of individuals in the full sample not included in the propensity score–trimmed sample) is the inverse of the propensity score–trimmed analysis sample shown in Table 15. Amongst new users in the overall female breast cancer cohorts, the proportion of new users trimmed out ranged from 16% to 39% for dapagliflozin new users and

15% to 31% for comparator AD new users. In the sex-combined overall bladder cancer cohort, the proportion of the sample trimmed out ranged from 16% to 50% for dapagliflozin new users and 11% to 32% for comparator AD new users. For the overall female breast cancer and the sex-combined bladder cancer cohorts, across all data sources, compared with users in the full (untrimmed) samples, users in the trimmed-out samples tended to have higher prevalence of heart failure and/or chronic kidney disease and higher prevalence of the most common indicators of diabetes severity, particularly amongst the dapagliflozin new users than those in the full sample. New users in the trimmed-out samples were also less likely than those in the full samples to initiate the index medication as an add-on therapy; more likely to have used at least three AD drug classes before the index date (although not uniformly seen in both exposure groups); and more likely to have at least four outpatient visits (HIRD and Medicare) or to have been hospitalised in the year before the index date (the HIRD) (in CPRD and PHARMO, new users in the trimmed-out samples were less likely to have at least four outpatient visits than in the full samples). Tables of results describing the baseline characteristics for the trimmed-out samples are available upon request.

10.2.4 Baseline Characteristics After Trimming: Primary Outcome Cohorts

10.2.4.1 Female Breast Cancer

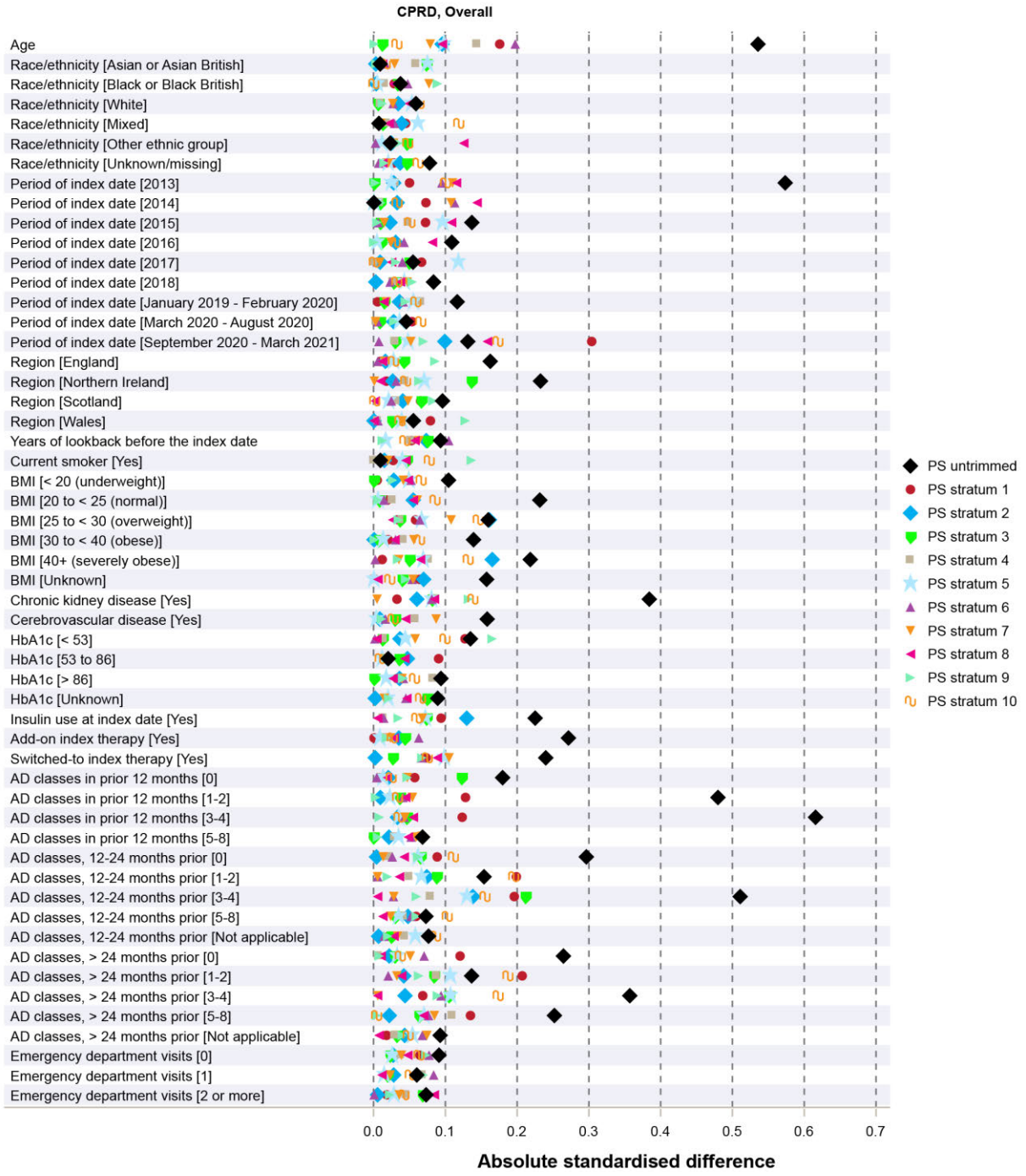
For the female breast cancer cohorts, baseline characteristics after propensity score trimming in the overall cohort and the cohorts stratified by insulin use at the index date are presented for each data source in Appendix J, BreastCa Table 1.2 (baseline demographic characteristics), BreastCa Table 2.2 (specific baseline medical conditions), BreastCa Table 3.2 (specific baseline medications), and BreastCa Table 4.2 (HCRU during the 180 days before the index date).

Figure 11 displays plots of the StDiff values for the covariates included in the propensity score models for the overall female breast cancer cohort for each data source and illustrates the balance of key covariates between the index treatment groups (dapagliflozin and comparator AD) before propensity score trimming and within propensity score strata after trimming. Overall, across data sources, propensity score trimming and stratification for the female breast cancer cohorts were effective in achieving balance between the index exposure groups for most variables in the propensity score models; StDiff values for most variables within most individual propensity score strata were ≤ 0.20 in all data sources, which correspond to small differences in the distribution of the variable between the dapagliflozin and comparator AD groups at the propensity score stratum level. In some data sources, after propensity score trimming, the imbalance of some variables amongst new users in some propensity score strata was higher (StDiff > 0.20) than the imbalance of the variable amongst new users in the full (untrimmed) sample (eg, in PHARMO and Medicare, some propensity score strata had a higher StDiff for age than the StDiff for age in the untrimmed sample). For the cohorts stratified by insulin use at the index date, plots of StDiff values before propensity score trimming and within propensity score strata after trimming are provided for each data source

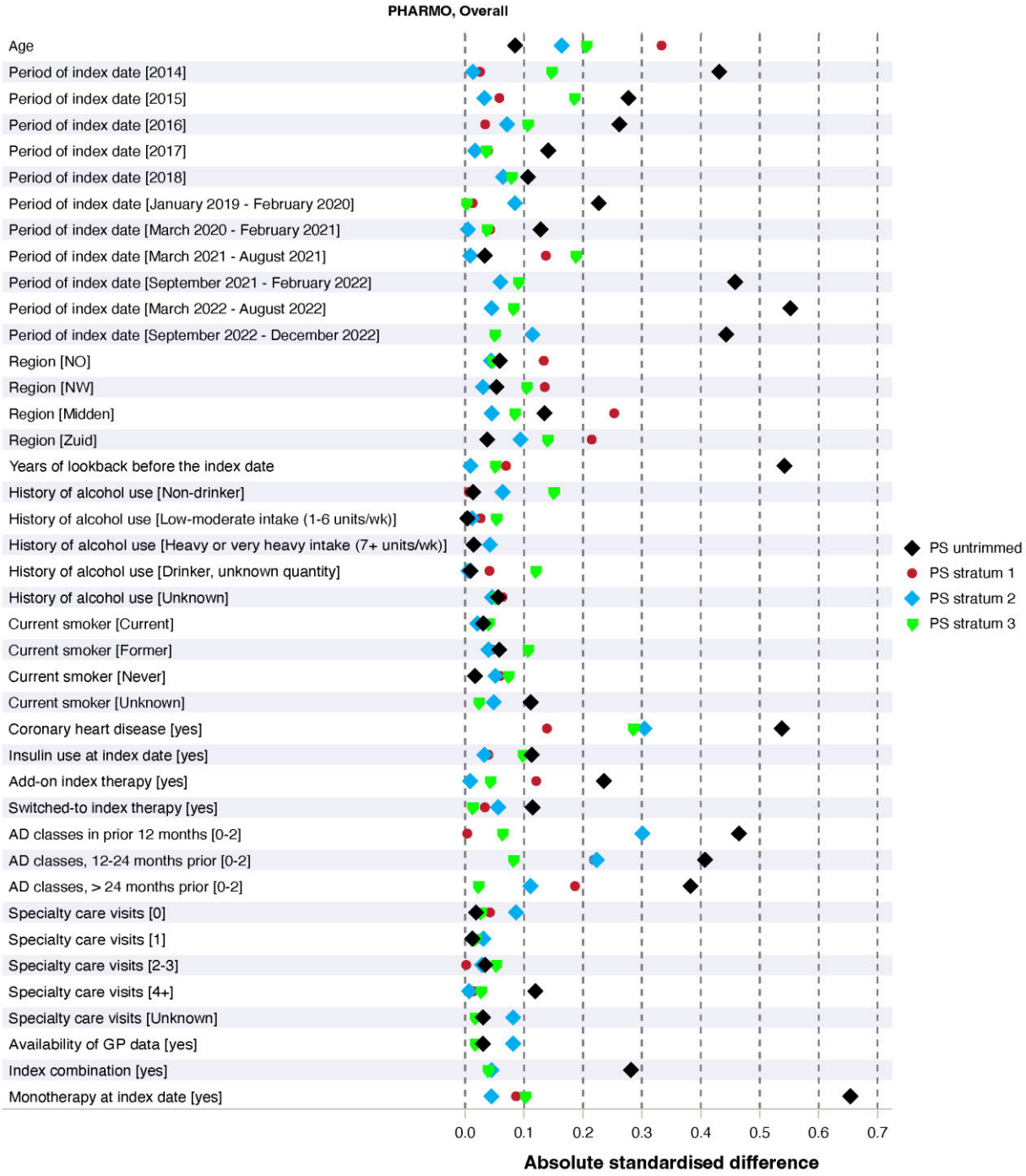
in Appendix J, BreastCa Figure 2, and the results are similar to the results in the overall cohort for all data sources.

Figure 11 Balance of Covariates in the Female Breast Cancer Cohort, Both in the Full Sample Before Trimming and Within Propensity Score Strata After Trimming, Overall Cohort, by Data Source

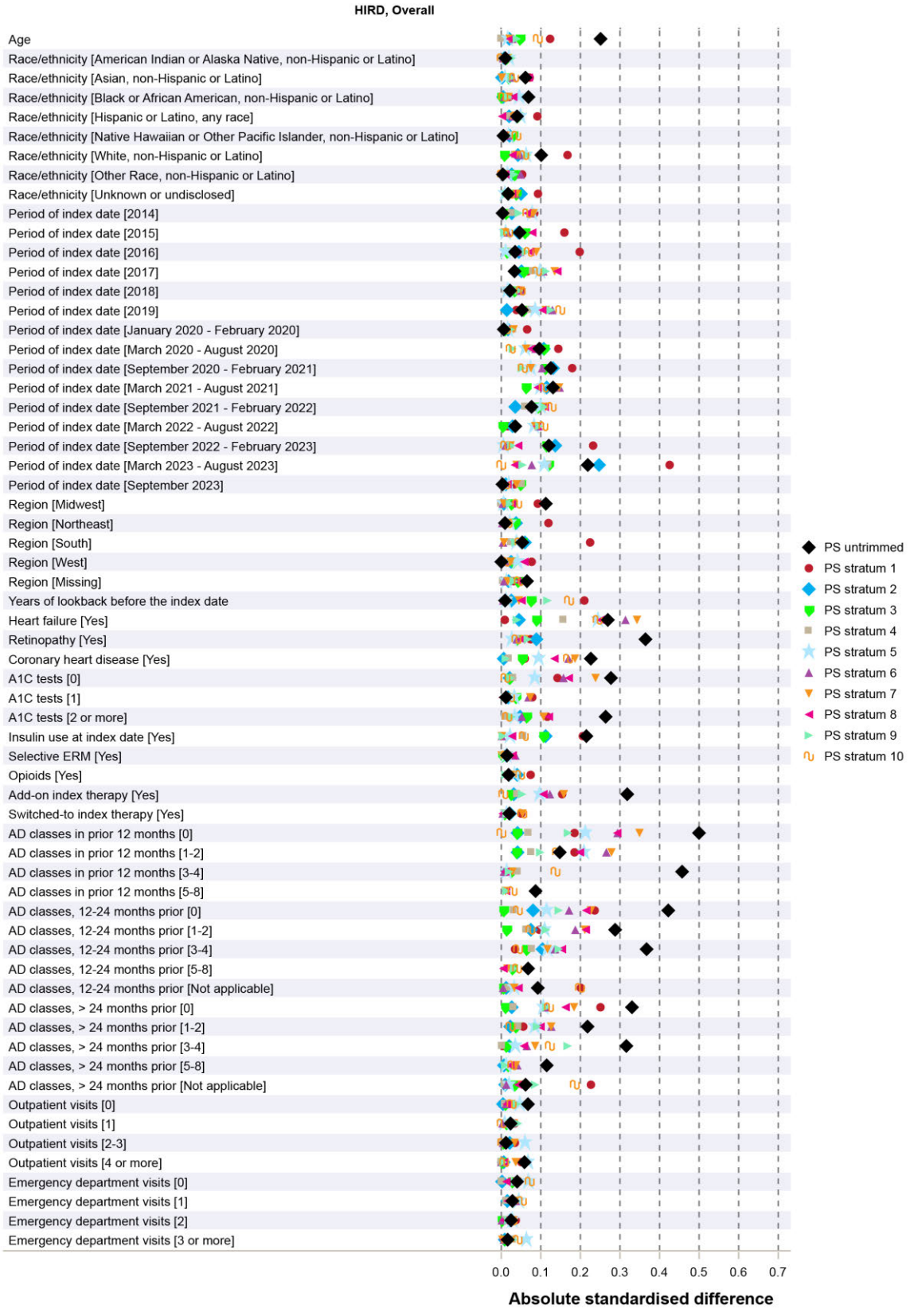
CPRD



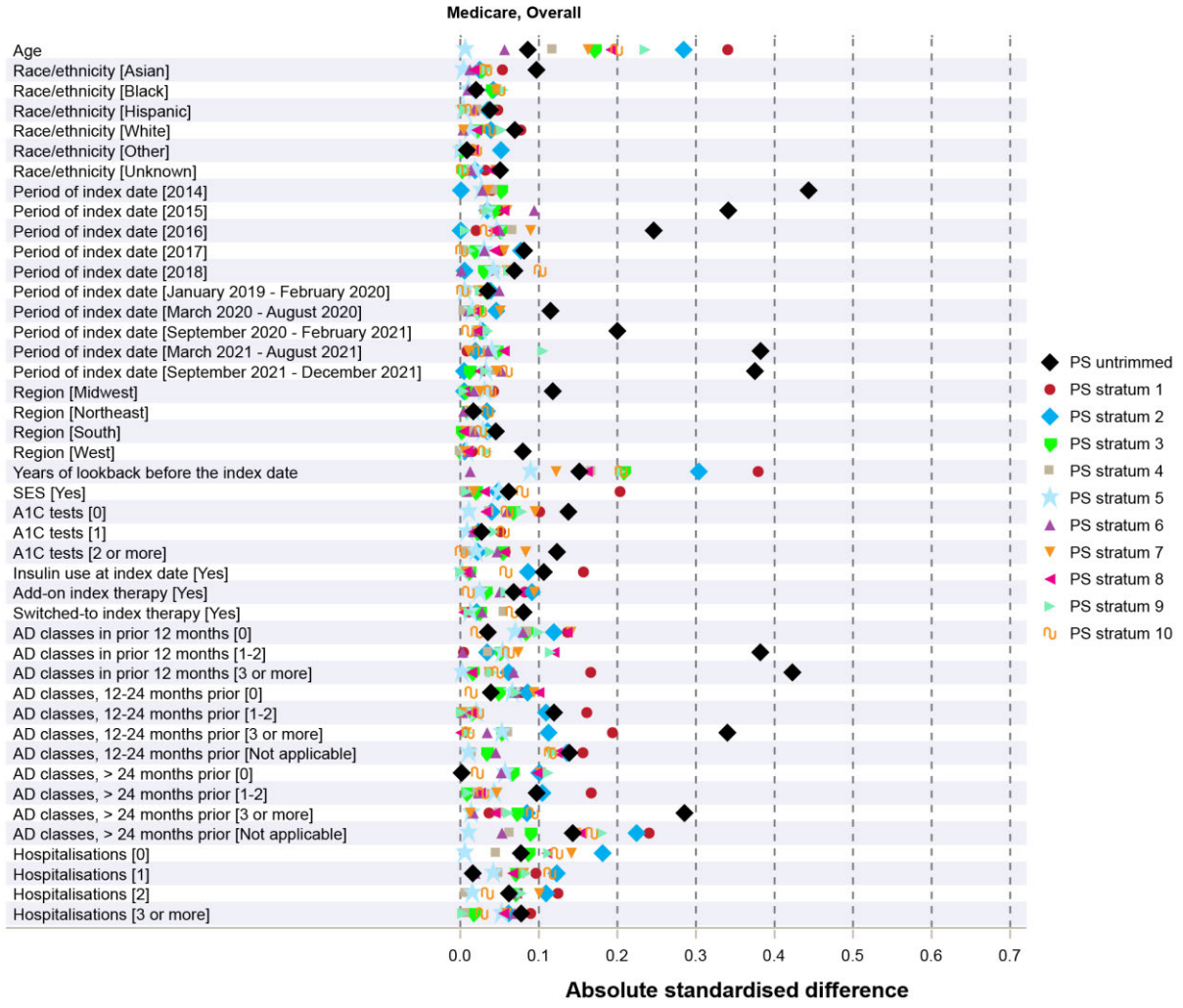
PHARMO


PHARMO, Overall
Age
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [January 2019 - February 2020]
Period of index date [March 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - December 2022]
Region [NO]
Region [NW]
Region [Midden]
Region [Zuid]
Years of lookback before the index date
History of alcohol use [Non-drinker]
History of alcohol use [Low-moderate intake (1-6 units/wk)]
History of alcohol use [Heavy or very heavy intake (7+ units/wk)]
History of alcohol use [Drinker, unknown quantity]
History of alcohol use [Unknown]
Current smoker [Current]
Current smoker [Former]
Current smoker [Never]
Current smoker [Unknown]
Coronary heart disease [yes]
Insulin use at index date [yes]
Add-on index therapy [yes]
Switched-to index therapy [yes]
AD classes in prior 12 months [0-2]
AD classes, 12-24 months prior [0-2]
AD classes, > 24 months prior [0-2]
Specialty care visits [0]
Specialty care visits [1]
Specialty care visits [2-3]
Specialty care visits [4+]
Specialty care visits [Unknown]
Availability of GP data [yes]
Index combination [yes]
Monotherapy at index date [yes]
PS untrimmed
PS stratum 1
PS stratum 2
PS stratum 3
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

HIRD


HIRD, Overall
Age
Race/ethnicity [American Indian or Alaska Native, non-Hispanic or Latino]
Race/ethnicity [Asian, non-Hispanic or Latino]
Race/ethnicity [Black or African American, non-Hispanic or Latino]
Race/ethnicity [Hispanic or Latino, any race]
Race/ethnicity [Native Hawaiian or Other Pacific Islander, non-Hispanic or Latino]
Race/ethnicity [White, non-Hispanic or Latino]
Race/ethnicity [Other Race, non-Hispanic or Latino]
Race/ethnicity [Unknown or undisclosed]
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [2019]
Period of index date [January 2020 - February 2020]
Period of index date [March 2020 - August 2020]
Period of index date [September 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - February 2023]
Period of index date [March 2023 - August 2023]
Period of index date [September 2023]
Region [Midwest]
Region [Northeast]
Region [South]
Region [West]
Region [Missing]
Years of lookback before the index date
Heart failure [Yes]
Retinopathy [Yes]
Coronary heart disease [Yes]
A1C tests [0]
A1C tests [1]
A1C tests [2 or more]
Insulin use at index date [Yes]
Selective ERM [Yes]
Opioids [Yes]
Add-on index therapy [Yes]
Switched-to index therapy [Yes]
AD classes in prior 12 months [0]
AD classes in prior 12 months [1-2]
AD classes in prior 12 months [3-4]
AD classes in prior 12 months [5-8]
AD classes, 12-24 months prior [0]
AD classes, 12-24 months prior [1-2]
AD classes, 12-24 months prior [3-4]
AD classes, 12-24 months prior [5-8]
AD classes, 12-24 months prior [Not applicable]
AD classes, > 24 months prior [0]
AD classes, > 24 months prior [1-2]
AD classes, > 24 months prior [3-4]
AD classes, > 24 months prior [5-8]
AD classes, > 24 months prior [Not applicable]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4 or more]
Emergency department visits [0]
Emergency department visits [1]
Emergency department visits [2]
Emergency department visits [3 or more]
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference
PS untrimmed
PS stratum 1
PS stratum 2
PS stratum 3
PS stratum 4
PS stratum 5
PS stratum 6
PS stratum 7
PS stratum 8
PS stratum 9
PS stratum 10

Medicare



ERM = oestrogen receptor modulator; PS = propensity score; SES = socioeconomic status.

10.2.4.2 Bladder Cancer

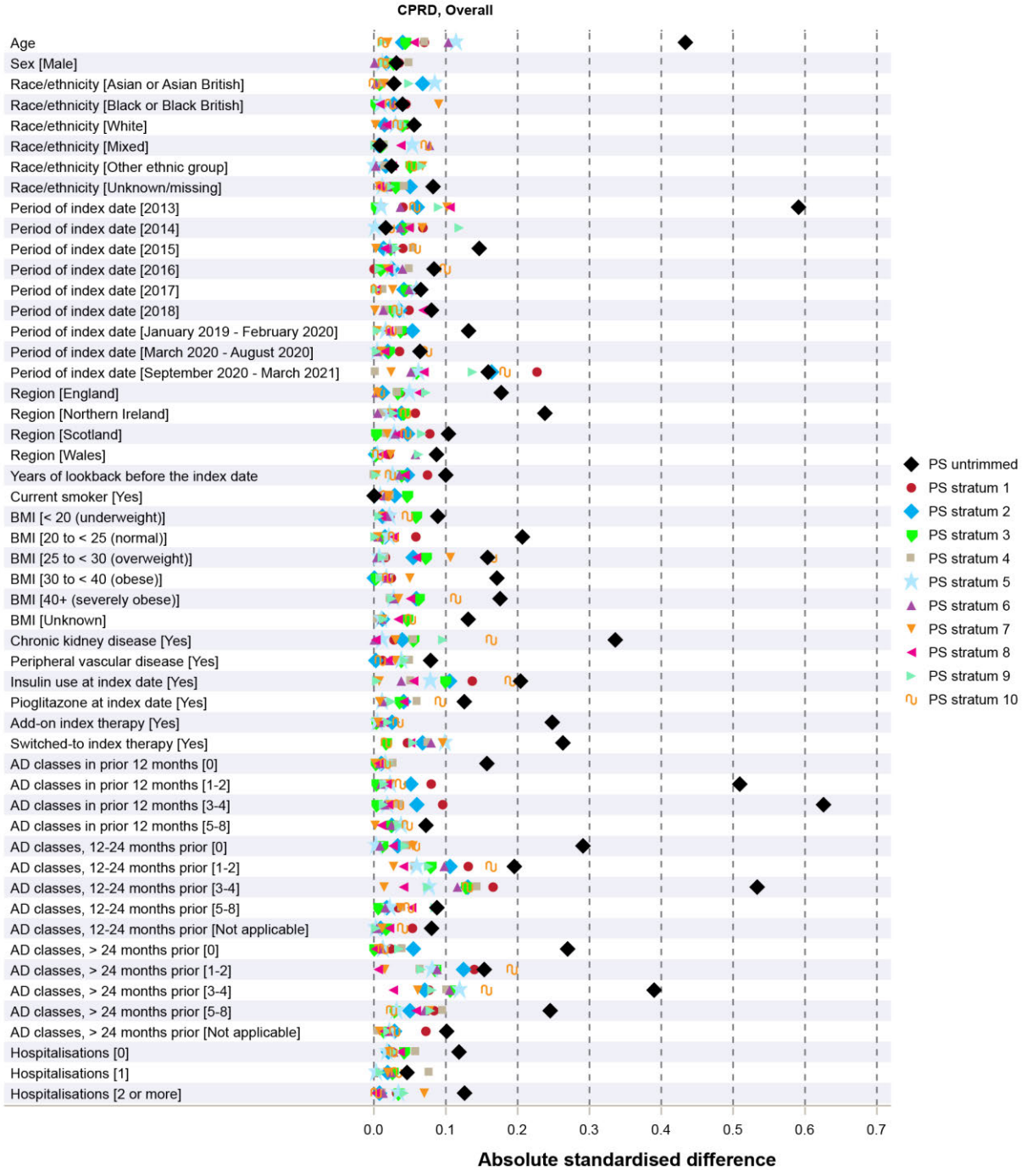
Sex-Combined Cohort

For the sex-combined bladder cancer cohort, baseline characteristics after propensity score trimming in the overall cohort and the cohorts stratified by insulin use at the index date are presented for each data source in Appendix J, BladderCa Table 1.2 (baseline demographic characteristics), BladderCa Table 2.2 (specific baseline medical conditions), BladderCa Table 3.2 (specific baseline medications), and BladderCa Table 4.2 (HCRU during the 180 days before the index date).

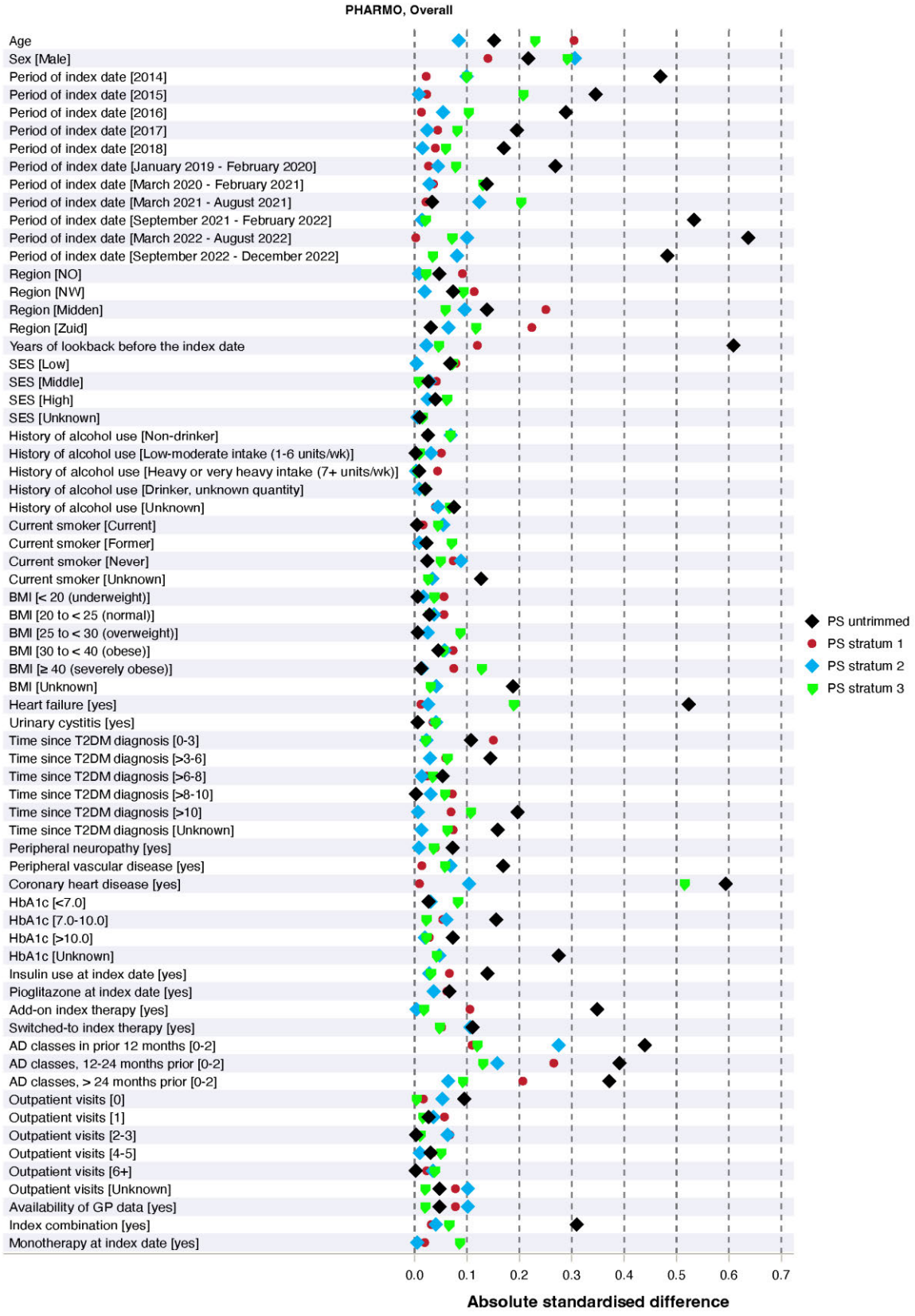
Figure 12 displays plots of the StDiff values for the overall sex-combined bladder cancer cohort for each data source to illustrate the balance of key covariates between the index treatment groups (dapagliflozin and comparator AD) before propensity score trimming and within propensity score strata after trimming. For the sex-combined bladder cancer cohort, propensity score trimming and stratification were effective in achieving balance between the index exposure groups for most variables in the propensity score model; StDiff values for most variables within most individual propensity score strata were ≤ 0.20 for all data sources, which correspond to small differences in the distribution of the variable between dapagliflozin and comparator AD groups at the propensity score stratum level. In some data sources, after propensity score trimming, the imbalance of some variables amongst new users in some propensity score strata was higher (StDiff > 0.20) than the imbalance of the variable amongst new users in the full (untrimmed) sample (eg, in PHARMO and Medicare, some propensity score strata had a higher StDiff for age than the StDiff for age in the untrimmed sample). For the cohorts stratified by insulin use at the index date, plots of StDiff values before propensity score trimming and within propensity score strata after trimming are provided for each data source in Appendix J, BladderCa Figure 2, and the results were similar to the results in the overall cohort for all data sources.

Figure 12 Balance of Covariates in the Sex-Combined Bladder Cancer Cohort, Both in the Full Sample Before Trimming and Within Propensity Score Strata After Trimming, Overall Cohort, by Data Source

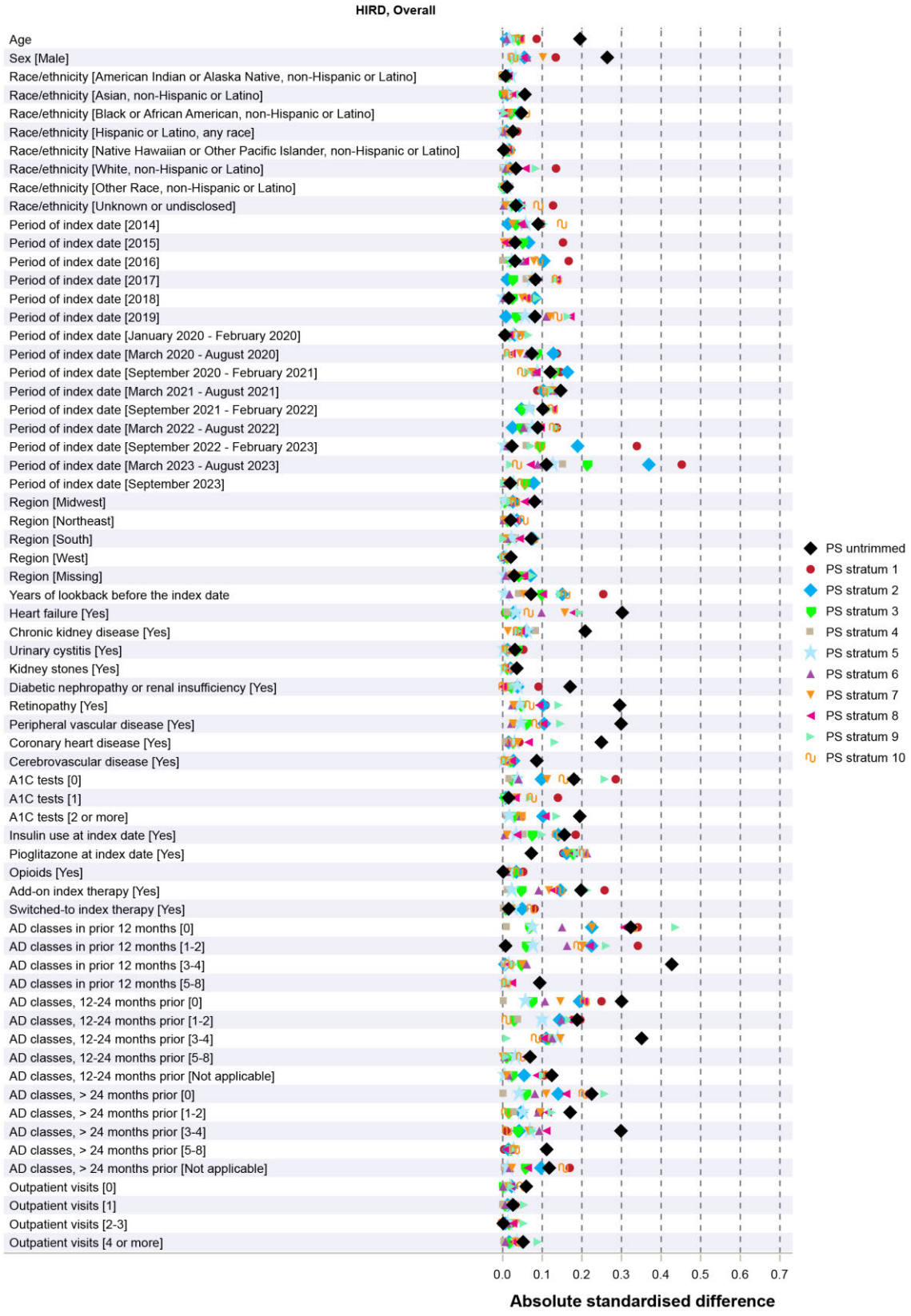
CPRD



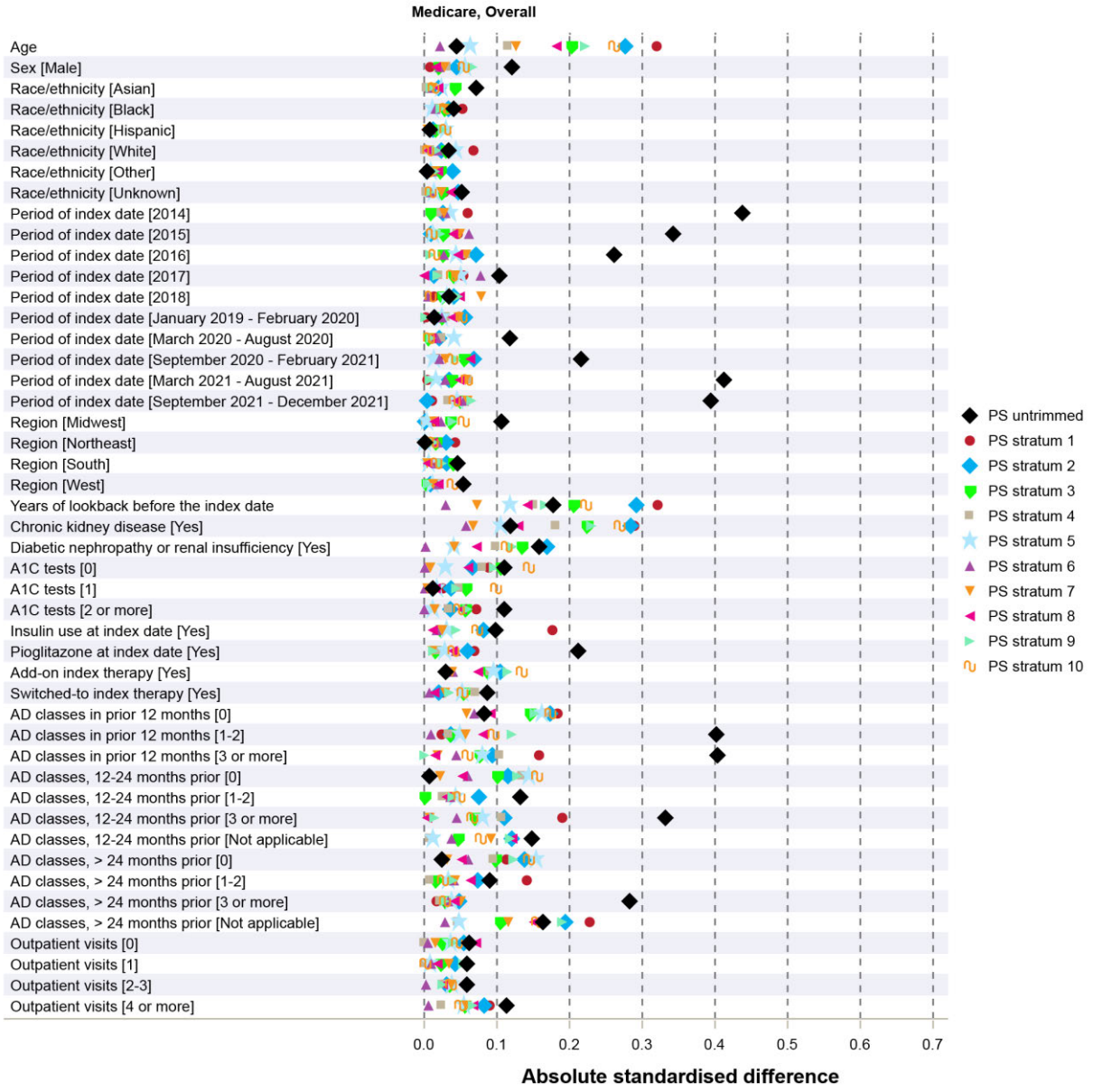
PHARMO


PHARMO, Overall
Age
Sex [Male]
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [January 2019 - February 2020]
Period of index date [March 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - December 2022]
Region [NO]
Region [NW]
Region [Midden]
Region [Zuid]
Years of lookback before the index date
SES [Low]
SES [Middle]
SES [High]
SES [Unknown]
History of alcohol use [Non-drinker]
History of alcohol use [Low-moderate intake (1-6 units/wk)]
History of alcohol use [Heavy or very heavy intake (7+ units/wk)]
History of alcohol use [Drinker, unknown quantity]
History of alcohol use [Unknown]
Current smoker [Current]
Current smoker [Former]
Current smoker [Never]
Current smoker [Unknown]
BMI [< 20 (underweight)]
BMI [20 to < 25 (normal)]
BMI [25 to < 30 (overweight)]
BMI [30 to < 40 (obese)]
BMI [≥ 40 (severely obese)]
BMI [Unknown]
Heart failure [yes]
Urinary cystitis [yes]
Time since T2DM diagnosis [0-3]
Time since T2DM diagnosis [>3-6]
Time since T2DM diagnosis [>6-8]
Time since T2DM diagnosis [>8-10]
Time since T2DM diagnosis [>10]
Time since T2DM diagnosis [Unknown]
Peripheral neuropathy [yes]
Peripheral vascular disease [yes]
Coronary heart disease [yes]
HbA1c [<7.0]
HbA1c [7.0-10.0]
HbA1c [>10.0]
HbA1c [Unknown]
Insulin use at index date [yes]
Pioglitazone at index date [yes]
Add-on index therapy [yes]
Switched-to index therapy [yes]
AD classes in prior 12 months [0-2]
AD classes, 12-24 months prior [0-2]
AD classes, > 24 months prior [0-2]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4-5]
Outpatient visits [6+]
Outpatient visits [Unknown]
Availability of GP data [yes]
Index combination [yes]
Monotherapy at index date [yes]
PS untrimmed
PS stratum 1
PS stratum 2
PS stratum 3
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

HIRD


HIRD, Overall
Age
Sex [Male]
Race/ethnicity [American Indian or Alaska Native, non-Hispanic or Latino]
Race/ethnicity [Asian, non-Hispanic or Latino]
Race/ethnicity [Black or African American, non-Hispanic or Latino]
Race/ethnicity [Hispanic or Latino, any race]
Race/ethnicity [Native Hawaiian or Other Pacific Islander, non-Hispanic or Latino]
Race/ethnicity [White, non-Hispanic or Latino]
Race/ethnicity [Other Race, non-Hispanic or Latino]
Race/ethnicity [Unknown or undisclosed]
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [2019]
Period of index date [January 2020 - February 2020]
Period of index date [March 2020 - August 2020]
Period of index date [September 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - February 2023]
Period of index date [March 2023 - August 2023]
Period of index date [September 2023]
Region [Midwest]
Region [Northeast]
Region [South]
Region [West]
Region [Missing]
Years of lookback before the index date
Heart failure [Yes]
Chronic kidney disease [Yes]
Urinary cystitis [Yes]
Kidney stones [Yes]
Diabetic nephropathy or renal insufficiency [Yes]
Retinopathy [Yes]
Peripheral vascular disease [Yes]
Coronary heart disease [Yes]
Cerebrovascular disease [Yes]
A1C tests [0]
A1C tests [1]
A1C tests [2 or more]
Insulin use at index date [Yes]
Pioglitazone at index date [Yes]
Opioids [Yes]
Add-on index therapy [Yes]
Switched-to index therapy [Yes]
AD classes in prior 12 months [0]
AD classes in prior 12 months [1-2]
AD classes in prior 12 months [3-4]
AD classes in prior 12 months [5-8]
AD classes, 12-24 months prior [0]
AD classes, 12-24 months prior [1-2]
AD classes, 12-24 months prior [3-4]
AD classes, 12-24 months prior [5-8]
AD classes, 12-24 months prior [Not applicable]
AD classes, > 24 months prior [0]
AD classes, > 24 months prior [1-2]
AD classes, > 24 months prior [3-4]
AD classes, > 24 months prior [5-8]
AD classes, > 24 months prior [Not applicable]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4 or more]
PS untrimmed
PS stratum 1
PS stratum 2
PS stratum 3
PS stratum 4
PS stratum 5
PS stratum 6
PS stratum 7
PS stratum 8
PS stratum 9
PS stratum 10
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

Medicare



PS = propensity score.

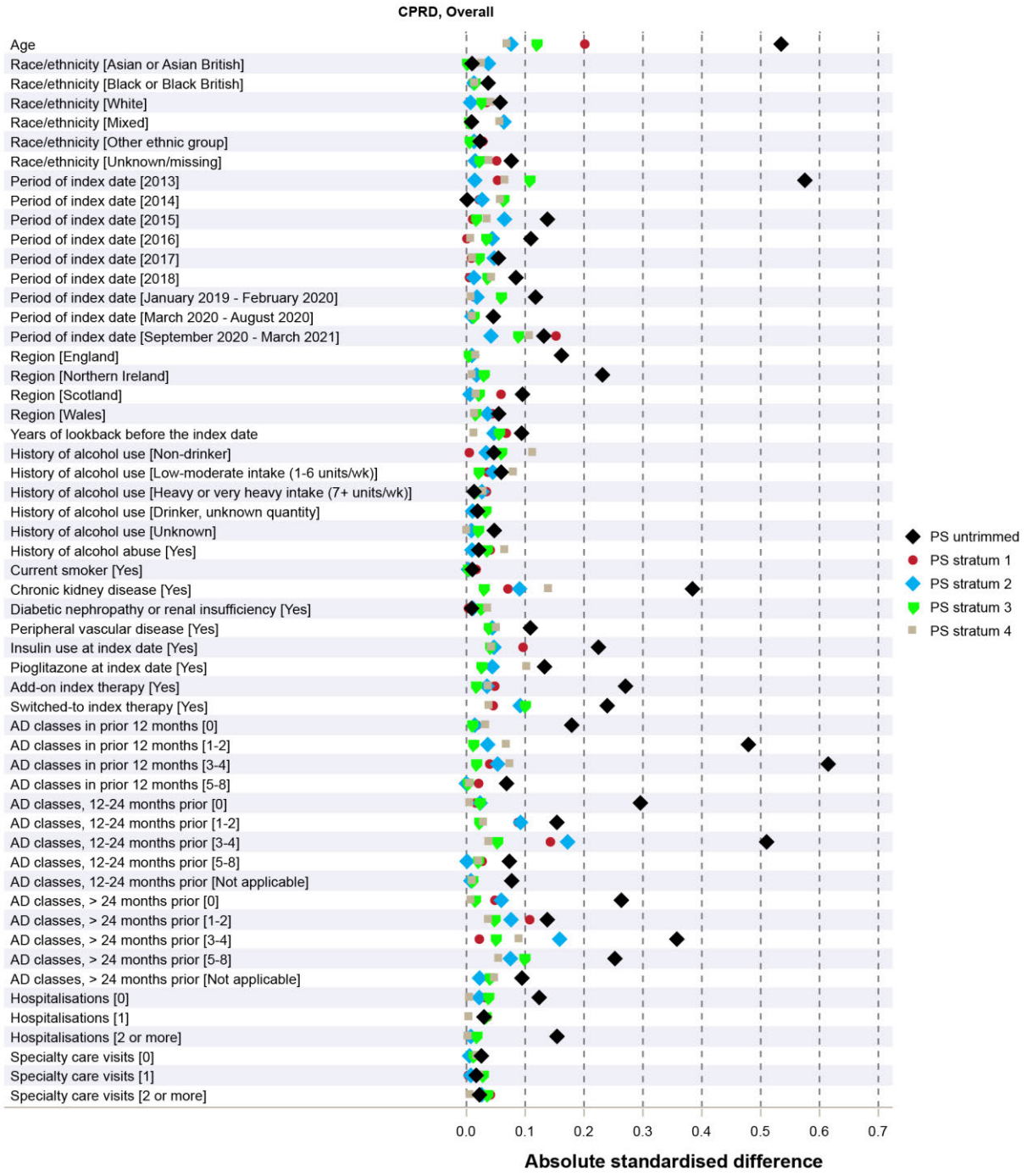
Female Cohort

For the female bladder cancer cohorts, baseline characteristics after propensity score trimming for the overall cohorts and the cohorts stratified by insulin use at the index date are presented for each data source in Appendix J, BladderCa Table 1.2F, BladderCa Table 2.2F, BladderCa Table 3.2F, and BladderCa Table 4.2F.

Figure 13 displays plots of the StDiff values for the covariates included in the propensity score models for the overall female bladder cancer cohort for each data source and illustrates the balance of key covariates between the index treatment groups before propensity score trimming and within propensity score strata after trimming. For cohorts where propensity score stratification was not feasible (ie, female bladder cancer cohort in PHARMO and the HIRD), the StDiff of covariates in the untrimmed sample and in the full propensity score–trimmed sample (without propensity score stratification) are displayed. In CPRD and Medicare, for the female bladder cancer cohorts, propensity score trimming and stratification were effective in achieving balance between the index exposure groups for most variables; StDiff values for most variables within most individual propensity score strata were ≤ 0.20. In Medicare, after propensity score trimming, the imbalance of some variables was higher (StDiff > 0.20) in some propensity score strata (eg, age, years of lookback, chronic kidney disease) than in the full (untrimmed) sample. Similar patterns observed in the balance of covariates in the overall female bladder cancer cohorts were also observed in the cohorts stratified by insulin use at the index date (Appendix J, BladderCa Figure 2F).

Figure 13 Balance of Covariates in the Female Bladder Cancer Cohort, Both in the Full Sample Before Trimming and Within Propensity Score Strata After Trimming, Overall Cohort, by Data Source

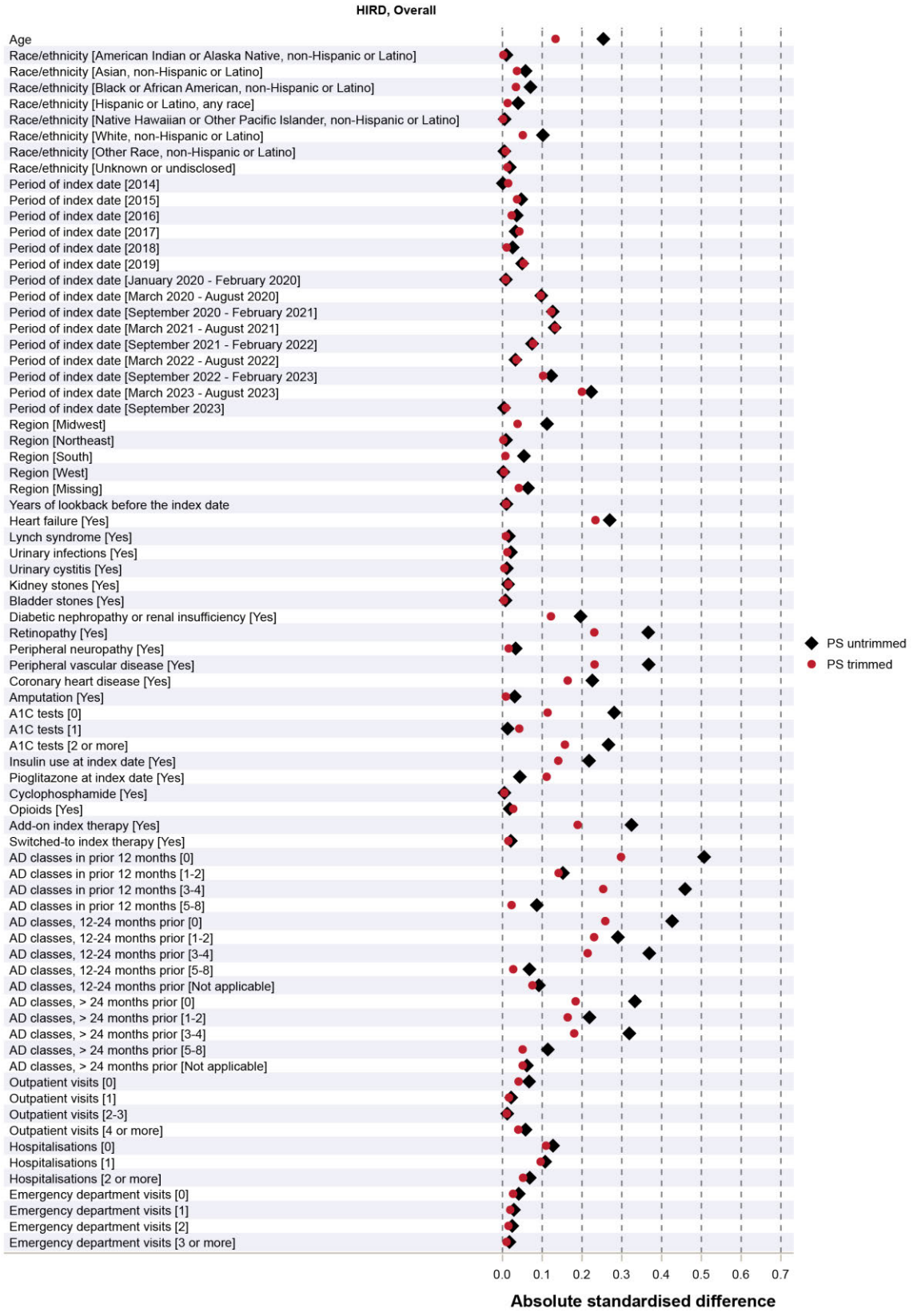
CPRD



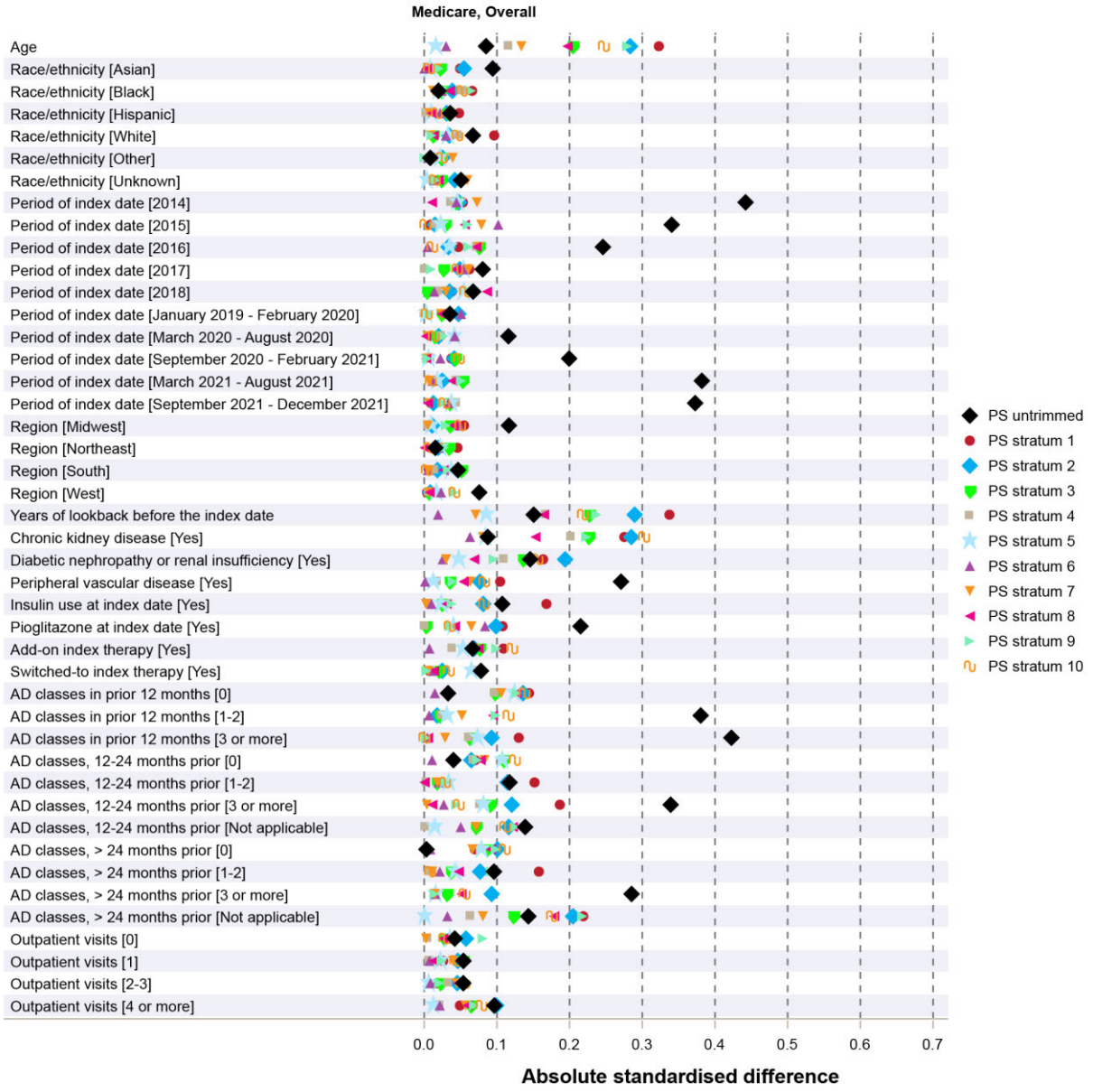
PHARMO


PHARMO, Overall
Age
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [January 2019 - February 2020]
Period of index date [March 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - December 2022]
Region [NO]
Region [NW]
Region [Midden]
Region [Zuid]
Years of lookback before the index date
SES [Low]
SES [Middle]
SES [High]
SES [Unknown]
History of alcohol use [Non-drinker]
History of alcohol use [Low-moderate intake (1-6 units/wk)]
History of alcohol use [Heavy or very heavy intake (7+ units/wk)]
History of alcohol use [Drinker, unknown quantity]
History of alcohol use [Unknown]
Current smoker [Current]
Current smoker [Former]
Current smoker [Never]
Current smoker [Unknown]
BMI [< 20 (underweight)]
BMI [20 to < 25 (normal)]
BMI [25 to < 30 (overweight)]
BMI [30 to < 40 (obese)]
BMI [≥ 40 (severely obese)]
BMI [Unknown]
Heart failure [yes]
Chronic kidney disease [yes]
Urinary infections [yes]
Urinary cystitis [yes]
Kidney stones [yes]
Time since T2DM diagnosis [0-3]
Time since T2DM diagnosis [>3-6]
Time since T2DM diagnosis [>6-8]
Time since T2DM diagnosis [>8-10]
Time since T2DM diagnosis [>10]
Time since T2DM diagnosis [Unknown]
Diabetic nephropathy or renal insufficiency [yes]
Retinopathy [yes]
Peripheral neuropathy [yes]
Peripheral vascular disease [yes]
Coronary heart disease [yes]
Cerebrovascular disease [yes]
Insulin use at index date [yes]
Pioglitazone at index date [yes]
Opioids [yes]
Add-on index therapy [yes]
Switched-to index therapy [yes]
AD classes in prior 12 months [0-2]
AD classes, 12-24 months prior [0-2]
AD classes, > 24 months prior [0-2]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4-5]
Outpatient visits [6+]
Outpatient visits [Unknown]
Hospitalisations [0]
Hospitalisations [1]
Hospitalisations [2-3]
Hospitalisations [4+]
Availability of GP data [yes]
Index combination [yes]
Monotherapy at index date [yes]
Non evaluable index therapy [yes]
PS untrimmed
PS trimmed
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

HIRD


HIRD, Overall
Age
Race/ethnicity [American Indian or Alaska Native, non-Hispanic or Latino]
Race/ethnicity [Asian, non-Hispanic or Latino]
Race/ethnicity [Black or African American, non-Hispanic or Latino]
Race/ethnicity [Hispanic or Latino, any race]
Race/ethnicity [Native Hawaiian or Other Pacific Islander, non-Hispanic or Latino]
Race/ethnicity [White, non-Hispanic or Latino]
Race/ethnicity [Other Race, non-Hispanic or Latino]
Race/ethnicity [Unknown or undisclosed]
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [2019]
Period of index date [January 2020 - February 2020]
Period of index date [March 2020 - August 2020]
Period of index date [September 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - February 2023]
Period of index date [March 2023 - August 2023]
Period of index date [September 2023]
Region [Midwest]
Region [Northeast]
Region [South]
Region [West]
Region [Missing]
Years of lookback before the index date
Heart failure [Yes]
Lynch syndrome [Yes]
Urinary infections [Yes]
Urinary cystitis [Yes]
Kidney stones [Yes]
Bladder stones [Yes]
Diabetic nephropathy or renal insufficiency [Yes]
Retinopathy [Yes]
Peripheral neuropathy [Yes]
Peripheral vascular disease [Yes]
Coronary heart disease [Yes]
Amputation [Yes]
A1C tests [0]
A1C tests [1]
A1C tests [2 or more]
Insulin use at index date [Yes]
Pioglitazone at index date [Yes]
Cyclophosphamide [Yes]
Opioids [Yes]
Add-on index therapy [Yes]
Switched-to index therapy [Yes]
AD classes in prior 12 months [0]
AD classes in prior 12 months [1-2]
AD classes in prior 12 months [3-4]
AD classes in prior 12 months [5-8]
AD classes, 12-24 months prior [0]
AD classes, 12-24 months prior [1-2]
AD classes, 12-24 months prior [3-4]
AD classes, 12-24 months prior [5-8]
AD classes, 12-24 months prior [Not applicable]
AD classes, > 24 months prior [0]
AD classes, > 24 months prior [1-2]
AD classes, > 24 months prior [3-4]
AD classes, > 24 months prior [5-8]
AD classes, > 24 months prior [Not applicable]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4 or more]
Hospitalisations [0]
Hospitalisations [1]
Hospitalisations [2 or more]
Emergency department visits [0]
Emergency department visits [1]
Emergency department visits [2]
Emergency department visits [3 or more]
PS untrimmed
PS trimmed
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

Medicare



PS = propensity score.

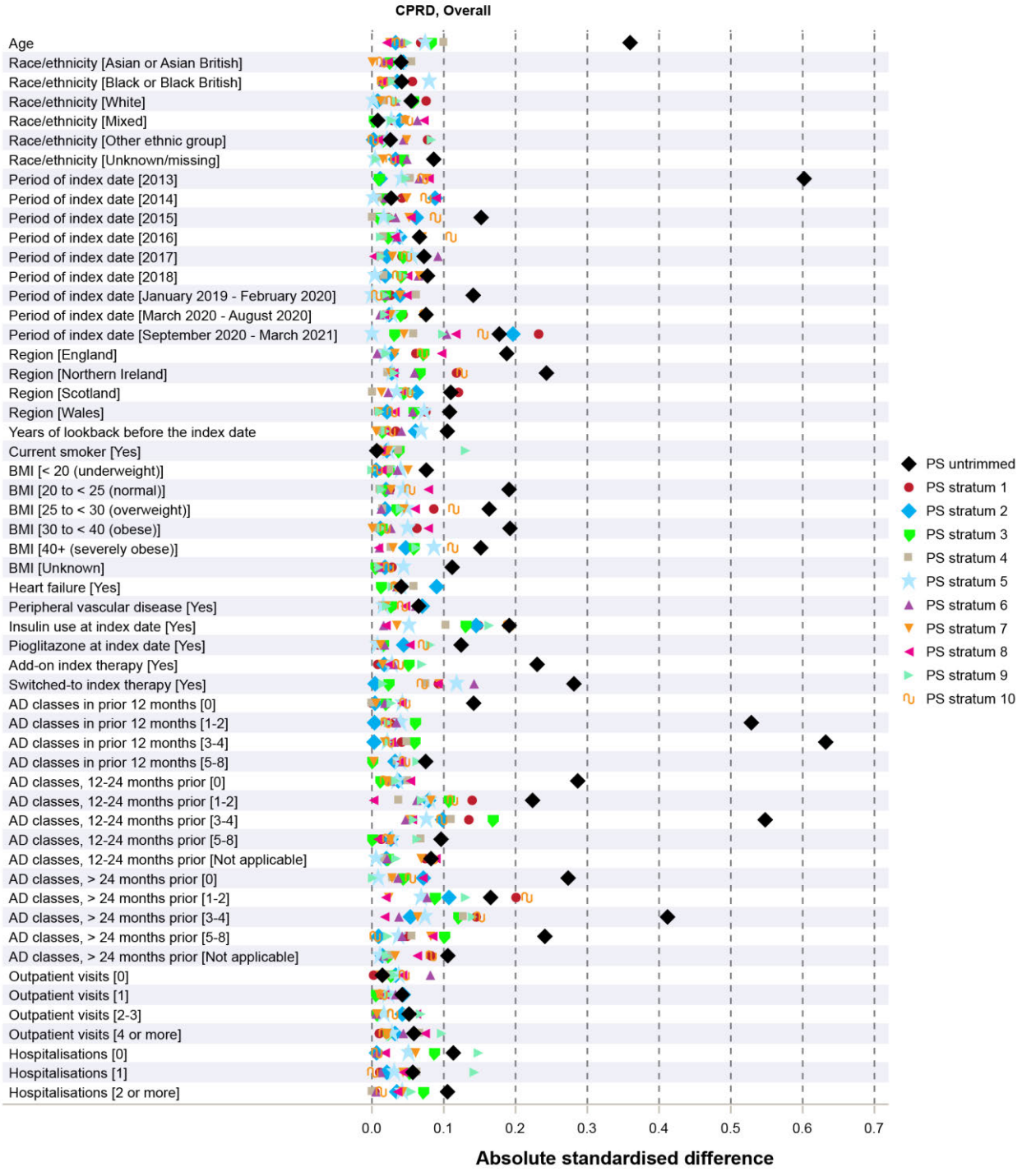
Male Cohort

For the male bladder cancer cohorts, baseline characteristics after propensity score trimming for the overall cohorts and the cohorts stratified by insulin use at the index date are presented for each data source in Appendix J, BladderCa Table 1.2M, BladderCa Table 2.2M, BladderCa Table 3.2M, and BladderCa Table 4.2M.

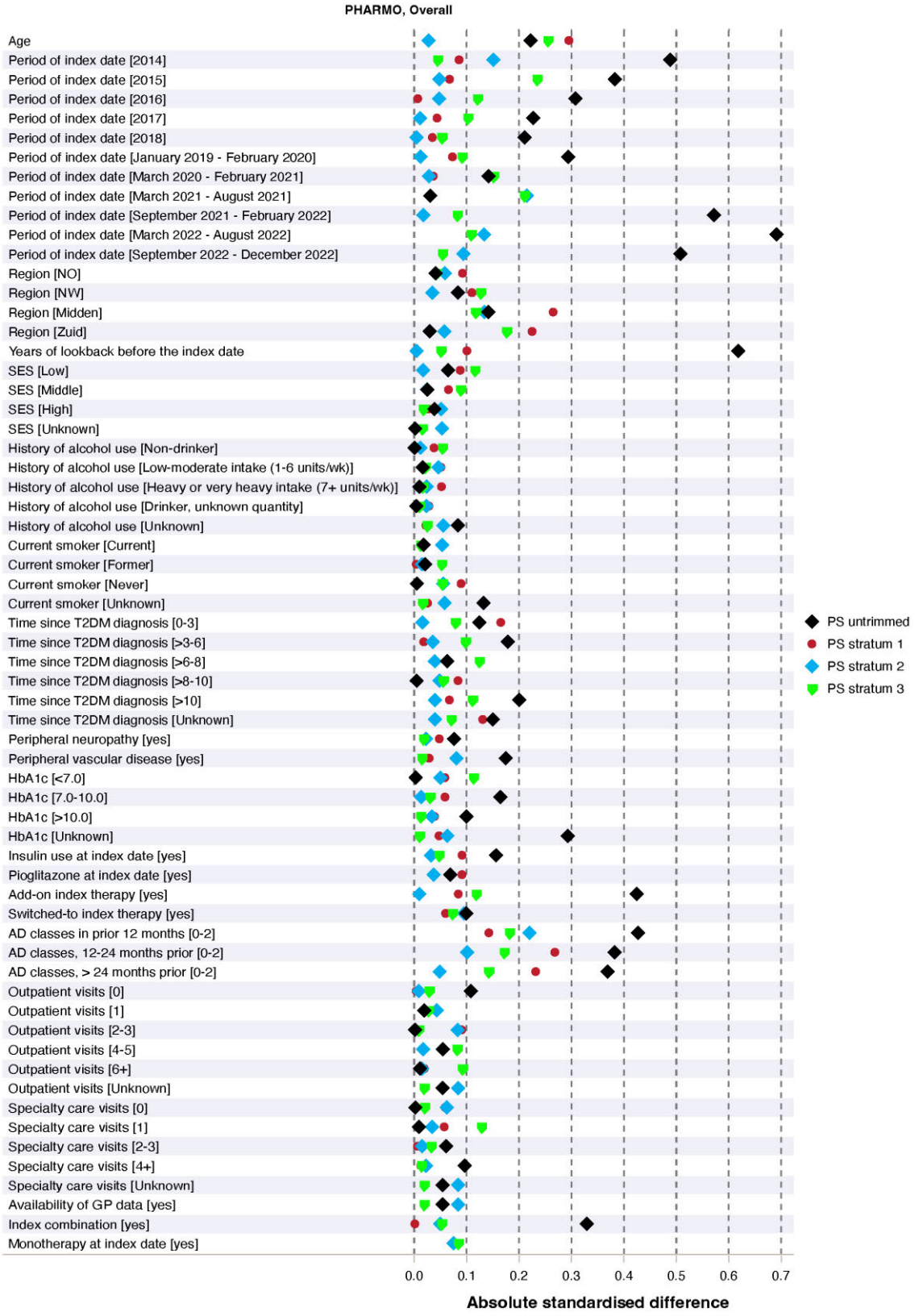
Figure 14 displays plots of the StDiff values for the overall male bladder cancer cohort for each data source to illustrate the balance of key covariates between the index treatment groups before propensity score trimming and within propensity score strata after trimming for all data sources. For the male bladder cancer cohorts, propensity score trimming and stratification were effective in achieving balance between the index exposure groups for most variables; StDiff values for most variables within most individual propensity score strata were ≤ 0.20 across all data sources. In some data sources, after propensity score trimming, the imbalance of some variables was higher (StDiff > 0.20) in some propensity score strata (eg, age in Medicare) than in the full (untrimmed) sample. Similar patterns observed in the balance of covariates in the overall male bladder cancer cohorts were also observed in the cohorts stratified by insulin use at the index date (Appendix J, BladderCa Figure 2M).

Figure 14 Balance of Covariates in the Male Bladder Cancer Cohort, Both in the Full Sample Before Trimming and Within Propensity Score Strata After Trimming, Overall Cohort, by Data Source

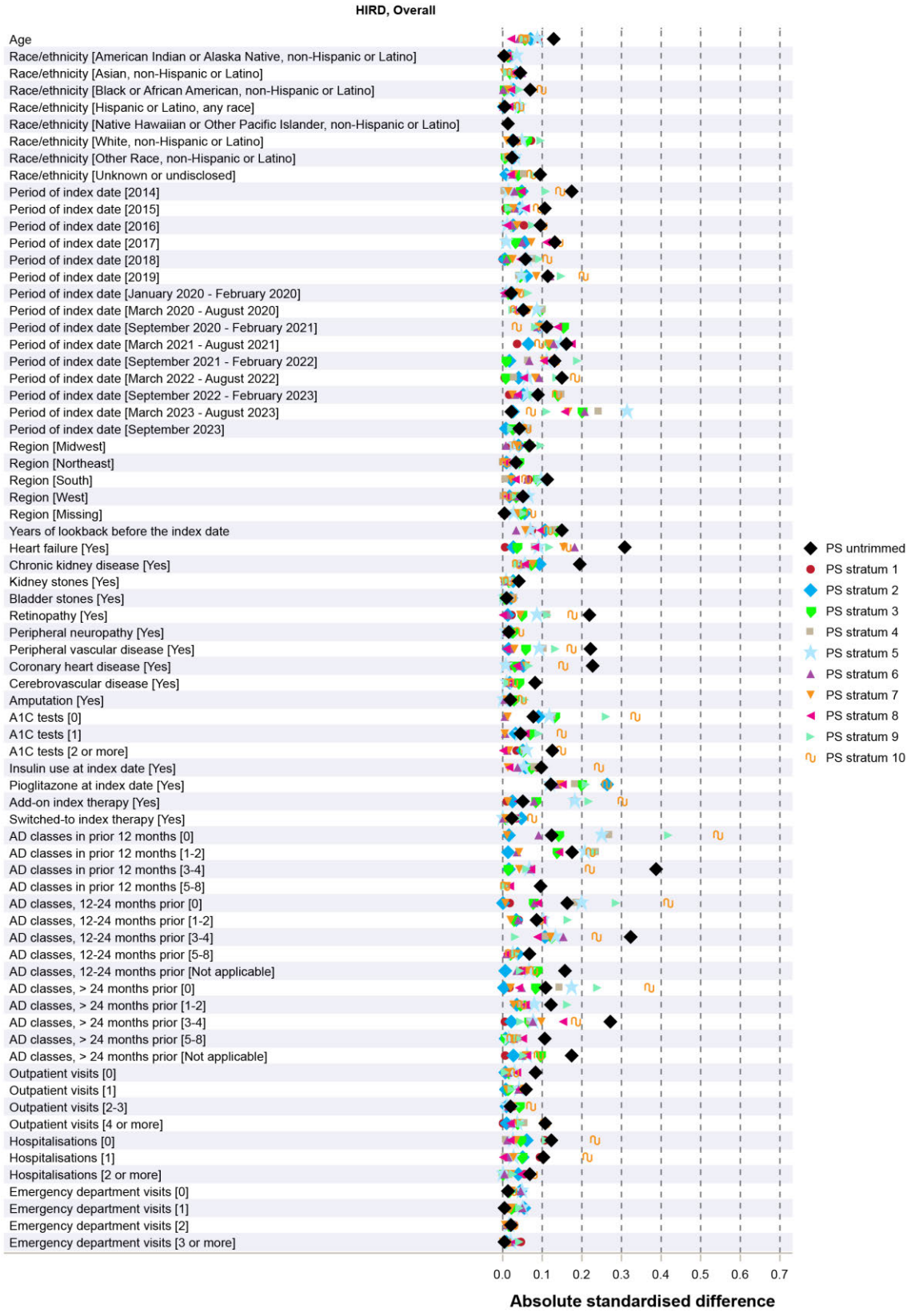
CPRD



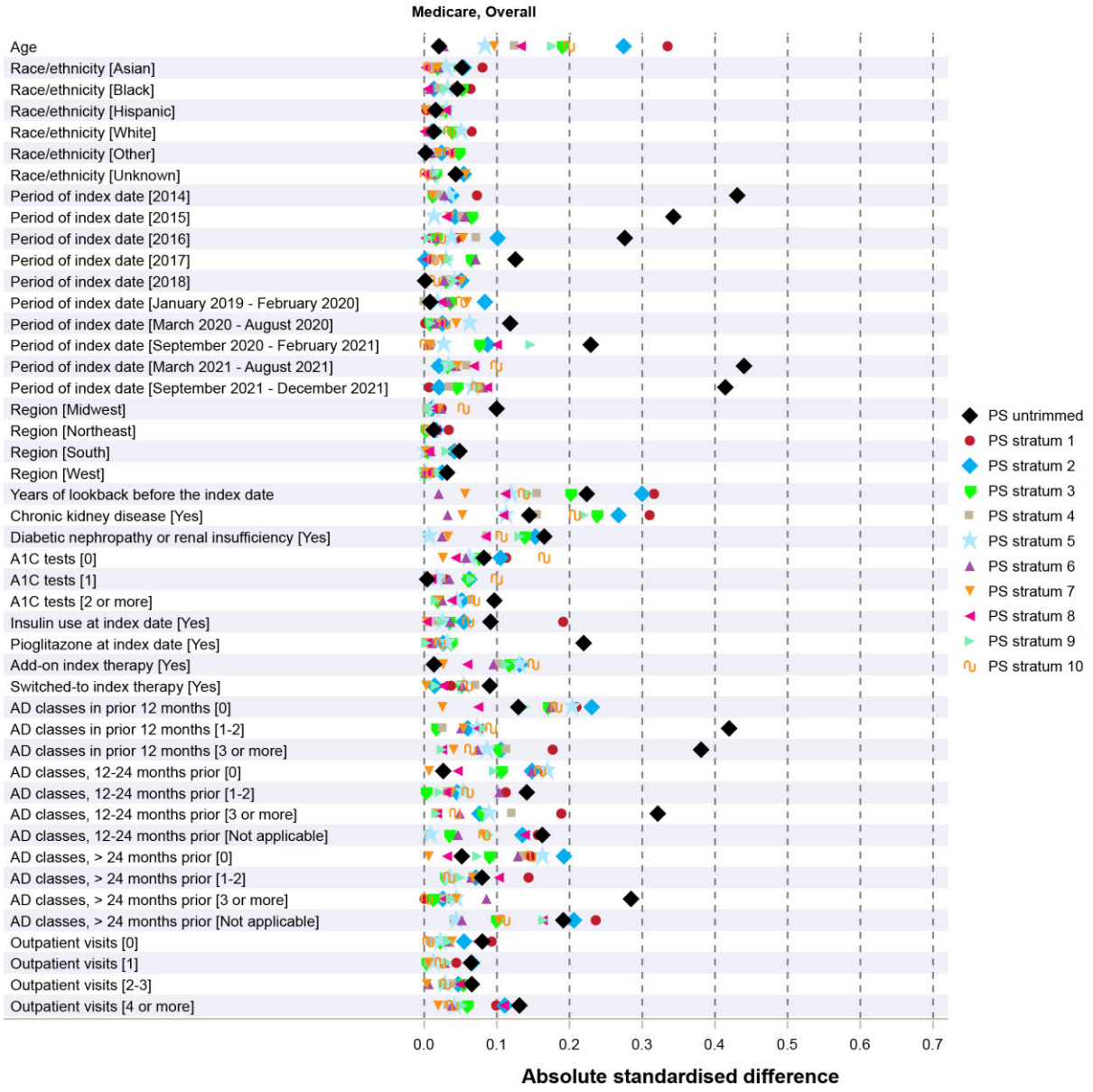
PHARMO


PHARMO, Overall
Age
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [January 2019 - February 2020]
Period of index date [March 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - December 2022]
Region [NO]
Region [NW]
Region [Midden]
Region [Zuid]
Years of lookback before the index date
SES [Low]
SES [Middle]
SES [High]
SES [Unknown]
History of alcohol use [Non-drinker]
History of alcohol use [Low-moderate intake (1-6 units/wk)]
History of alcohol use [Heavy or very heavy intake (7+ units/wk)]
History of alcohol use [Drinker, unknown quantity]
History of alcohol use [Unknown]
Current smoker [Current]
Current smoker [Former]
Current smoker [Never]
Current smoker [Unknown]
Time since T2DM diagnosis [0-3]
Time since T2DM diagnosis [>3-6]
Time since T2DM diagnosis [>6-8]
Time since T2DM diagnosis [>8-10]
Time since T2DM diagnosis [>10]
Time since T2DM diagnosis [Unknown]
Peripheral neuropathy [yes]
Peripheral vascular disease [yes]
HbA1c [<7.0]
HbA1c [7.0-10.0]
HbA1c [>10.0]
HbA1c [Unknown]
Insulin use at index date [yes]
Pioglitazone at index date [yes]
Add-on index therapy [yes]
Switched-to index therapy [yes]
AD classes in prior 12 months [0-2]
AD classes, 12-24 months prior [0-2]
AD classes, > 24 months prior [0-2]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4-5]
Outpatient visits [6+]
Outpatient visits [Unknown]
Specialty care visits [0]
Specialty care visits [1]
Specialty care visits [2-3]
Specialty care visits [4+]
Specialty care visits [Unknown]
Availability of GP data [yes]
Index combination [yes]
Monotherapy at index date [yes]
PS untrimmed
PS stratum 1
PS stratum 2
PS stratum 3
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

HIRD


HIRD, Overall
Age
Race/ethnicity [American Indian or Alaska Native, non-Hispanic or Latino]
Race/ethnicity [Asian, non-Hispanic or Latino]
Race/ethnicity [Black or African American, non-Hispanic or Latino]
Race/ethnicity [Hispanic or Latino, any race]
Race/ethnicity [Native Hawaiian or Other Pacific Islander, non-Hispanic or Latino]
Race/ethnicity [White, non-Hispanic or Latino]
Race/ethnicity [Other Race, non-Hispanic or Latino]
Race/ethnicity [Unknown or undisclosed]
Period of index date [2014]
Period of index date [2015]
Period of index date [2016]
Period of index date [2017]
Period of index date [2018]
Period of index date [2019]
Period of index date [January 2020 - February 2020]
Period of index date [March 2020 - August 2020]
Period of index date [September 2020 - February 2021]
Period of index date [March 2021 - August 2021]
Period of index date [September 2021 - February 2022]
Period of index date [March 2022 - August 2022]
Period of index date [September 2022 - February 2023]
Period of index date [March 2023 - August 2023]
Period of index date [September 2023]
Region [Midwest]
Region [Northeast]
Region [South]
Region [West]
Region [Missing]
Years of lookback before the index date
Heart failure [Yes]
Chronic kidney disease [Yes]
Kidney stones [Yes]
Bladder stones [Yes]
Retinopathy [Yes]
Peripheral neuropathy [Yes]
Peripheral vascular disease [Yes]
Coronary heart disease [Yes]
Cerebrovascular disease [Yes]
Amputation [Yes]
A1C tests [0]
A1C tests [1]
A1C tests [2 or more]
Insulin use at index date [Yes]
Pioglitazone at index date [Yes]
Add-on index therapy [Yes]
Switched-to index therapy [Yes]
AD classes in prior 12 months [0]
AD classes in prior 12 months [1-2]
AD classes in prior 12 months [3-4]
AD classes in prior 12 months [5-8]
AD classes, 12-24 months prior [0]
AD classes, 12-24 months prior [1-2]
AD classes, 12-24 months prior [3-4]
AD classes, 12-24 months prior [5-8]
AD classes, 12-24 months prior [Not applicable]
AD classes, > 24 months prior [0]
AD classes, > 24 months prior [1-2]
AD classes, > 24 months prior [3-4]
AD classes, > 24 months prior [5-8]
AD classes, > 24 months prior [Not applicable]
Outpatient visits [0]
Outpatient visits [1]
Outpatient visits [2-3]
Outpatient visits [4 or more]
Hospitalisations [0]
Hospitalisations [1]
Hospitalisations [2 or more]
Emergency department visits [0]
Emergency department visits [1]
Emergency department visits [2]
Emergency department visits [3 or more]
PS untrimmed
PS stratum 1
PS stratum 2
PS stratum 3
PS stratum 4
PS stratum 5
PS stratum 6
PS stratum 7
PS stratum 8
PS stratum 9
PS stratum 10
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Absolute standardised difference

Medicare



PS = propensity score.

Sex-Combined Cohort, Stratified by Pioglitazone Use

For the sex-combined bladder cancer cohort stratified by pioglitazone use at the index date, baseline characteristics after propensity score trimming are presented for CPRD, the HIRD, and Medicare in Appendix J, BladderCa Table 17.2 (baseline demographic characteristics), BladderCa Table 18.2 (specific baseline medical conditions), BladderCa Table 19.2 (specific baseline medications), and BladderCa Table 20.2 (HCRU during the 180 days before the index date); these tables were not generated for PHARMO, as described below.

For the sex-combined bladder cancer cohorts stratified by pioglitazone use in CPRD, the HIRD, and Medicare, propensity score trimming and stratification were effective in achieving balance between the index exposure groups for most but not all of the variables in the propensity score model; most StDiff values for the individual propensity score strata were ≤ 0.20 for all data sources (Appendix J, BladderCa Figure 7). In PHARMO, the propensity score modelling and trimming were not conducted in the cohorts stratified by pioglitazone use at the index date due to the small number of dapagliflozin new users and bladder cancer events (dapagliflozin, n = 0; comparator AD, n = 1) amongst pioglitazone users; therefore, the propensity score–trimmed analysis samples stratified by pioglitazone use at the index date were not generated for the PHARMO sex-combined bladder cancer cohort.

10.2.4.3 Female Composite Cancer

For the female composite cancer cohorts, plots of the StDiff values for the covariates included in the propensity score models for the overall cohort and the insulin use–stratified cohorts, illustrating the balance of key covariates between the index treatment groups (dapagliflozin and comparator AD) before propensity score trimming and within propensity score strata after trimming, are displayed for each data source in Appendix J, FemaleCompositeCa Figure 2.

Across data sources, propensity score trimming and stratification for the female composite cancer cohorts were effective in achieving balance between the index exposure groups for most variables in the propensity score models; StDiff values for most variables within most individual propensity score strata were ≤ 0.20 in all data sources, which correspond to small differences in the distribution of the variable between the dapagliflozin and comparator AD groups at the propensity score stratum level. The balance is similar to that observed for the female breast cancer cohorts (described in Section 10.2.4.1).

10.2.4.4 Male Composite Cancer

For the male composite cancer cohorts, plots of the StDiff values for the covariates included in the propensity score models for the overall cohort and the insulin use–stratified cohorts, illustrating the balance of key covariates between the index treatment groups (dapagliflozin and comparator AD) before propensity score trimming and within propensity score strata after trimming, are displayed for each data source in Appendix J, MaleCompositeCa Figure 2.

Across data sources, propensity score trimming and stratification for the male composite cancer cohorts were effective in achieving balance between the index exposure groups for most variables in the propensity score models; StDiff values for most variables within most individual propensity score strata were ≤ 0.20 in all data sources, which correspond to small differences in the distribution of the variable between the dapagliflozin and comparator AD

groups at the propensity score stratum level. The balance is similar to that observed for the male bladder cancer cohorts (described in Section 10.2.4.2).

10.2.5 Cumulative Dapagliflozin Dose

The estimated cumulative dose of dapagliflozin within the propensity score–trimmed populations for the overall cohorts and the insulin use–stratified cohorts is presented for each data source in Table 16. For the bladder cancer cohorts only, the estimated cumulative dose of dapagliflozin within the propensity score–trimmed population stratified by pioglitazone use at the index date is presented for each data source in Appendix J, BladderCa Table 22.

The cumulative dapagliflozin dose for the propensity score–trimmed cohorts for the female breast cancer cohorts and the sex-combined bladder cancer cohorts by data source are summarised below:

- The mean cumulative dose of dapagliflozin in the overall female breast cancer cohorts ranged from approximately 2,500 mg in Medicare to approximately 5,880 mg in CPRD. For the sex-combined bladder cancer cohorts, the mean cumulative dose of dapagliflozin ranged from approximately 2,710 mg in Medicare to approximately 6,180 mg in CPRD.
- Across the primary cancer outcomes, the mean cumulative dose of dapagliflozin in insulin users compared with insulin non-users was higher in CPRD, lower in PHARMO, and similar in the HIRD and Medicare.

Table 16 Dapagliflozin Cumulative Dose for the Female Breast Cancer and Sex-Combined Bladder Cancer Cohorts, Propensity Score–Trimmed Analysis Samples, Overall and Stratified by Insulin Use at the Index Date, by Data Source

Outcome cohort	Insulin use at the index date	CPRD	PHARMO	HIRD	Medicare
Female breast cancer					
New users of dapagliflozin, n	Insulin	1,082	172	3,857	6,188
	No insulin	4,534	1,116	22,030	25,405
	Overall	5,713	1,382	25,988	31,656
Cumulative dose (mg)					
Mean (SD)	Insulin	6,537.8 (6,351.8)	3,898.0 (4,267.9)	3,041.9 (3,873.75)	2,651.4 (3,419.03)
	No insulin	5,723.7 (5,800.2)	4,714.3 (5,990.8)	3,083.4 (4,015.72)	2,462.3 (3,351.93)
	Overall	5,878.6 (5,924.3)	4,500.2 (5,701.2)	3,067.4 (3,989.42)	2,499.0 (3,383.48)
Median (IQR)	Insulin	4,400.0 (8,490.0)	2,835.0 (4,450.0)	1,620.0 (3,320.0)	NR (2,965.0)
	No insulin	3,620.0 (7,630.0)	2,160.0 (5,030.0)	1,590.0 (3,360.0)	NR (2,520.0)
	Overall	3,730.0 (7,800.0)	2,185.0 (4,735.0)	1,580.0 (3,330.0)	NR (2,600.0)

Outcome cohort	Insulin use at the index date	CPRD	PHARMO	HIRD	Medicare
Minimum; maximum	Insulin	30.0; 26,935.0	30.0; 31,760.0	10.0; 27,210.0	NR; NR
	No insulin	5.0; 29,070.0	20.0; 32,420.0	5.0; 31,080.0	NR; NR
	Overall	5.0; 29,070.0	20.0; 32,420.0	5.0; 31,200.0	NR; NR
Sex-combined bladder cancer					
New users of dapagliflozin, n	Insulin	2,344	499	8,807	12,003
	No insulin	11,257	2,908	54,231	50,325
	Overall	14,050	3,298	63,525	62,856
Cumulative dose (mg)					
Mean (SD)	Insulin	6,814.1 (6,507.8)	4,145.8 (4,845.4)	3,385.6 (4,180.05)	2,784.9 (3,517.48)
	No insulin	6,087.7 (5,917.6)	4,700.2 (5,774.3)	3,354.7 (4,223.17)	2,707.2 (3,518.27)
	Overall	6,175.6 (6,017.8)	4,774.4 (5,738.2)	3,350.0 (4,219.23)	2,713.2 (3,518.45)
Median (IQR)	Insulin	4,540.0 (9,030.0)	2,670.0 (4,590.0)	1,805.0 (3,750.0)	NR (3,140.0)
	No insulin	4,090.0 (8,100.0)	2,410.0 (5,045.0)	1,790.0 (3,780.0)	NR (2,940.0)
	Overall	4,140.0 (8,230.0)	2,550.0 (5,340.0)	1,790.0 (3,760.0)	NR (2,970.0)
Minimum; maximum	Insulin	30.0; 28,530.0	20.0; 31,760.0	5.0; 31,550.0	NR; NR
	No insulin	5.0; 29,070.0	10.0; 32,560.0	2.5; 40,810.0	NR; NR
	Overall	5.0; 29,070.0	20.0; 32,560.0	2.5; 40,810.0	NR; NR

Note: The reporting of minima, maxima, medians, modes, and percentiles is not permitted under the Medicare data use agreement.

IQR = interquartile range; NR = not reportable; SD = standard deviation.

10.3 Outcome Data

Detailed outcome data are shown in Appendix J, BreastCa Table 5, BladderCa Table 5, FemaleCompositeCa Table 5, and MaleCompositeCa Table 5, for each data source and are summarised below.

10.3.1 Primary Cancer Outcomes

Preliminary counts of the female breast cancer outcome, the bladder cancer outcome (sex-combined and amongst males and females separately), in both exposure groups combined, before propensity score trimming, are shown for each data source in Table 17. In the overall cohorts and in both exposure groups combined across all data sources, a total of 10,025 provisional cases of female breast cancer were identified by the algorithm. Amongst insulin users at the index date, the total number of provisional breast cancer cases was 1,306 across both exposure groups in all data sources. In the overall cohorts in both exposure groups combined across all data sources, a total of 3,546 provisional bladder cancer cases were

identified by the algorithm. Amongst insulin users at the index date, the total number of provisional bladder cancer cases was 433 across both exposure groups in all data sources.

10.3.2 Secondary Cancer Outcomes

Preliminary counts of provisional cases of the selected cancers comprising the female composite cancers and the male composite cancers, in both exposure groups combined, before propensity score trimming, are shown for each data source in Table 17. The total number of provisional cases of selected cancers in females identified by the algorithm in the overall cohorts was 11,780. Amongst new users in both exposure groups combined that had insulin use at the index date, the total number of provisional cases of selected cancers in females was 1,533 across all data sources. The total number of provisional cases of selected cancers in males identified by the algorithm was 21,970. Amongst new users in both exposure groups combined that had insulin use at the index date, the total number of provisional cases of selected cancers in males was 2,738 across all data sources.

Table 17 Number of Provisional Cases of Cancer Outcomes Identified Using the Electronic Algorithm; All Outcomes, Overall and by Insulin Use at the Index Date, by Data Source; Full Samples

Outcome cohort	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]
Female breast cancer	Insulin	4,848	45	2,280	26	29,663	165	101,504	1,070
	No insulin	37,283	343	18,163	239	323,844	1,657	562,225	6,480
	Overall	42,131	388	20,443	265	353,507	1,822	663,729	7,550
Bladder cancer	Insulin	10,784	27	5,145	21	63,392	23	180,755	362
	No insulin	91,238	191	42,000	179	606,553	232	1,021,401	2,511
	Overall	102,022	218	47,145	200	669,945	255	1,202,156	2,873
Female	Insulin	4,823	8	2,275	6	29,047	1-4	98,501	76
	No insulin	37,098	40	18,141	24	318,022	47	545,611	534
	Overall	41,921	48	20,416	30	347,069	48-51	644,112	610
Male	Insulin	5,961	19	2,870	15	34,345	1-4	82,254	286
	No insulin	54,140	151	23,859	155	288,531	185	475,790	1,977
	Overall	60,101	170	26,729	170	322,876	186-189	558,044	2,263
Female composite cancer[b]	Insulin	4,848	77	2,281	59	29,790	238	101,833	1,159
	No insulin	37,295	578	18,181	448	325,327	1,823	563,871	7,398
	Overall	42,143	655	20,462	507	355,117	2,061	665,704	8,557
Colon/rectum	Insulin	NA	13	NA	20	NA	50	NA	299
	No insulin	NA	126	NA	140	NA	338	NA	1,879
	Overall	NA	139	NA	160	NA	388	NA	2,178
Lung	Insulin	NA	20	NA	15	NA	42	NA	306
	No insulin	NA	175	NA	119	NA	253	NA	1,949
	Overall	NA	195	NA	134	NA	295	NA	2,255

Outcome cohort	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]
Corpus uteri	Insulin	NA	16	NA	8	NA	88	NA	265
	No insulin	NA	125	NA	64	NA	637	NA	1,680
	Overall	NA	141	NA	72	NA	725	NA	1,945
Ovary	Insulin	NA	9	NA	4	NA	14	NA	29
	No insulin	NA	39	NA	19	NA	149	NA	164
	Overall	NA	48	NA	23	NA	163	NA	193
Stomach	Insulin	NA	6	NA	0	NA	1-4	NA	30
	No insulin	NA	13	NA	15	NA	39	NA	184
	Overall	NA	19	NA	15	NA	40-43	NA	214
Non-Hodgkin lymphoma	Insulin	NA	8	NA	5	NA	24	NA	137
	No insulin	NA	46	NA	59	NA	208	NA	1,026
	Overall	NA	54	NA	64	NA	232	NA	1,163
Melanoma of skin	Insulin	NA	5	NA	7	NA	18	NA	95
	No insulin	NA	54	NA	32	NA	208	NA	522
	Overall	NA	59	NA	39	NA	226	NA	617
Male composite cancer[b]	Insulin	6,016	135	2,889	110	35,103	375	86,057	2,118
	No insulin	54,584	1,287	23,969	955	293,942	2,769	497,794	14,221
	Overall	60,600	1,422	26,858	1,065	329,045	3,144	583,851	16,339
Prostate	Insulin	NA	43	NA	34	NA	153	NA	1,095
	No insulin	NA	482	NA	301	NA	1,273	NA	7,467
	Overall	NA	525	NA	335	NA	1,426	NA	8,562
Colon/rectum	Insulin	NA	39	NA	21	NA	83	NA	314
	No insulin	NA	297	NA	222	NA	524	NA	1,949
	Overall	NA	336	NA	243	NA	607	NA	2,263

Outcome cohort	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]	Patients in cohort (N)	Provisional cases (N)[a]
Lung	Insulin	NA	32	NA	29	NA	46	NA	304
	No insulin	NA	281	NA	234	NA	334	NA	2,219
	Overall	NA	313	NA	263	NA	380	NA	2,523
Stomach	Insulin	NA	6	NA	5	NA	15	NA	41
	No insulin	NA	34	NA	35	NA	92	NA	238
	Overall	NA	40	NA	40	NA	107	NA	279
Non-Hodgkin lymphoma	Insulin	NA	9	NA	14	NA	43	NA	194
	No insulin	NA	110	NA	102	NA	288	NA	1,226
	Overall	NA	119	NA	116	NA	331	NA	1,420
Melanoma of skin	Insulin	NA	8	NA	7	NA	46	NA	171
	No insulin	NA	87	NA	62	NA	311	NA	1,128
	Overall	NA	95	NA	69	NA	357	NA	1,299

[a] Outcome assessed by electronic algorithm, not validated.

[b] Sum of provisional cases within each individual cancer type included in the female and male composite cancer outcome cohorts does not equal the total number of provisional cases for the female or male composite cancer outcome cohorts as some patients have multiple cancer diagnoses recorded on the same day.

Note 1: For HIRD data, any cell with a value of 1-4, or any cell that allows a value of 1-4 to be derived from other reported cells or information, cannot be reported. To comply with these reporting requirements, values that could be used to derive small count sizes in other cells are presented as ranges in this table.

Note 2: For each cancer type included in the female and male composite cancer outcome cohorts, the number of patients to assess each specific cancer type was not enumerated (and indicated as "NA") because no analyses assessed these cancer types independently; therefore, cohorts specific to these cancer types (ie, applying specific eligibility criteria for each of these cancer types) were not generated.

NA = not applicable.

10.4 Main Results

10.4.1 Female Breast Cancer

10.4.1.1 Incidence Analysis

10.4.1.1.1 Unadjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, BreastCa Table 7, for the results overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) after propensity score trimming for each data source. Unadjusted incidence rates for female breast cancer for each propensity score stratum are presented for each data source in Appendix J, BreastCa Table 8. This analysis addresses primary objective 1 (see Section 7.2.1).

The estimated accumulated person-time, number of provisional female breast cancer cases, and unadjusted incidence rates within the female breast cancer cohorts—overall cohorts and propensity score–trimmed cohorts stratified by insulin use—are presented for each data source in Table 18. The female breast cancer cohorts after propensity score trimming included a total of 130,358 person-years of accumulated dapagliflozin exposure for the overall cohorts across the data sources (Table 18). The mean [SD] duration of follow-up for the overall cohorts was slightly longer amongst dapagliflozin than comparator AD new users in CPRD (2.9 [2.06] years and 2.6 [2.00] years, respectively) and the HIRD (1.9 [1.89] years and 1.6 [1.71] years, respectively) whilst the opposite was observed in PHARMO (2.5 [2.31] years and 3.3 [2.33] years, respectively) and Medicare (1.9 years [1.72] and 2.7 [2.00] years, respectively). The overall unadjusted incidence rate for female breast cancer was lower in new users of dapagliflozin than in new users of comparator AD in CPRD and PHARMO, higher in dapagliflozin than in comparator AD in HIRD, and similar in both exposure groups in Medicare. Overall, unadjusted incidence rates for breast cancer were generally similar in both exposure groups in CPRD, PHARMO, and the HIRD. The highest unadjusted incidence rates for breast cancer were observed in Medicare for both the dapagliflozin and comparator AD groups.

Amongst insulin users, the unadjusted incidence rate of female breast cancer was higher in dapagliflozin new users than in comparator AD new users in CPRD and the HIRD, lower in PHARMO, and similar in both groups in Medicare. In PHARMO, a single case of female breast cancer occurred amongst dapagliflozin-exposed insulin users. Amongst insulin non-users, differences in the unadjusted incidence rate for female breast cancer between dapagliflozin new users and comparator AD new users were similar to the trends observed for the unadjusted incidence rates in the overall cohorts across all data sources. See Table 18 for details.

Table 18 Unadjusted and Propensity Score–Adjusted Incidence Rates (per 10,000 Person-years) for Female Breast Cancer, Overall and Stratified by Insulin Use at the Index Date, by Data Source; Propensity Score–Trimmed Analysis Samples

	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of breast cancer events	Insulin	12	17	1	17	33	114	52	846
	No insulin	40	190	8	148	146	1,252	184	4,572
	Overall	51	210	10	162	173	1,377	236	5,422
Person-years	Insulin	3,660	6,707	480	4,392	7,240	38,193	12,097	196,376
	No insulin	12,706	55,258	2,866	42,400	42,472	390,336	48,034	1,139,756
	Overall	16,555	61,534	3,481	46,844	49,973	432,127	60,349	1,340,593
Unadjusted incidence rate (95% CI)	Insulin	32.78 (16.94-57.24)	25.35 (14.77-40.58)	20.83 (0.53-116.03)	38.70 (22.55-61.97)	45.58 (31.37-64.01)	29.85 (24.62-35.86)	42.99 (32.10-56.37)	43.08 (40.23-46.08)
	No insulin	31.48 (22.49-42.87)	34.38 (29.67-39.64)	27.91 (12.05-55.00)	34.91 (29.51-41.00)	34.38 (29.03-40.43)	32.07 (30.32-33.90)	38.31 (32.97-44.26)	40.11 (38.96-41.29)
	Overall	30.81 (22.94-40.51)	34.13 (29.67-39.07)	28.73 (13.78-52.83)	34.58 (29.46-40.34)	34.62 (29.65-40.18)	31.87 (30.20-33.59)	39.11 (34.28-44.43)	40.44 (39.38-41.54)
Adjusted incidence rate[a] (95% CI)	Insulin	32.78 (16.94-57.27)	23.75 (9.55-44.20)	NE	NE	45.58 (31.37-64.01)	32.40 (25.92-39.84)	42.99 (32.10-56.37)	40.62 (37.15-44.27)
	No insulin	31.48 (22.49-42.87)	30.89 (25.20-37.22)	27.91 (12.05-55.00)	38.50 (29.91-48.21)	34.38 (29.03-40.43)	33.23 (31.17-35.37)	38.31 (32.97-44.26)	40.24 (38.66-41.86)
	Overall	30.81 (22.94-40.51)	28.74 (23.32-34.74)	28.73 (13.78-52.83)	34.49 (27.48-42.36)	34.62 (29.65-40.18)	32.98 (30.99-35.06)	39.11 (34.28-44.43)	40.34 (38.88-41.82)

[a] Adjusted incidence rates for propensity score strata were standardised to the person-time of the dapagliflozin cohort; therefore, for the dapagliflozin group, the adjusted incidence rates are the same (or nearly the same) as the unadjusted incidence rates.

Note: The size of the overall sample may not equal the sum of the insulin use–stratified samples, as propensity score estimation and trimming were performed separately in each sample.

NE = not estimable; No. = number.

10.4.1.1.2 Propensity Score–Adjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, BreastCa Table 10, for each data source, for the results overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) after propensity score trimming. This analysis addresses primary objective 1 (see Section 7.2.1).

Propensity score–adjusted incidence rates were estimated for each exposure group using the propensity score–stratified rates standardised to the person-years in the dapagliflozin group; therefore, for the dapagliflozin group, the propensity score–adjusted incidence rates are the same (or nearly the same) as the unadjusted incidence rates. Propensity score–adjusted incidence rates for female breast cancer are presented for each data source in Table 18. In the overall cohorts not stratified by insulin use at the index date, the propensity score–adjusted incidence rate per 10,000 person-years point estimate for female breast cancer in CPRD and the HIRD, respectively, was higher in dapagliflozin new users (30.81 and 34.62) than in comparator AD new users (28.74 and 32.98); the propensity score–incidence rate in PHARMO was lower in dapagliflozin new users (28.73) than in comparator AD new users (34.49) and was similar in both exposure groups in Medicare (Table 18). The propensity score–adjusted incidence rates for breast cancer in both exposure groups were higher in females aged 65 years or older than in females aged younger than 65 years in CPRD, whilst in PHARMO, propensity score–adjusted incidence rates amongst dapagliflozin new users aged 65 years or older were lower than in females aged younger than 65 years; amongst comparator AD new users, propensity score–adjusted incidence rates were similar in females aged < 65 years and those aged ≥ 65 years (Appendix J, BreastCa Table 10, for each data source). In CPRD, propensity score–adjusted incidence rates in the dapagliflozin group were slightly higher than those in the comparator AD group in females younger than 65 years of age; in PHARMO, propensity score–adjusted incidence rates in the dapagliflozin group were lower than those in the comparator AD for females aged 65 years or older.

Amongst insulin users, the propensity score–adjusted incidence rates per 10,000 person-years for female breast cancer were higher in dapagliflozin new users than in comparator AD new users in CPRD and the HIRD, and comparable in dapagliflozin new users and comparator AD new users in Medicare. In PHARMO, the propensity score–adjusted incidence rate for insulin users was not estimable in either exposure group due the small number of cases (Table 18). In CPRD, the propensity score–adjusted incidence rates for breast cancer in both the dapagliflozin and comparator AD groups were higher in females aged 65 years or older than in females younger than 65 years of age. In both age groups in CPRD, propensity score–adjusted incidence rates of breast cancer were slightly higher in new users of dapagliflozin than in new users of the comparator AD medications (Appendix J, BreastCa Table 10, for each data source).

Amongst insulin non-users, the propensity score–adjusted incidence rates per 10,000 person-years for female breast cancer were similar in new users of dapagliflozin and comparator ADs in CPRD, the HIRD, and Medicare and lower in dapagliflozin new users than in comparator AD new users in PHARMO (Table 18). In age-stratified results in CPRD, the propensity score–adjusted incidence rate of breast cancer was generally comparable in new users of dapagliflozin and new users of comparator AD in patients 65 years of age or older and in patients younger than 65 years of age, although propensity score–adjusted incidence rates were slightly higher in both exposure groups in patients 65 years or older than the younger patients (Appendix J, BreastCa Table 10, for each data source). Conversely, in age-stratified analyses within PHARMO, propensity score–incidence rates were comparable in new users of dapagliflozin and new users of comparator ADs in patients younger than 65 years of age but lower in the dapagliflozin group than in the comparator AD group in patients 65 years or older.

10.4.1.2 Cumulative Treatment Duration

The descriptive results of the female breast cancer incidence rates (in addition to the number of patients, person-years, and breast cancer events) for each exposure group by categories of cumulative treatment duration (in years), after propensity score trimming for the overall and insulin use–stratified cohorts, are presented for each data source in Appendix J, BreastCa Table 9.

As described in Section 9.4.2.4, crude incidence rates for female breast cancer (in the overall cohort and the cohorts stratified by insulin use at the index date) were calculated for each of the mutually exclusive categories of cumulative treatment duration—less than 1 year, 1 to < 2 years, 2 to < 5 years, and ≥ 5 years. Descriptive results showing the crude incidence rates for female breast cancer by treatment duration categories and stratified by insulin use are summarised in Table 19. In the overall cohorts, for the treatment duration of less than 1 year, dapagliflozin new users had a lower crude incidence rate of female breast cancer than comparator AD new users in CPRD, PHARMO, and Medicare. For the treatment duration of 1 to < 2 years, the incidence rates were similar in dapagliflozin new users and comparator AD new users in CPRD, higher in dapagliflozin new users than comparator AD new users in PHARMO, and lower in dapagliflozin new users than comparator AD new users in Medicare. For the treatment duration of 2 to < 5 years, rates were similar in dapagliflozin new users and comparator AD new users in CPRD and Medicare and lower in dapagliflozin new users than comparator new users in PHARMO. In the HIRD, for the treatment duration of less than 1 year, 1 to < 2 years, and 2 to < 5 years, crude incidence rates for breast cancer were higher amongst dapagliflozin new users than comparator AD new users. In the treatment duration category of ≥ 5 years, crude incidence rates were higher amongst dapagliflozin new users than comparator AD new users in PHARMO and Medicare, whilst the opposite was observed in the HIRD; 95% CIs for these estimates were wide due to the low number of

cancer events. In CPRD, for the treatment duration of ≥ 5 years, there were no female breast cancer events in the dapagliflozin group (Table 19).

In PHARMO, for the insulin users cohort, there were no female breast cancer events amongst dapagliflozin new users in the treatment duration categories of less than 1 year, 2 to < 5 years, and ≥ 5 years; therefore, the crude incidence rates were zero. Similarly, for insulin users in the treatment duration category of ≥ 5 years, there were no female breast cancer events in either exposure group across all data sources, except for the comparator AD group in Medicare. In the insulin users cohort, for the treatment duration 1 to < 2 years, although based on limited accumulated person-years, crude incidence rates for female breast cancer were higher in dapagliflozin new users than in comparator AD new users across all data sources.

Table 19 Unadjusted Incidence Rates (per 10,000 Person-years) for Female Breast Cancer by Treatment Duration, Overall and Stratified by Insulin Use at the Index Date, by Data Source; Propensity Score–Trimmed Analysis Samples

Treatment duration	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
< 1 year, incidence rate (95% CI)	Insulin	10.20 (0.26-56.82)	36.89 (15.93-72.68)	0.00 (0.00-174.35)	55.31 (29.45-94.58)	81.13 (35.03-159.86)	70.29 (50.44-95.36)	54.67 (37.64-76.78)	61.32 (56.22-66.75)
	No insulin	68.03 (43.59-101.22)	58.02 (46.89-71.01)	43.15 (14.01-100.70)	58.94 (45.67-74.85)	75.44 (53.90-102.73)	60.15 (54.36-66.38)	51.28 (42.84-60.89)	60.45 (58.24-62.72)
	Overall	55.20 (35.72-81.49)	58.16 (47.61-70.34)	27.97 (7.62-71.62)	52.36 (41.32-65.45)	72.94 (53.40-97.30)	61.05 (55.52-66.98)	51.77 (44.13-60.35)	60.63 (58.59-62.71)
1 to < 2 years, incidence rate (95% CI)	Insulin	102.27 (37.53-222.59)	22.61 (4.66-66.08)	72.65 (1.84-404.76)	21.93 (0.56-122.16)	46.92 (15.24-109.50)	23.05 (11.91-40.26)	57.50 (31.44-96.48)	46.88 (40.23-54.32)
	No insulin	28.14 (10.33-61.25)	44.65 (32.57-59.75)	24.14 (0.61-134.49)	58.63 (41.07-81.17)	31.49 (18.96-49.18)	24.00 (19.89-28.73)	29.68 (19.39-43.49)	44.42 (41.57-47.41)
	Overall	43.60 (22.53-76.16)	43.88 (32.57-57.85)	69.09 (18.83-176.90)	54.50 (38.17-75.45)	32.28 (20.46-48.44)	23.69 (19.80-28.11)	34.98 (24.88-47.82)	44.53 (41.91-47.27)
2 to < 5 years, incidence rate (95% CI)	Insulin	39.37 (12.78-91.89)	23.59 (8.66-51.35)	0.00 (0.00-335.92)	26.71 (5.51-78.06)	27.50 (7.49-70.41)	9.43 (3.46-20.53)	16.18 (5.25-37.75)	21.49 (17.85-25.66)
	No insulin	21.04 (10.09-38.70)	23.59 (17.51-31.11)	12.44 (0.31-69.30)	24.05 (16.93-33.15)	22.82 (13.94-35.24)	14.30 (11.40-17.70)	21.56 (13.96-31.83)	21.80 (20.30-23.38)
	Overall	23.16 (12.66-38.86)	22.57 (16.91-29.52)	10.36 (0.26-57.74)	24.81 (17.72-33.78)	21.56 (13.51-32.64)	13.68 (10.99-16.83)	20.92 (14.21-29.69)	21.76 (20.36-23.22)
≥ 5 years, incidence rate (95% CI)	Insulin	0.00 (0.00-44.82)	0.00 (0.00-55.24)	0.00 (0.00-1,744.77)	0.00 (0.00-79.73)	0.00 (0.00-75.55)	0.00 (0.00-30.75)	0.00 (0.00-68.85)	9.38 (5.13-15.74)
	No insulin	0.00 (0.00-16.08)	1.28 (0.03-7.15)	20.44 (0.52-113.90)	8.42 (3.63-16.58)	2.87 (0.07-15.97)	6.45 (2.95-12.25)	12.88 (2.66-37.63)	6.09 (4.73-7.72)
	Overall	0.00 (0.00-11.42)	1.19 (0.03-6.61)	19.72 (0.50-109.87)	9.56 (4.37-18.16)	2.50 (0.06-13.91)	5.99 (2.74-11.38)	10.37 (2.14-30.30)	6.54 (5.21-8.09)

Note: The size of the overall sample may not equal the sum of the insulin use–stratified samples because propensity score estimation and trimming were performed separately in each sample. NR = not reportable.

10.4.1.3 Comparative Analysis

10.4.1.3.1 Unadjusted Comparisons, Overall and by Insulin Use at the Index Date

See Appendix J, BreastCa Table 7, for results overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) after propensity score trimming for each data source. Unadjusted IRRs for female breast cancer for each propensity score stratum are presented for each data source in Appendix J, BreastCa Table 8. This analysis addresses primary objective 1 (see Section 7.2.1).

For female breast cancer, based on the comparison of incidence rates in the dapagliflozin group with incidence rates in the comparator AD group, the overall unadjusted IRR was 0.90 (95% CI, 0.65-1.23) in CPRD, 0.83 (95% CI, 0.39-1.57) in PHARMO, 1.09 (95% CI, 0.92-1.27) in the HIRD, and 0.97 (95% CI, 0.85-1.10) in Medicare.

In the insulin users cohort, the unadjusted IRR was 1.29 (95% CI, 0.56-2.87) in CPRD, 0.54 (95% CI, 0.01-3.43) in PHARMO, 1.53 (95% CI, 1.00-2.27) in the HIRD, and 1.00 (95% CI, 0.74-1.32) in Medicare. Across all data sources, the CIs for estimates in the insulin users cohort were wide due to the small number of cases, especially in the dapagliflozin group.

The same pattern observed in the IRRs for the overall cohorts was observed in insulin non-users, as they comprised a large proportion of the overall cohort.

10.4.1.3.2 Propensity Score–Adjusted Comparisons, Overall and by Insulin Use at the Index Date

See Appendix J, BreastCa Table 10, for the propensity score–adjusted IRR results for female breast cancer overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) for each data source. Results from the pooled analysis are shown in Appendix J, BreastCa Table 26. These analyses address primary objective 1 (see Section 7.2.1).

The propensity score–adjusted IRR estimates for female breast cancer for the overall and insulin use–stratified cohorts in each data source and the pooled adjusted IRRs across all data sources (with an estimated data source–specific propensity score–adjusted IRR) are presented in Figure 15. A pooled adjusted IRR for female breast cancer was reported for the overall cohort and for each insulin use–stratified cohort given that, for each cohort, there was minimal statistical heterogeneity across the data sources (criteria described in Section 9.9.7).

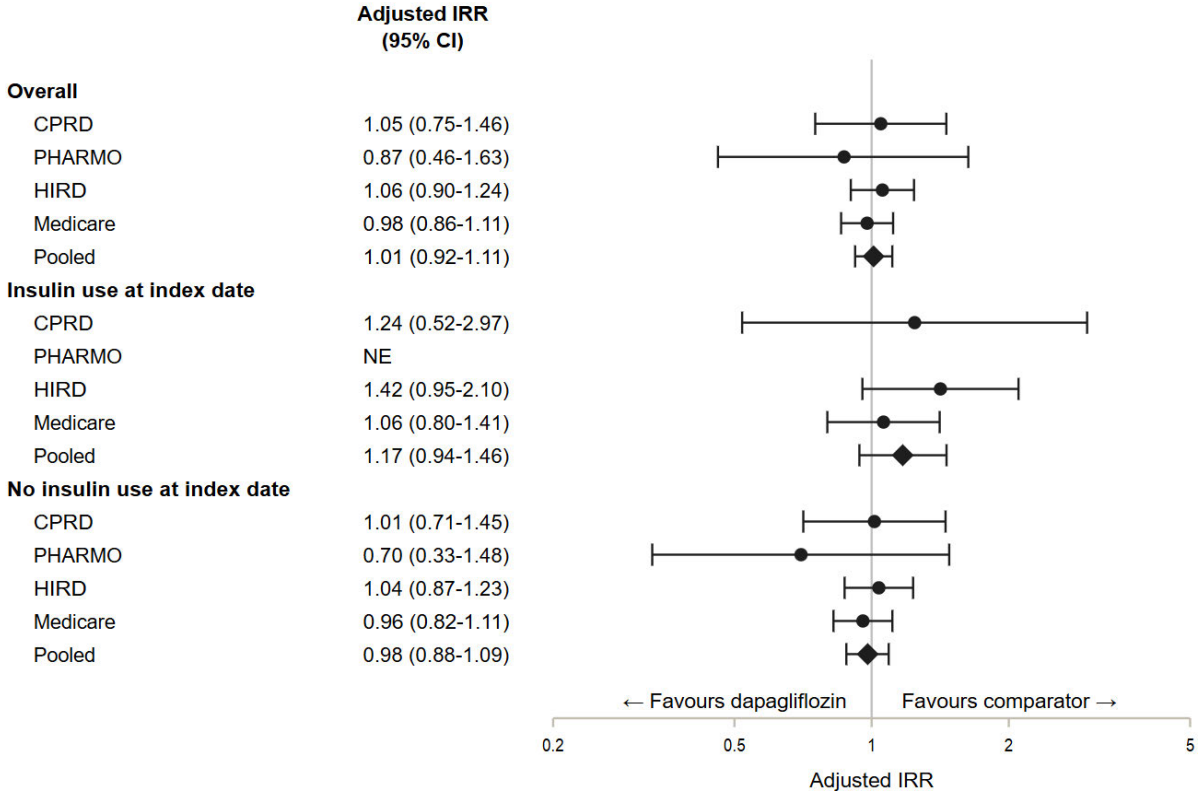
For female breast cancer, the overall propensity score–adjusted IRR estimates comparing dapagliflozin with comparator AD new users ranged from 0.87 (95% CI, 0.46-1.63) in PHARMO to 1.06 (95% CI, 0.90-1.24) in HIRD, with a pooled adjusted IRR estimate of 1.01

(95% CI, 0.92-1.11). The 95% CI was widest for the PHARMO estimate, indicating less precision due to the smaller number of new users and breast cancer cases.

Amongst insulin users, the propensity score–adjusted IRR estimates ranged from 1.06 (95% CI, 0.80-1.41) in Medicare to 1.42 (95% CI, 0.95-2.10) in the HIRD, with wide 95% CIs due to the low number of breast cancer cases, particularly in the dapagliflozin group. In PHARMO, the propensity score–adjusted IRR amongst insulin users was not estimable due to the low number of breast cancer cases amongst insulin users. The pooled adjusted IRR estimate for insulin users (1.17; 95% CI, 0.94-1.46) was based on data from CPRD, the HIRD, and Medicare.

For all data sources, estimates in the insulin non-users cohorts were similar to those observed in the overall cohorts, as insulin non-users comprised the largest portion of the overall cohort in each data source.

Figure 15 Propensity Score–Adjusted Incidence Rate Ratios for Female Breast Cancer, Overall and Stratified by Insulin Use at the Index Date, by Data Source



Note: The pooled IRR estimates were calculated using Mantel-Haenszel methods, stratified by propensity score stratum, as described in Section 9.9.7.

NE = not estimable.

10.4.1.4 HCRU During Follow-up: Secondary Outcome

Assessment of HCRU during follow-up was relevant for assessing the potential role of HCRU as a mediator of the total effect of dapagliflozin on breast cancer incidence if an effect was observed (See Section 9.9.3.7 for the criterion for conducting the mediation analysis).

10.4.1.4.1 Descriptive Assessment

See Appendix J, BreastCa Table 15, for yearly HCRU descriptive results for each exposure group, overall and stratified by insulin use at the index date, in the propensity score–trimmed cohorts for each data source. This analysis addresses secondary objective 1 (see Section 7.2.2).

No consistent differences in HCRU between dapagliflozin new users and comparator AD new users were observed across years of follow-up in any data source. The first year of follow-up had the highest average number of outpatient visits/encounters compared with subsequent follow-up years for both dapagliflozin new users and comparator AD new users, in CPRD (approximately 10 for both exposure groups), the HIRD (24-25, respectively), and Medicare (12 for both exposure groups), whilst in PHARMO, the average number of visits (only amongst patients with GP data) was consistent across first and subsequent follow-up years (19-20 for both exposure groups). The higher number of outpatient visits/encounters in the HIRD than in other data sources was driven by the broader definition including all outpatient encounters, eg, laboratory tests. For both dapagliflozin and comparator AD, the first year of follow-up had the highest proportions of new users with at least one gynaecologist visit, the magnitude varied by data source, with the smallest proportions in CPRD and PHARMO (~ 2%-3%) and higher proportions in the HIRD (27%-28%) and Medicare (10%-12%). Similarly, the proportion of dapagliflozin and comparator AD new users with a mammogram was higher in the first year of follow-up (CPRD, 10%-12%; the HIRD, ~52% for both groups; Medicare, 43%-46%) than in subsequent follow-up years. Very few patients (≤ 1%) had a breast biopsy during follow-up. Information on mammograms and breast biopsies was not available for patients in PHARMO.

10.4.1.4.2 Mediation Analysis to Assess HCRU as a Potential Mediator

Because the criterion for conducting the mediation analysis was not met (ie, the lower bounds of the 95% CI of the propensity score–adjusted IRRs for female breast cancer were not above 1.5 in any data source [Section 10.4.1.3]), the assessment of HCRU as a potential mediator of the association between exposure to a treatment group and cancer was not conducted in the female breast cancer cohort.

10.4.2 Bladder Cancer

10.4.2.1 Incidence Analysis

10.4.2.1.1 Unadjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, BladderCa Table 7, for the sex-combined bladder cancer cohorts after propensity score trimming, for the results overall, stratified by insulin use at the index date, and stratified by sex (for all data sources) and age (in CPRD and PHARMO only). Unadjusted incidence rates for bladder cancer for each propensity score stratum are presented in Appendix J, BladderCa Table 8, for each data source. This analysis addresses primary objective 2 (see Section 7.2.1).

For the sex-combined bladder cancer cohorts, after propensity score trimming, the estimated accumulated person-time, number of provisional bladder cancer cases, and unadjusted incidence rates within the overall and insulin use–stratified bladder cancer cohorts are presented for each data source in Table 20. The sex-combined bladder cancer cohorts after propensity score trimming included a total of 286,950 person-years of accumulated dapagliflozin exposure for the overall cohorts across the data sources (Table 20). The mean [SD] duration of follow-up was slightly longer amongst dapagliflozin than comparator AD new users in CPRD (2.9 [2.05] years and 2.6 [1.96] years, respectively) and the HIRD (1.9 [1.84] years and 1.6 [1.72] years, respectively), whilst the opposite was observed in PHARMO (2.5 years [2.30] and 3.4 [2.32] years, respectively) and Medicare (1.9 [1.69] years and 2.6 [1.98] years, respectively). The number of bladder cancer events amongst dapagliflozin new users was low in PHARMO and the HIRD and very low in the insulin-use cohorts across all data sources. The overall unadjusted incidence rates for sex-combined bladder cancer were similar in new users of dapagliflozin and comparator AD in PHARMO and the HIRD. In CPRD and Medicare, unadjusted incidence rates for bladder cancer were lower in dapagliflozin new users than comparator AD new users. Unadjusted incidence rates for both exposure groups were generally similar in CPRD and Medicare, with the highest unadjusted incidence rates for both exposure groups observed in PHARMO and the lowest unadjusted incidence rates for both exposure groups observed in the HIRD.

The pattern observed for unadjusted incidence rates in the overall cohorts was also observed in both the insulin users and insulin non-users cohorts.

Table 20 Unadjusted and Propensity Score–Adjusted Incidence Rates (per 10,000 Person-years) for Sex-Combined Bladder Cancer, Overall and Stratified by Insulin Use at the Index Date, by Data Source; Propensity Score–Trimmed Analysis Samples

	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of bladder cancer events	Insulin	6	14	2	8	1-4	20	1-10	289
	No insulin	11	93	9	138	25	194	68	1,751
	Overall	93	106	11	146	27	212	79	2,042
Person-years	Insulin	7,604	14,995	1,240	7,935	16,487	81,581	NR	339,229
	No insulin	31,449	129,615	6,940	94,551	102,133	779,981	95,527	1,960,653
	Overall	40,125	144,720	8,284	102,669	119,300	861,444	119,241	2,319,664
Unadjusted incidence rate (95% CI)	Insulin	7.89 (2.90-17.17)	9.34 (5.10-15.66)	16.13 (1.95-58.26)	10.08 (4.35-19.87)	1.82 (0.38-5.32)	2.45 (1.50-3.79)	3.06 (1.23-6.31)	8.52 (7.57-9.56)
	No insulin	3.50 (1.75-6.26)	7.18 (5.79-8.79)	12.97 (5.93-24.62)	14.60 (12.26-17.24)	2.45 (1.58-3.61)	2.49 (2.15-2.86)	7.12 (5.53-9.02)	8.93 (8.52-9.36)
	Overall	4.74 (2.85-7.39)	7.32 (6.00-8.86)	13.28 (6.63-23.76)	14.22 (12.01-16.72)	2.26 (1.49-3.29)	2.46 (2.14-2.82)	6.63 (5.25-8.26)	8.80 (8.43-9.19)
Adjusted incidence rate[a] (95% CI)	Insulin	7.89 (2.90-17.17)	6.64 (2.48-12.85)	NE	NE	1.82 (0.38-5.32)	2.26 (1.26-3.66)	3.06 (1.23-6.31)	8.35 (7.18-9.62)
	No insulin	3.50 (1.75-6.26)	5.45 (3.99-7.15)	12.97 (5.93-24.62)	14.51 (10.97-18.52)	2.45 (1.58-3.61)	2.84 (2.40-3.34)	7.12 (5.53-9.02)	8.79 (8.24-9.37)
	Overall	4.74 (2.85-7.39)	6.79 (4.91-8.96)	13.28 (6.63-23.76)	14.94 (11.57-18.74)	2.26 (1.49-3.29)	2.73 (2.33-3.19)	6.63 (5.25-8.26)	8.77 (8.26-9.30)

[a] Adjusted incidence rates for propensity score strata were standardised to the person-time of the dapagliflozin cohort; therefore, for the dapagliflozin group, the adjusted incidence rates are the same (or nearly the same) as the unadjusted incidence rates.

Note 1: Person-years and bladder cancer events in the overall sample may not equal the sum of person-years and bladder cancer events across samples stratified by insulin use because propensity score estimation and trimming were performed separately in the overall cohort and the cohorts stratified by insulin use.

Note 2: HIRD privacy rules do not allow displaying cells with values of 1-4 or cells that allow a value of 1-4 to be derived from other reported cells. Medicare privacy rules do not allow displaying cells with frequency values of 1-10 or any cell that allows a frequency of 1-10 to be derived from other reported cells.

NE = not estimable; No. = number; NR = not reportable.

10.4.2.1.2 PROPENSITY SCORE–ADJUSTED INCIDENCE RATES, OVERALL AND BY INSULIN USE AT THE INDEX DATE

Propensity score–adjusted incidence rates for sex-combined and sex-specific bladder cancer, for the overall and insulin use–stratified cohorts after propensity score trimming and also stratified by age (in CPRD and PHARMO only), are presented for each data source in Appendix J, BladderCa Table 10. This analysis addresses primary objective 1 (see Section 7.2.1).

Females and Males Combined

In addition to Appendix J, BladderCa Table 10, propensity score–adjusted incidence rates for sex-combined bladder cancer are presented for the overall and insulin use–stratified cohorts for each data source in Table 20. Adjustment of incidence rates was accomplished when propensity score–stratified rates in each exposure group were standardised to the person-years in the dapagliflozin group; for the dapagliflozin group, this resulted in propensity score–adjusted incidence rates that were identical or very similar to the unadjusted incidence rates. In the overall cohorts, the estimated propensity score–adjusted incidence rate for sex-combined bladder cancer per 10,000 person-years was consistently lower in dapagliflozin new users (range, 2.26-13.28) than in comparator AD new users (range, 2.73-14.94), across all data sources. In CPRD and PHARMO, the data sources where propensity score–adjusted incidence rates for bladder cancer were estimable by age group (< 65 years and ≥ 65 years), the propensity score–adjusted incidence rates for both exposure groups were higher in those aged at least 65 years than in those aged younger than 65 years. Amongst new users aged younger than 65 years, the propensity score–adjusted incidence rates were similar in both exposure groups in CPRD whilst in PHARMO the propensity score–adjusted incidence rates were slightly higher in dapagliflozin new users than in comparator AD new users. For new users aged 65 years or older, the adjusted incidence of bladder cancer per 10,000 person-years in CPRD was lower amongst dapagliflozin new users (8.03) than amongst comparator AD new users (15.20), whilst the propensity score–adjusted incidence rates in PHARMO were similar in both exposure groups (20.75 and 23.74, respectively) (Appendix J, BladderCa Table 10 CPRD; Appendix J, BladderCa Table 10 PHARMO).

The pattern observed for the propensity score–adjusted incidence rates for sex-combined bladder cancer in the overall cohorts was also observed in the insulin non-users cohorts across all data sources. Amongst insulin users, the propensity score–adjusted incidence rate for sex-combined bladder cancer was lower in dapagliflozin new users than in comparator AD new users in the HIRD and Medicare, higher in dapagliflozin new users than comparator AD new users in CPRD and was not estimable in dapagliflozin new users in PHARMO due to the small number of cases in both exposure groups.

Females

In the overall cohorts (not stratified by insulin use at the index date), the propensity score–adjusted incidence rate for bladder cancer per 10,000 person-years in females was consistently lower in dapagliflozin new users than in comparator AD new users in CPRD and Medicare. In the HIRD and PHARMO, the adjusted female bladder cancer incidence rates were not estimable in the overall or insulin use–stratified cohorts due to the low number of cases (Appendix J, BladderCa Table 10, for each data source).

The propensity score–adjusted bladder cancer incidence rates per 10,000 person-years amongst females for dapagliflozin new users and comparator AD new users, respectively, were 2.46 (95% CI, 0.67-6.30) and 3.10 (95% CI, 1.43-5.40) in CPRD, and 2.22 (95% CI, 1.18-3.79) and 3.09 (95% CI, 2.71-3.49) in Medicare (Appendix J, BladderCa Table 10, for each data source).

Amongst female insulin users, the propensity score–adjusted bladder cancer incidence rates per 10,000 person-years for dapagliflozin new users and comparator new users, respectively, were 8.49 (95% CI, 1.75-24.80) and 2.04 (95% CI, 0.42-5.97) in CPRD and were not estimable in Medicare (in both exposures groups) due to the low numbers of bladder cancer cases.

Amongst insulin non-users, the propensity score–adjusted bladder cancer incidence rates per 10,000 person-years amongst females were slightly lower in dapagliflozin new users (2.79; 95% CI, 1.49-4.77) than in comparator AD new users (3.05; 95% CI, 2.66-3.47) in Medicare. In CPRD, the propensity score–adjusted incidence rates in both exposure groups were not estimable due to the low number of female bladder cancer cases.

Males

The overall propensity score–adjusted incidence rates for bladder cancer per 10,000 person-years amongst males was lower in dapagliflozin new users than in comparator AD new users in CPRD and Medicare, and similar in both exposure groups in PHARMO and the HIRD. The adjusted rate per 10,000 amongst males for new users of dapagliflozin and new users of comparator AD, respectively, was 7.02 (95% CI, 4.09-11.24) and 8.24 (95% CI, 5.81-11.07) in CPRD, 20.6 (95% CI, 9.88-37.88) and 20.15 (95% CI, 14.63-26.55) in PHARMO, 3.53 (95% CI, 2.29-5.21) and 3.51 (95% CI, 2.91-4.19) in the HIRD, and 11.08 (95% CI, 8.59-14.07) and 14.30 (95% CI, 13.34-15.29) in Medicare (Appendix J, BladderCa Table 10, for each data source).

In CPRD and Medicare, the patterns observed in the propensity score–adjusted incidence rates in the overall cohorts were also observed in the insulin use–stratified cohorts. Amongst insulin users, the propensity score–adjusted incidence rates were not estimable in PHARMO (for dapagliflozin new users) and the HIRD (in both exposure groups) due to the low number of bladder cancer cases.

10.4.2.1.3 Unadjusted and Propensity Score–Adjusted Incidence Rates, by Pioglitazone Use at the Index Date

See Appendix J, BladderCa Table 23, for the sex-combined bladder cancer cohort incidence rate results after propensity score trimming for the analysis stratified by pioglitazone use and also by sex for CPRD, the HIRD, and Medicare (propensity score modelling was unable to be conducted in the pioglitazone use group in PHARMO due to the low number of dapagliflozin new users and bladder cancer events amongst the pioglitazone users; therefore, the propensity score–trimmed samples in the pioglitazone use–stratified cohorts were not generated for this data source, and this analysis was not conducted). This incidence analysis addresses primary objective 2 (see Section 7.2.2).

The estimated accumulated person-time, number of provisional bladder cancer cases, and unadjusted and propensity score–adjusted incidence rates per 10,000 person-years stratified by pioglitazone use at the index date within the propensity score–trimmed bladder cancer cohorts are displayed for each data source in Table 21 in addition to Appendix J, BladderCa Table 23. Overall, the number of pioglitazone users amongst dapagliflozin new users was small and represented a small proportion of all dapagliflozin new users: CPRD, n = 629 (4.4%); the HIRD, n = 2,412 (3.8%); and Medicare, n = 3,726 (5.9%) (Appendix J, BladderCa Table 23). In PHARMO, the pioglitazone use–stratified analyses were not conducted due to the low number of pioglitazone users and bladder cancer events.

Unadjusted Incidence Rates. Amongst pioglitazone users at the index date, the unadjusted incidence rates of bladder cancer per 10,000 person-years for the sex-combined cohorts were higher (CPRD) or slightly higher (the HIRD) amongst dapagliflozin new users than amongst comparator AD new users and slightly lower in dapagliflozin new users than in comparator AD new users in Medicare. However, the incidence rate estimates were imprecise due to the very small number of bladder cancer cases observed, particularly amongst dapagliflozin new users in the three data sources. Amongst pioglitazone non-users at the index date, the unadjusted incidence rates for sex-combined bladder cancer per 10,000 person-years were lower amongst dapagliflozin new users than amongst comparator AD new users in CPRD, the HIRD, and Medicare.

Propensity Score–Adjusted Incidence Rates. In the pioglitazone users cohort, the overall propensity score–adjusted incidence rate for sex-combined bladder cancer was estimable only amongst comparator AD new users in CPRD (11.01; 95 CI, 2.26-24.81) and Medicare (11.79; 95% CI, 9.64-14.15) and was not estimable in either exposure group in the HIRD due to the low number of bladder cancer cases.

In the pioglitazone non-users cohort, the propensity score–adjusted incidence rates for sex-combined bladder cancer were similar amongst dapagliflozin new users and comparator AD new users in the HIRD and were estimable only amongst the comparator AD group in CPRD

and Medicare. Propensity score–adjusted incidence rates of bladder cancer in the comparator AD group were higher for males than for females (Appendix J, BladderCa Table 23, for CPRD, the HIRD, and Medicare).

Table 21 Unadjusted and Propensity Score–Adjusted Incidence Rates (per 10,000 Person-years) for Sex-Combined Bladder Cancer, Stratified by Pioglitazone Use at the Index Date, by Data Source; Propensity Score–Trimmed Analysis Samples

	Pioglitazone use at the index date	CPRD		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
No. of bladder cancer events	Pioglitazone	1-4	11	1-4	24	1-10	280
	No pioglitazone	17	108	25	187	67	1,787
Person-years	Pioglitazone	NR	12,255	5,624	56,464	NR	262,592
	No pioglitazone	38,711	135,966	112,973	796,293	111,114	2,074,341
Unadjusted incidence rate (95% CI)	Pioglitazone	14.62 (3.02-42.73)	8.98 (4.48-16.06)	5.33 (1.10-15.59)	4.25 (2.72-6.32)	9.50 (4.10-18.73)	10.66 (9.45-11.99)
	No pioglitazone	4.39 (2.56-7.03)	7.94 (6.52-9.59)	2.21 (1.43-3.27)	2.35 (2.02-2.71)	6.03 (4.67-7.66)	8.61 (8.22-9.02)
Adjusted incidence rate[a] (95% CI)	Pioglitazone	NE	11.01 (2.26-24.81)	NE	NE	NE	11.79 (9.64-14.15)
	No pioglitazone	NE	5.55 (4.11-7.22)	2.21 (1.43-3.27)	2.61 (2.20-3.07)	NE	8.42 (7.90-8.95)

[a] Adjusted incidence rates for propensity score strata were standardised to the person-time of the dapagliflozin cohort.

Note 1: Person-years and bladder cancer events in the overall cohort may not equal the sum of person-years and bladder cancer events across cohorts stratified by pioglitazone use because propensity score estimation and trimming were performed separately in the overall cohort and the cohorts stratified by pioglitazone use.

Note 2: For CPRD data, any cell with a count of 1-4, or any cell that allows a count of 1-4 to be derived from other reported cells or information, cannot be reported. HIRD privacy rules do not allow displaying cells with values of 1-4 or cells that allow a value of 1-4 to be derived from other reported cells. Medicare privacy rules do not allow displaying cells with frequency values of 1-10 or any cell that allows a frequency of 1-10 to be derived from other reported cells. To comply with these reporting requirements, values that could be used to derive small count sizes in other cells are reported as ranges or are masked in this table.

NE = not estimable; No. = number; NR = not reportable.

10.4.2.2 Cumulative Treatment Duration

As described in Section 9.4.2.4, crude incidence rates for sex-combined bladder cancer (in the overall cohort, the cohorts stratified by insulin use at the index date, and the cohorts stratified by pioglitazone use at the index date) were calculated for each of the mutually exclusive categories of cumulative treatment duration—less than 1 year, 1 to < 2 years, 2 to < 5 years, and ≥ 5 years.

10.4.2.2.1 Overall and by Insulin Use at the Index Date

The descriptive results on the sex-combined bladder cancer incidence rates (in addition to the number of patients, person-years, and bladder cancer events) for each exposure group by category of cumulative treatment duration (in years), after propensity score trimming for the overall and insulin use–stratified cohorts, are presented for each data source in Appendix J, BladderCa Table 9.

Descriptive results showing the crude incidence rates for sex-combined bladder cancer by treatment duration categories, in the overall cohort and the cohorts stratified by insulin use at the index date are summarised in Table 22. In the overall cohort, for treatment duration categories of less than 1 year, 1 to < 2 years, and 2 to < 5 years, dapagliflozin new users had a lower crude incidence rate for bladder cancers than comparator AD new users in CPRD and Medicare. In PHARMO, for the less than 1 year treatment duration category, the crude incidence rate for bladder cancer was similar in both exposure groups whilst for the treatment duration categories 1 to < 2 years and 2 to < 5 years, dapagliflozin new users had lower crude incidence rates of bladder cancer than comparator AD new users. In the HIRD, for all treatment duration categories, crude incidence rates were similar between both exposure groups. Across all data sources, for the treatment duration category ≥ 5 years, there were few or no cases of bladder cancers in the dapagliflozin group.

Amongst insulin users, the limited number of bladder cancer cases across treatment duration categories in all data sources resulted in wide CIs for the crude incidence rates. Amongst insulin non-users, the trends of crude incidence rates were similar to those observed in the overall sex-combined bladder cancer cohort across treatment duration categories in all data sources.

Table 22 Unadjusted Incidence Rates (per 10,000 Person-years) for Sex-Combined Bladder Cancer by Treatment Duration, Overall and Stratified by Insulin Use at the Index Date, by Data Source; Propensity Score–Trimmed Analysis Samples

Treatment duration	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
< 1 year, incidence rate (95% CI)	Insulin	17.07 (3.52-49.89)	16.99 (7.33-33.47)	20.09 (0.51-111.96)	11.98 (3.89-27.96)	0.00 (0.00-16.81)	3.27 (0.89-8.36)	5.52 (2.03-12.01)	10.87 (9.24-12.69)
	No insulin	4.04 (0.83-11.82)	12.52 (9.06-16.86)	24.06 (8.83-52.36)	25.77 (19.52-33.39)	3.90 (1.27-9.10)	3.63 (2.66-4.84)	7.76 (5.43-10.74)	12.59 (11.80-13.41)
	Overall	6.32 (2.32-13.75)	12.29 (9.07-16.30)	23.12 (9.30-47.64)	23.83 (18.39-30.37)	3.30 (1.07-7.70)	3.66 (2.72-4.81)	7.60 (5.52-10.21)	12.25 (11.54-12.99)
1 to < 2 years, incidence rate (95% CI)	Insulin	17.66 (2.14-63.78)	9.29 (1.92-27.15)	32.76 (0.83-182.55)	11.61 (0.29-64.68)	0.00 (0.00-14.27)	1.80 (0.22-6.50)	2.14 (0.05-11.94)	9.77 (7.54-12.46)
	No insulin	3.89 (0.47-14.04)	10.16 (6.58-15.00)	8.26 (0.21-46.01)	20.97 (14.04-30.11)	3.17 (1.03-7.40)	2.73 (1.81-3.95)	10.00 (6.02-15.62)	10.14 (9.13-11.24)
	Overall	9.27 (3.40-20.17)	9.37 (6.12-13.73)	13.24 (1.60-47.82)	18.41 (12.13-26.79)	2.71 (0.88-6.32)	2.64 (1.78-3.77)	8.47 (5.17-13.08)	10.05 (9.13-11.05)
2 to < 5 years, incidence rate (95% CI)	Insulin	3.61 (0.09-20.10)	3.68 (0.45-13.30)	0.00 (0.00-117.68)	10.93 (1.32-39.48)	2.59 (0.07-14.41)	0.69 (0.02-3.84)	0.00 (0.00-5.88)	5.85 (4.45-7.54)
	No insulin	3.22 (0.88-8.23)	4.63 (2.97-6.89)	9.49 (1.15-34.27)	11.72 (8.44-15.84)	1.25 (0.26-3.64)	1.63 (1.02-2.46)	5.20 (2.77-8.89)	6.03 (5.45-6.66)
	Overall	3.21 (1.04-7.49)	5.24 (3.53-7.47)	8.06 (0.98-29.12)	12.02 (8.77-16.08)	1.07 (0.22-3.14)	1.54 (0.98-2.31)	4.75 (2.66-7.84)	5.99 (5.45-6.56)
≥ 5 years, incidence rate (95% CI)	Insulin	0.00 (0.00-18.99)	6.16 (0.16-34.31)	0.00 (0.00-297.99)	0.00 (0.00-34.45)	0.00 (0.00-26.38)	0.00 (0.00-13.24)	0.00 (0.00-35.24)	2.35 (0.86-5.12)
	No insulin	3.10 (0.38-11.21)	0.53 (0.01-2.96)	0.00 (0.00-32.75)	4.39 (2.11-8.08)	1.04 (0.03-5.78)	0.60 (0.07-2.18)	0.00 (0.00-7.19)	1.96 (1.39-2.69)
	Overall	2.33 (0.28-8.43)	0.97 (0.12-3.50)	0.00 (0.00-29.17)	3.86 (1.77-7.34)	0.89 (0.02-4.97)	0.56 (0.07-2.01)	0.00 (0.00-5.96)	1.87 (1.35-2.53)

Note: The size of the overall sample may not equal the sum of the insulin use–stratified samples because propensity score estimation and trimming were performed separately in each sample. NR = not reportable.

10.4.2.2.2 By Pioglitazone Use at the Index Date

The descriptive results on the sex-combined bladder cancer incidence rates (in addition to the number of patients, person-years, and bladder cancer events) for each exposure group by category of cumulative treatment duration (in years), after propensity score trimming for the pioglitazone use–stratified cohorts, are presented for CPRD, the HIRD, and Medicare in Appendix J, BladderCa Table 21. In PHARMO, propensity score modelling was unable to be conducted in the pioglitazone use group due to the low number of dapagliflozin new users and bladder cancer events amongst the pioglitazone users; therefore, the propensity score–trimmed samples in the pioglitazone use–stratified cohorts were not generated for this data source and this analysis was not conducted.

In the pioglitazone user cohorts in CPRD, the HIRD, and Medicare, there were few or no bladder cancer events (ie, not reportable due to small cell sizes) in the dapagliflozin group in each treatment duration category, resulting in crude incidence rates of zero or imprecise incidence rates with wide 95% CIs across all treatment duration categories.

In the pioglitazone non-user cohorts, across all treatment duration categories, dapagliflozin new users had a lower crude incidence rate for bladder cancer than comparator AD new users in CPRD (except for the ≥ 5 years treatment duration category) and Medicare. In the HIRD, for all treatment duration categories, crude incidence rates for bladder cancer were similar between both exposure groups. Across all data sources, for the treatment duration category ≥ 5 years, there were a limited number of bladder cancer cases in the dapagliflozin group.

10.4.2.3 Comparative Analysis

10.4.2.3.1 Unadjusted Comparisons, Overall and by Insulin Use at the Index Date

For each data source, the unadjusted IRR results for sex-combined bladder cancer, including number of person-years and number of cancer events, are presented by exposure group in Appendix J, BladderCa Table 7; results are presented for the overall and insulin use–stratified cohorts using the propensity score–trimmed samples for each data source and also stratified by age (CPRD and PHARMO only). Unadjusted IRRs for sex-combined bladder cancer for each propensity score stratum are presented in Appendix J, BladderCa Table 8, for each data source. The unadjusted comparative analysis addresses primary objective 1 (see Section 7.2.1).

Based on the comparison of incidence rates in the dapagliflozin group with incidence rates in the comparator AD group, the overall unadjusted IRR for sex-combined bladder cancer was 0.65 (95% CI, 0.37-1.06) in CPRD, 0.92 (95% CI, 0.59-1.38) in HIRD, 0.75 (95% CI, 0.59-0.94) in Medicare, and 0.93 (95% CI, 0.46-1.72) in PHARMO.

In the insulin users cohort, the unadjusted IRR was 0.85 (95% CI, 0.27-2.34) in CPRD, 1.60 (95% CI, 0.17-8.02) in PHARMO, 0.74 (95% CI, 0.14-2.50) in the HIRD, and 0.36 (0.14-0.75) in Medicare. In CPRD, PHARMO, and the HIRD, the CIs for estimates in the insulin users cohort were wide due to the small number of cases, especially in the dapagliflozin group.

The same pattern observed in the IRRs for the overall cohorts was observed in insulin non-users, as they comprised a large proportion of the overall cohort.

The unadjusted IRR was slightly higher in patients aged younger than 65 years than amongst those aged ≥ 65 years or older, respectively, in CPRD (0.99; 95% CI, 0.43-2.09 and 0.62; 95% CI, 0.27-1.24) and PHARMO (1.15; 95% CI, 0.30-3.18 and 0.95; 95% CI, 0.37-2.02).

10.4.2.3.2 Propensity Score–Adjusted Comparisons, Overall and by Insulin Use at the Index Date

See Appendix J, BladderCa Table 10, for the propensity score–adjusted IRR results for sex-combined bladder cancer for the overall and insulin use–stratified cohorts for each data source. Results are also stratified by sex for each data source and also stratified by age in CPRD and PHARMO only. Results from the pooled analysis are shown in Appendix J, BladderCa Table 26, BladderCa Table 26F, and BladderCa Table 26M. These analyses address primary objective 1 (see Section 7.2.1).

The propensity score–adjusted IRR estimates for bladder cancer for the overall and insulin use–stratified cohorts in each data source and the pooled adjusted IRRs across all data sources (with an estimated data source–specific propensity score–adjusted IRR) are presented in Figure 16 (sex-combined bladder cancer), Figure 17 (female bladder cancer), and Figure 18 (male bladder cancer). A pooled adjusted IRR for bladder cancer is reported for the overall cohort and for each insulin use–stratified cohort given that, for each cohort, minimal statistical heterogeneity was observed across the data sources (criteria described in Section 9.9.7). For female bladder cancer, pooled adjusted IRRs were not estimated for the insulin use–stratified cohorts given that data source–specific, propensity score–adjusted IRRs were estimable for only one data source each for insulin users (CPRD) and for insulin non-users (Medicare).

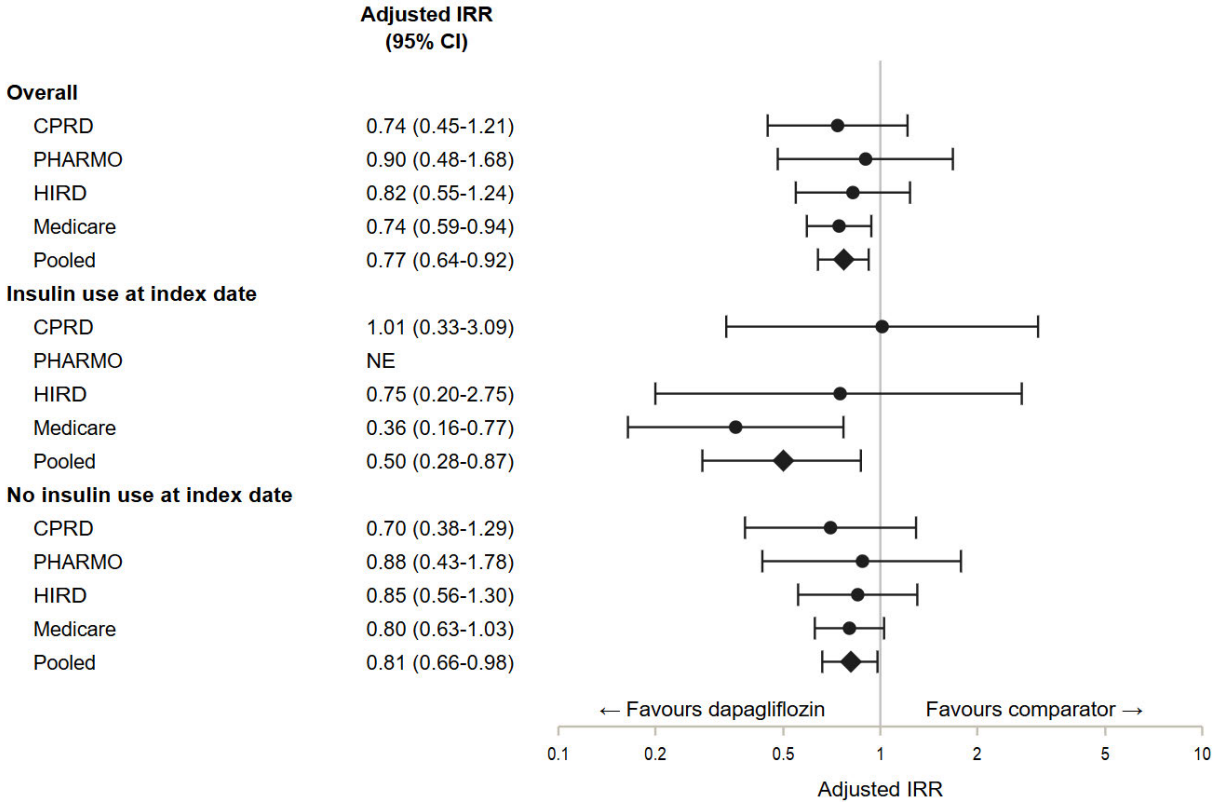
Females and Males Combined

As presented in Figure 16, in the overall cohorts, the propensity score–adjusted IRRs for sex-combined bladder cancer ranged from 0.74 (95% CI, 0.45-1.21) in CPRD to 0.90 (95% CI, 0.48-1.68) in PHARMO, with a pooled adjusted IRR estimate of 0.77 (95% CI, 0.64-0.92). The 95% CIs were wide due to the low number of bladder cancer cases, particularly amongst dapagliflozin new users.

Amongst insulin users, the propensity score–adjusted IRRs for bladder cancer were slightly higher than in the overall cohorts in CPRD and lower than in the overall cohorts in the HIRD and Medicare, with a pooled adjusted IRR of 0.50 (95% CI, 0.28-0.87) based on data from CPRD, the HIRD, and Medicare. The effect estimates were imprecise due to the very low numbers of events observed, particularly amongst dapagliflozin new users. In PHARMO, the propensity score–adjusted IRR in the insulin users cohort was not estimable due to the very small number of bladder cancer cases in both exposure groups.

For all data sources, estimates in the insulin non-users cohorts were similar to those observed in the overall cohorts, as insulin non-users comprised the largest portion of the overall cohort in each data source.

Figure 16 Propensity Score–Adjusted Incidence Rate Ratios for Sex-Combined Bladder Cancer, Overall and Stratified by Insulin Use at the Index Date, by Data Source



Note: The pooled IRR estimates were calculated using Mantel-Haenszel methods, stratified by propensity score stratum, as described in Section 9.9.7.

NE = not estimable.

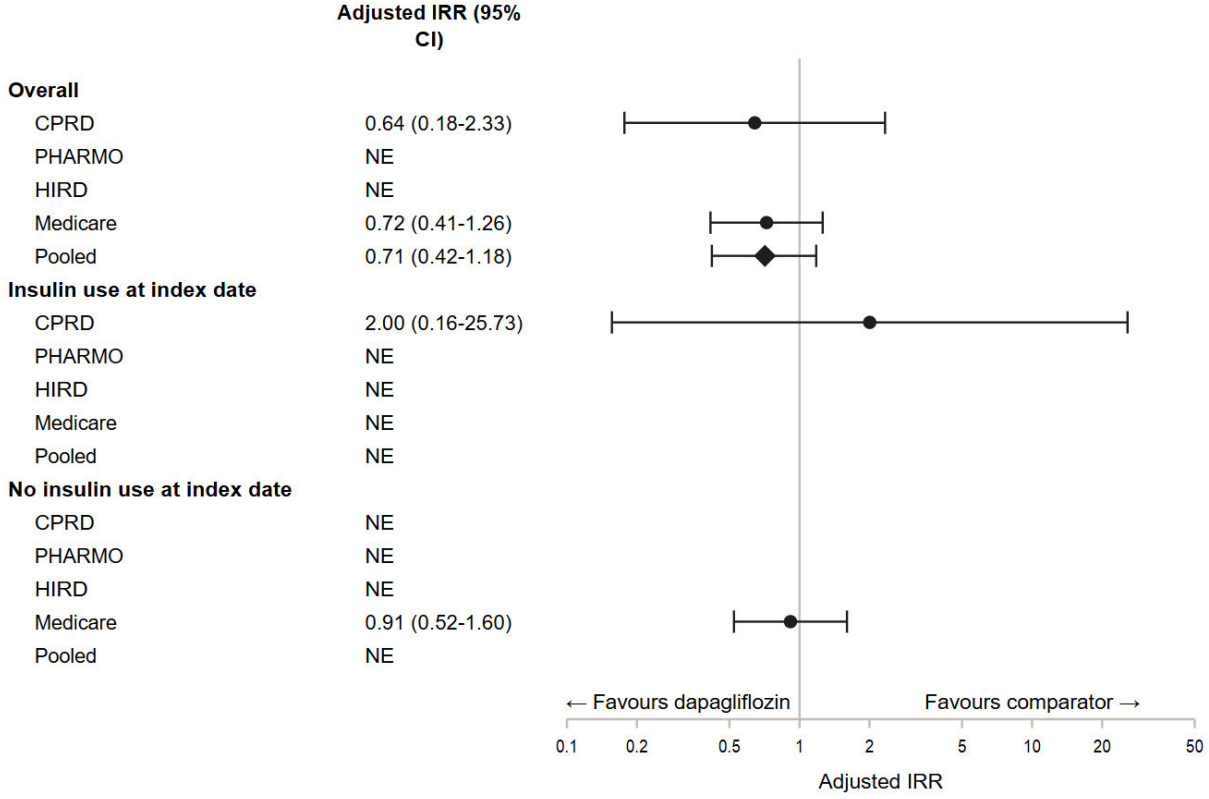
Females

As presented in Figure 17, in the overall cohort, the pooled adjusted IRR for bladder cancer amongst females (0.71; 95% CI, 0.42-1.18) was based on propensity score–adjusted IRRs

from CPRD and Medicare, both of which had wide 95% CIs due to the low number of bladder cancer cases in each of these data sources. In PHARMO and the HIRD, propensity score–adjusted IRRs for bladder cancer amongst females were not estimable for the overall and insulin use–stratified cohorts due to the low number of bladder cancer cases.

Amongst insulin users, a propensity score–adjusted IRR for bladder cancer amongst females was estimable only in CPRD, and, amongst insulin non-users, a propensity score–adjusted IRR was estimable only in Medicare; both IRR estimates had wide 95% CIs.

Figure 17 Propensity Score–Adjusted Incidence Rate Ratios for Female Bladder Cancer, Overall and Stratified by Insulin Use at the Index Date, by Data Source



Note: The pooled IRR estimates were calculated using Mantel-Haenszel methods, stratified by propensity score stratum, as described in Section 9.9.7.

NE = not estimable.

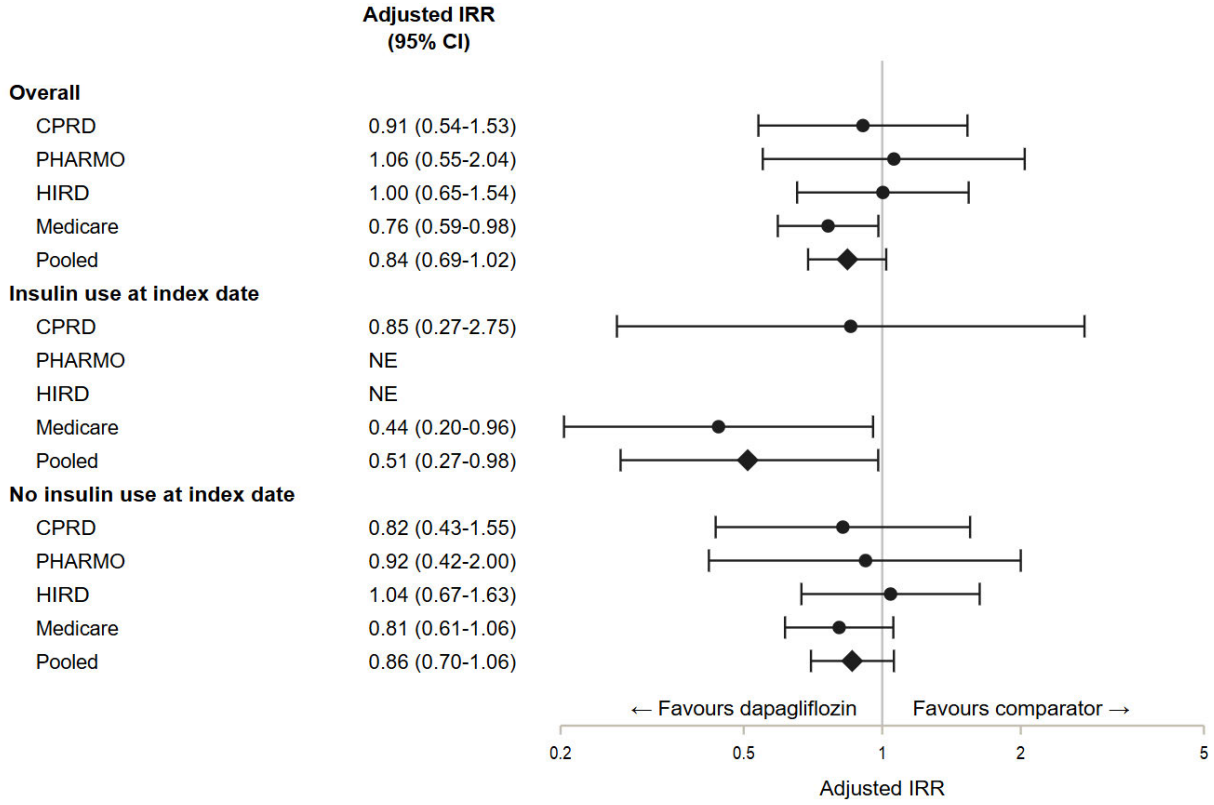
Males

As presented in Figure 18, in the overall cohort, the propensity score–adjusted IRR for bladder cancer amongst males ranged from 0.76 (95% CI, 0.59-0.98) in Medicare to 1.06 (95% CI, 0.55-2.04) in PHARMO, with a pooled adjusted IRR of 0.84 (95% CI, 0.69-1.02). The 95% CIs were wide for CPRD, PHARMO, and the HIRD due to the small number of bladder cancer cases.

Amongst insulin users, the pooled adjusted IRR was 0.51 (95% CI, 0.27-0.98) and was based on the propensity score–adjusted IRRs from CPRD and Medicare, both of which had wide 95% CIs. Propensity score–adjusted IRRs were not estimable amongst insulin users in PHARMO and the HIRD because of the low number of bladder cancer cases amongst dapagliflozin new users.

For all data sources, estimates in the insulin non-users cohorts were similar to those observed in the overall cohorts, as insulin non-users comprised the largest portion of the overall cohort in each data source.

Figure 18 Propensity Score–Adjusted Incidence Rate Ratios for Male Bladder Cancer, Overall and Stratified by Insulin Use at the Index Date, by Data Source



Note: The pooled IRR estimates were calculated using Mantel-Haenszel methods, stratified by propensity score stratum, as described in Section 9.9.7.

NE = not estimable.

10.4.2.3.3 Unadjusted and Propensity Score–Adjusted Comparisons, by Pioglitazone Use at the Index Date

See results Appendix J, BladderCa Table 23, for the sex-combined bladder cancer cohorts after propensity score trimming stratified by pioglitazone use and by sex for each data source. This analysis addresses primary objective 2 (see Section 7.2.1).

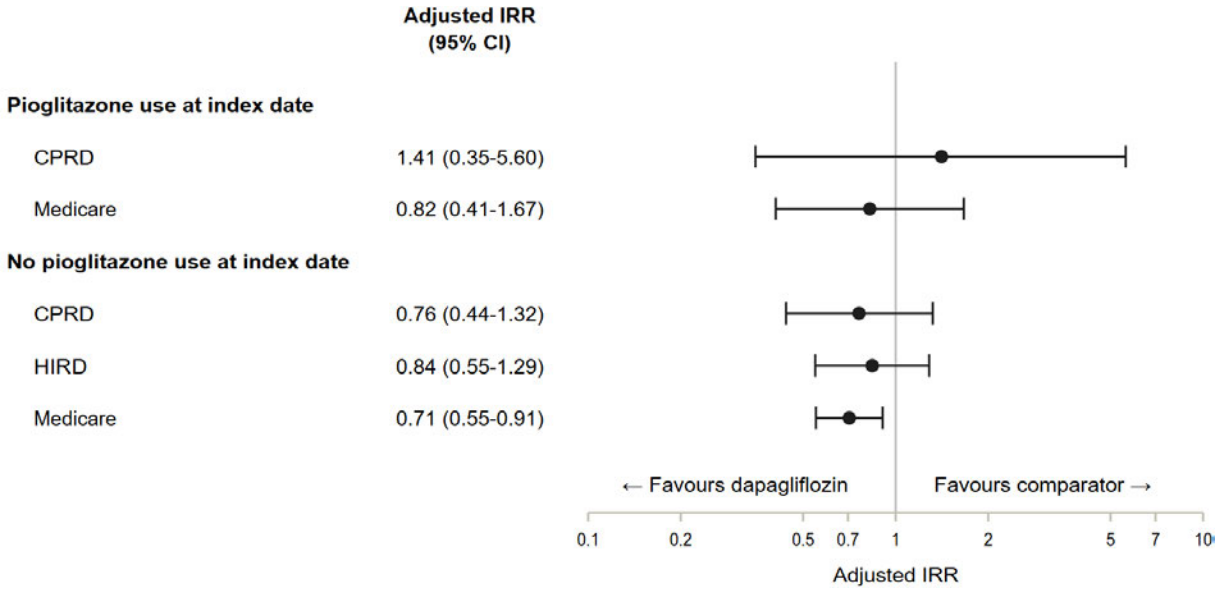
In PHARMO, the pioglitazone use–stratified analyses were not conducted because no bladder cancer cases occurred in the dapagliflozin exposure group and a single bladder cancer case occurred in the comparator AD exposure group in the pioglitazone users cohort.

Unadjusted Comparisons. In the pioglitazone users cohort, the unadjusted IRR for sex-combined bladder cancer was 1.63 (95% CI, 0.29-6.17) in CPRD, 1.26 (95% CI, 0.24-4.13) in the HIRD, and 0.89 (95% CI, 0.38-1.78) in Medicare.

In the pioglitazone non-users cohort, the unadjusted IRR for sex-combined bladder cancer was 0.55 (95% CI, 0.31-0.93) in CPRD, 0.94 (95% CI, 0.59-1.43) in the HIRD, and 0.70 (95% CI, 0.54-0.89) in Medicare.

Adjusted Comparisons. The propensity score–adjusted IRR estimates for sex-combined bladder cancer by pioglitazone use at the index date are presented in Figure 19. In the pioglitazone users cohort, the propensity score–adjusted IRR for the sex-combined bladder cancer outcome was 1.41 (95% CI, 0.35-5.60) in CPRD and 0.82 (95% CI, 0.41-1.67) in Medicare; the effect estimates were imprecise due to the low number of bladder cancer cases. In the HIRD, the propensity score–adjusted IRR was not estimable amongst pioglitazone users due to the low number of bladder cancer cases in this group. Amongst pioglitazone non-users, the propensity score–adjusted IRRs for sex-combined bladder cancer ranged from 0.71 (95% CI, 0.55-0.91) in Medicare to 0.84 (95% CI, 0.55-1.29) in the HIRD.

Figure 19 Propensity Score–Adjusted Incidence Rate Ratios for Sex-Combined Bladder Cancer Stratified by Pioglitazone Use at the Index Date, by Data Source



Note: Data sources in which the propensity score–adjusted IRRs were not estimable are not displayed.

Stratification by sex resulted in non-estimable propensity score–adjusted IRR estimates for female pioglitazone users. For female pioglitazone non-users, the propensity score–adjusted IRR point estimates were 0.70 (95% CI, 0.19-2.59) in CPRD, 0.38 (95% CI, 0.09-1.68) in the HIRD, and 0.72 (95% CI, 0.41-1.29) in Medicare. All propensity score–adjusted IRR estimates were imprecise because of the small number of bladder cancer cases.

For males, the propensity score–adjusted IRR estimates amongst pioglitazone users were 1.81 (95% CI, 0.43-7.68) in CPRD and 0.93 (95% CI, 0.46-1.89) in Medicare and were not estimable in the HIRD. Amongst male pioglitazone non-users, the propensity score–adjusted IRR was 0.78 (95% CI, 0.43-1.42) in CPRD, 0.98 (95% CI, 0.63-1.53) in the HIRD, and 0.72 (95% CI, 0.55-0.95) in Medicare.

10.4.2.4 HCRU During Follow-up: Secondary Outcome

See Appendix J, BladderCa Table 15, for yearly HCRU descriptive results for each exposure group overall and stratified by insulin use at the index date in the propensity score–trimmed cohorts for each data source. This analysis addresses secondary objective 1 (see Section 7.2.2).

Assessment of HCRU during follow-up was relevant for assessing the potential role of HCRU as a mediator of the total effect of dapagliflozin on bladder cancer incidence if an effect was observed (see Section 9.9.3.7 for the criterion for conducting the mediation analysis).

10.4.2.4.1 Descriptive Assessment

In the sex-combined bladder cancer cohorts, the frequency of HCRU measures—including outpatient visits, urologist visits, other specialist visits, and hospitalisations—did not show relevant differences between dapagliflozin users and comparator AD users across years of follow-up in all data sources. Overall, higher average numbers of outpatient visits/encounters were reported in both exposure groups during the first year of follow-up in CPRD (9), in the HIRD (21), and in Medicare (14) than in subsequent years of follow-up. In PHARMO, the average number of outpatient visits, assessed only amongst the subset of patients with GP data available, were higher during the last years of follow-up (up to 21 visits in the eighth and ninth years in both exposure groups) although based on a small number of new users. The higher number of outpatient visits/encounters in the HIRD than in other data sources was driven by the broader definition including all outpatient encounters, eg, laboratory tests. The proportion of new users with at least one urologist visit across years of follow-up had slight differences across data sources (~1% in CPRD, up to ~7% in PHARMO, up to 6% in the HIRD, and up to 11% in Medicare), with similar results found in females and males. Across years of follow-up in all data sources, the proportion of new users with at least 1 bladder biopsy, urinary cytology, or cystoscopy was < 1% in CPRD, up to 1% in the HIRD, and up to 2% in Medicare (the latter referring to new users with at least 1 cystoscopy). Information on bladder biopsies, urinary cytologies, and cystoscopies was not available for patients in PHARMO.

10.4.2.4.2 Mediation Analysis to Assess HCRU as a Potential Mediator

The assessment of HCRU as a potential mediator of the association between exposure to a treatment group and cancer was not conducted in the bladder cancer cohort because the lower bound of the 95% CI of the propensity score–adjusted IRRs for bladder cancer was not above 1.5 in any data source (Section 10.4.1.4.2). Therefore, the criterion for conducting the mediation analysis was not met.

10.4.3 Female Composite Cancer

10.4.3.1 Incidence Analysis

10.4.3.1.1 Unadjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, FemaleCompositeCa Table 7, for results for the female composite cancer outcome after propensity score trimming for the overall and insulin use–stratified cohorts for each data source (and stratified by age for CPRD and PHARMO only). Unadjusted incidence rates for the female composite cancer outcome for each propensity score stratum are presented for each data source in Appendix J, FemaleCompositeCa Table 8. This analysis addresses secondary objective 4 (see Section 7.2.2).

The incidence analyses for the female composite cancer outcome were based on the following person-years of exposure to dapagliflozin in each data source: 16,783 person-years in CPRD, 3,449 person-years in PHARMO, 49,962 person-years in the HIRD, and 60,593 person-years in Medicare. In the overall cohorts, the lowest unadjusted incidence rates for the female composite cancer outcome were observed in the HIRD (dapagliflozin, 42.63 [95% CI, 37.10-48.76]; comparator AD, 36.48 [95% CI, 34.70-38.34]), and the highest unadjusted incidence rates were observed in PHARMO (dapagliflozin, 66.69 [95% CI, 42.27-100.06]; comparator AD, 72.55 [95% CI, 64.97-80.78]) (Appendix J, FemaleCompositeCa Table 7). Overall, the unadjusted incidence rates for the female composite cancer outcome were higher in the dapagliflozin group than in the comparator AD group in the HIRD. Conversely, the unadjusted incidence rates were lower in the dapagliflozin group than in the comparator AD group in CPRD and PHARMO and were similar in both exposure groups in Medicare.

Amongst insulin users, the unadjusted incidence rates for the female composite cancer outcome in both exposure groups were generally higher than in the overall unstratified cohorts in CPRD and the HIRD; in PHARMO, this was true only amongst dapagliflozin new users. In Medicare, the unadjusted incidence rate amongst dapagliflozin new users was lower than the incidence rate in the overall stratified cohort. The pattern observed in the overall cohorts was also observed in the insulin non-users cohort in all data sources.

10.4.3.1.2 Propensity Score–Adjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, FemaleCompositeCa Table 10, for the propensity score–adjusted incidence rate results for the female composite cancer outcome in the overall cohort and the insulin use–stratified cohorts for each data source (and stratified by age for CPRD and PHARMO). This analysis addresses secondary objective 4 (see Section 7.2.2).

Adjustment of incidence rates was accomplished when propensity score–stratified rates in each exposure group were standardised to the person-years in the dapagliflozin group; this approach resulted in rates for the dapagliflozin group which were the same as, or comparable to, the unadjusted rates. For the overall cohort (not stratified by insulin use), the propensity score–adjusted incidence rates for the female composite cancer outcome per 10,000 person-years were generally similar for dapagliflozin and comparator AD new users, respectively, in CPRD (42.90; 95% CI, 33.57-54.03 and 43.79; 95% CI, 36.58-51.63), the HIRD (42.63; 95% CI, 37.10-48.76 and 38.90; 95% CI, 36.73-41.16), and Medicare (44.39; 95% CI, 39.25-50.03 and 43.76; 95% CI, 42.28-45.28). In PHARMO, the highest propensity score–adjusted incidence rates were observed in both exposure groups, with lower rates in dapagliflozin new users (66.69; 95% CI, 42.27-100.06) than in comparator AD new users (72.11; 95% CI, 62.28-82.78). The overall propensity score–adjusted incidence rates for the female composite cancer outcome were generally higher in those aged 65 years or older than in those aged younger than 65 years in CPRD; in PHARMO, this was true only amongst the comparator AD new users.

Amongst insulin users, the propensity score–adjusted incidence rate for the female composite cancer outcome was higher in dapagliflozin than in comparator AD new users, respectively, in PHARMO (126.89; 95% CI, 46.56-276.18 and 77.41; 95% CI, 31.38-137.38) and the HIRD (53.76; 95% CI, 38.23-73.49 and 45.16; 95% CI, 37.58-53.65). Conversely, the propensity score–adjusted incidence rates were lower in dapagliflozin than in comparator AD new users, respectively, in CPRD (46.03; 95% CI, 26.82-73.70 and 53.50; 95% CI, 30.28-83.20) and Medicare (37.71; 95% CI, 27.61-50.30 and 42.05; 95% CI, 38.54-45.73).

For insulin non-users, the propensity score–adjusted incidence rates for the female composite cancer outcome were generally similar to those in the overall unstratified cohorts.

10.4.3.2 Cumulative Treatment Duration

The descriptive results on the female composite cancer outcome incidence rates (in addition to the number of patients, person-years, and female composite cancer outcome events) for each exposure group by category of cumulative treatment duration (in years), after propensity score trimming for the overall and insulin use–stratified cohorts, are presented for each data source in Appendix J, FemaleCompositeCa Table 9.

As described in Section 9.4.2.4, person-time and crude incidence rates for the female composite cancer outcome were calculated for each of the mutually exclusive categories of cumulative treatment duration, categorised as < 1 year, 1 to < 2 years, 2 to < 5 years, and ≥ 5 years.

In the overall cohorts, across all treatment duration categories in CPRD (except for treatment duration 2 to < 5 yrs), PHARMO, and Medicare, dapagliflozin new users had a lower incidence rate for the female composite cancer outcome than comparator AD new users. In the HIRD, for all treatment duration categories, dapagliflozin new users had higher incidence rates for female composite cancers than comparator AD new users. Across all data sources, for the treatment duration category ≥ 5 years, there were few or no cases of the female composite cancer outcome in the dapagliflozin group.

Amongst insulin users, the lower number of events across treatment duration categories in CPRD, PHARMO, and the HIRD resulted in wide CIs for the crude incidence rates. In the treatment duration categories 1 to < 2 years and 2 to < 5 years, the incidence rate was lower in dapagliflozin new users than in comparator AD new users in Medicare and higher in the HIRD. In CPRD, the incidence rate was lower in dapagliflozin new users than in comparator AD new users in the category 1 to < 2 years and similar in both groups in the category 2 to < 5 years. In PHARMO, the incidence rate was higher in dapagliflozin new users than in comparator AD new users across the treatment duration categories < 1 year, 1 to < 2 years and 2 to < 5 years. In the treatment duration category ≥ 5 years, there were no cases in the dapagliflozin group across all data sources.

10.4.3.3 Comparative Analysis

10.4.3.3.1 Unadjusted Comparisons, Overall and by Insulin Use at the Index Date

See Appendix J, FemaleCompositeCa Table 7, for the results overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) after propensity score trimming for each data source. Unadjusted IRRs for female composite cancers for each propensity score stratum are displayed in Appendix J, FemaleCompositeCa Table 8, for each data source. This analysis addresses secondary objective 4 (see Section 7.2.2).

The overall unadjusted IRR for the female composite cancer outcome was 0.82 (95% CI, 0.63-1.06) in CPRD, 0.92 (95% CI, 0.57-1.40) in PHARMO, 1.17 (95% CI, 1.01-1.35) in the HIRD, and 0.98 (95% CI, 0.86-1.10) in Medicare. The same pattern was observed for the unadjusted IRR estimates in both insulin use–stratified cohorts in CPRD, the HIRD, and Medicare. In PHARMO, higher unadjusted IRRs for female composite cancers were observed amongst insulin users (1.78; 95% CI, 0.61-4.31) than amongst insulin non-users (0.82; 95% CI, 0.47-1.33), although IRRs were imprecise due to the low number of new users and cases, particularly amongst dapagliflozin new users. Analyses stratified by age yielded the

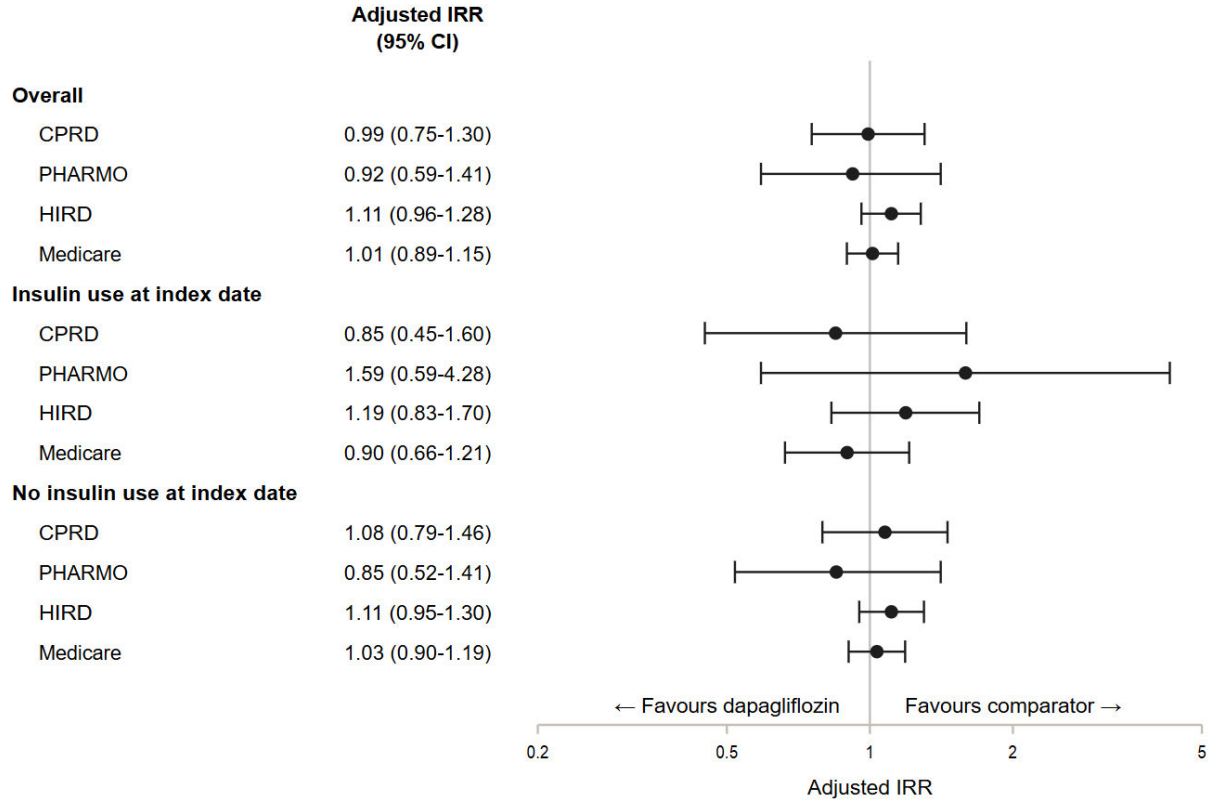
following unadjusted IRRs amongst those aged < 65 years: 0.99 (95% CI, 0.68-1.41) in CPRD and 1.27 (95% CI, 0.66-2.26) in PHARMO; and amongst those aged ≥ 65 years: 0.83 (95% CI, 0.54-1.23) in CPRD and 0.73 (95% CI, 0.35-1.37) in PHARMO.

10.4.3.3.2 Propensity Score–Adjusted Comparisons, Overall and by Insulin Use at the Index Date

See Appendix J, FemaleCompositeCa Table 10, for the propensity score–adjusted IRR results for the female composite cancer outcome overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) for each data source. This analysis addresses secondary objective 4 (see Section 7.2.2).

For the overall cohorts, propensity score–adjusted IRRs for the female composite cancer outcome ranged from 0.92 (95% CI, 0.59-1.41) in PHARMO to 1.11 (95% CI, 0.96-1.28) in the HIRD (Figure 20). Similar findings were observed for the propensity score–adjusted IRR estimates in both insulin use–stratified cohorts in all data sources, except for a higher IRR in PHARMO amongst the insulin users cohort than amongst the overall and insulin non-users cohorts. Analyses stratified by age yielded the following propensity score–adjusted IRRs amongst those aged < 65 years: 1.10 (95% CI, 0.76-1.60) in CPRD and 1.24 (95% CI, 0.68-2.25) in PHARMO; and amongst those aged ≥ 65 years: 0.91 (95% CI, 0.61-1.36) in CPRD and 0.75 (95% CI, 0.40-1.44) in PHARMO.

Figure 20 Propensity Score–Adjusted Incidence Rate Ratios for the Female Composite Cancer Outcome Stratified by Insulin Use at the Index Date, by Data Source



10.4.4 Male Composite Cancer

10.4.4.1 Incidence Analysis

10.4.4.1.1 Unadjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, MaleCompositeCa Table 7, for the male composite cancer outcome after propensity score trimming for the overall and insulin use–stratified cohorts for each data source (and stratified by age for CPRD and PHARMO only). Unadjusted incidence rates for the male composite cancer outcome for each propensity score stratum are displayed for each data source in Appendix J, MaleCompositeCa Table 8. This analysis addresses secondary objective 3 (see Section 7.2.2).

The incidence analyses for the male composite cancer outcome were based on the following person-years of exposure to dapagliflozin in each data source: 25,148 person-years in CPRD, 4,805 person-years in PHARMO, 72,100 person-years in the HIRD, and 63,255 person-years in Medicare. In the overall cohorts, the lowest unadjusted incidence rates for the male composite cancer outcome were observed in the HIRD (dapagliflozin, 49.24 [95% CI, 44.25-54.63]; comparator AD, 53.66 [95% CI, 51.53-55.86]), and the highest unadjusted incidence

rates were observed in Medicare (dapagliflozin, 100.86 [95% CI, 93.19-109.00]; comparator AD, 103.45 [95% CI, 101.59-105.34]) and PHARMO (dapagliflozin, 93.66 [95% CI, 68.32-125.32]; comparator AD, 115.13 [95% CI, 106.17-124.64]). The unadjusted incidence rate for male composite cancers was lower for dapagliflozin than for comparator AD across all data sources.

Amongst insulin users, the unadjusted incidence rates for male composite cancers in dapagliflozin new users were lower than in comparator AD new users in CPRD and the HIRD and higher in dapagliflozin new users than in comparator AD new users in PHARMO and Medicare. The pattern observed for the unadjusted incidence rates in the overall cohorts was also observed in the insulin non-users cohorts across all data sources.

10.4.4.1.2 Propensity Score–Adjusted Incidence Rates, Overall and by Insulin Use at the Index Date

See Appendix J, MaleCompositeCa Table 10, for the propensity score–adjusted incidence rate results for the male composite cancer outcome in the overall cohort and the insulin use–stratified cohorts for each data source (and stratified by age for CPRD and PHARMO only). This analysis addresses secondary objective 3 (see Section 7.2.2).

Adjustment of incidence rates was accomplished when propensity score–stratified rates in each exposure group were standardised to the person-years in the dapagliflozin group. For the overall cohort (not stratified by insulin use), the propensity score–adjusted incidence rates for male composite cancers per 10,000 person-years were somewhat lower for dapagliflozin new users than for comparator AD new users across all data sources. The propensity score–adjusted incidence rates for dapagliflozin and comparator AD new users, respectively, were highest in PHARMO (93.66; 95% CI, 68.32- 125.32 and 125.89; 95% CI, 112.15-140.47) and Medicare (100.86; 95% CI, 93.19-109.00 and 102.41; 95% CI, 99.89-104.97) and were somewhat lower in CPRD (60.84; 95% CI, 51.58-71.28 and 73.05; 95% CI, 64.62-81.97) and the HIRD (49.24; 95% CI, 44.25-54.63 and 55.90; 95% CI, 53.23-58.66). In CPRD and PHARMO, the overall propensity score–adjusted incidence rates stratified by age were considerably higher in males aged 65 years or older than in males aged younger than 65 years.

Amongst insulin users, the propensity score–adjusted incidence rate for male composite cancers was lower in dapagliflozin new users than in comparator AD new users, respectively, in CPRD (55.56; 95% CI, 35.22-83.36 and 80.02; 95% CI, 55.34-109.51) and the HIRD (47.53; 95% CI, 34.80-63.40 and 65.50; 95% CI, 56.88-74.92) and higher in dapagliflozin new users than in comparator AD new users, respectively, in Medicare (106.91; 95% CI, 88.78-127.65 and 100.11; 95% CI, 93.72-106.75). In PHARMO, propensity score–adjusted incidence rates amongst insulin users were not estimable due to a low number of new users and cases. In the insulin non-users cohort, the propensity score–adjusted incidence rates were

similar to the propensity score–adjusted incidence rates in the overall cohorts in all data sources.

10.4.4.2 Cumulative Treatment Duration

The descriptive results on the male composite cancer outcome incidence rates (in addition to the number of patients, person-years, and male composite cancer outcome events) for each exposure group by category of cumulative treatment duration (in years), after propensity score trimming, for the overall and insulin use–stratified cohorts, are presented for each data source in Appendix J, MaleCompositeCa Table 9.

In the overall cohorts, across all treatment duration categories in CPRD (except for the category ≥ 5 years), PHARMO, and Medicare, dapagliflozin new users had a lower incidence rate for the male composite cancer outcome than comparator AD new users. In the HIRD, dapagliflozin new users had higher incidence rates for the male composite cancer outcome than comparator AD new users in the treatment duration categories ≤ 1 year, 2 to < 5 years, and ≥ 5 years. Across all data sources, for the treatment duration category ≥ 5 years, there were few cases of the male composite cancer outcome in the dapagliflozin group.

Amongst insulin users, for all data sources, the lower number of male composite cancer outcome events across treatment duration categories resulted in wide CIs for the crude incidence rates, notably in CPRD, PHARMO, and the HIRD.

10.4.4.3 Comparative Analysis

10.4.4.3.1 Unadjusted Comparison, Overall and by Insulin Use at the Index Date

See Appendix J, MaleCompositeCa Table 7, for unadjusted IRR estimates overall, stratified by insulin use, and stratified by age (in CPRD and PHARMO only) after propensity score trimming for each data source. Unadjusted IRRs for the male composite cancer outcome for each propensity score stratum are presented for each data source in Appendix J, MaleCompositeCa Table 8. This analysis addresses secondary objective 3 (see Section 7.2.2).

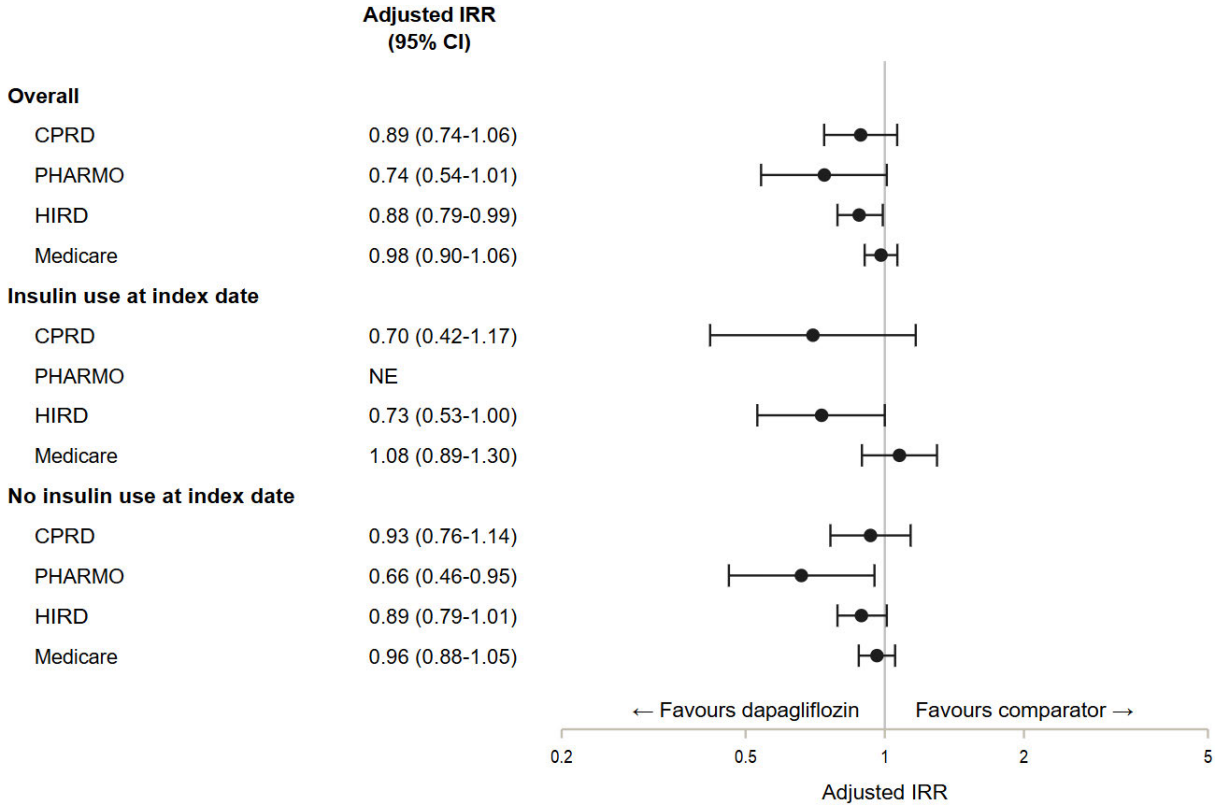
The overall unadjusted IRR estimates for the male composite cancer outcome was 0.76 (95% CI, 0.63-0.91) in CPRD, 0.81 (95% CI, 0.59-1.10) in PHARMO, 0.92 (95% CI, 0.82-1.03) in the HIRD, and 0.97 (95% CI, 0.90-1.06) in Medicare. The same pattern was observed for the unadjusted IRR estimates in both insulin use–stratified cohorts in CPRD, the HIRD and Medicare. In PHARMO, a higher IRR was observed amongst insulin users (1.42; 95% CI, 0.66-2.82). Analyses stratified by age yielded the following unadjusted IRRs amongst those aged < 65 years: 1.08 (95% CI, 0.82-1.41) in CPRD and 0.68 (95% CI, 0.35-1.19) in PHARMO; amongst those aged ≥ 65 years: 0.75 (95% CI, 0.58-0.95) in CPRD and 0.99 (95% CI, 0.67-1.42) in PHARMO.

10.4.4.3.2 PROPENSITY SCORE–ADJUSTED COMPARISON, OVERALL AND BY INSULIN USE AT THE INDEX DATE

See Appendix J, MaleCompositeCa Table 10, for propensity score–adjusted IRRs for the male composite cancer outcome overall, stratified by insulin use at the index date, and stratified by age (in CPRD and PHARMO only) for each data source. This analysis addresses secondary objective 3 (see Section 7.2.2).

For the overall cohorts, propensity score–adjusted IRR estimates for the male composite cancer outcome ranged from 0.74 (95% CI, 0.54-1.01) in PHARMO to 0.98 (95% CI, 0.90-1.06) in Medicare (Figure 21). The same pattern was observed for the propensity score–adjusted IRR estimates in both insulin use–stratified cohorts in all data sources, except for insulin users in PHARMO, for which the propensity score–adjusted IRR was not estimable. Analyses stratified by age yielded the following propensity score–adjusted IRRs amongst those aged < 65 years: 1.18 (95% CI, 0.90-1.55) in CPRD and 0.65 (95% CI, 0.37-1.15) in PHARMO; and amongst those aged ≥ 65 years: 0.76 (95% CI, 0.60-0.98) in CPRD and 0.93 (95% CI, 0.64-1.35) in PHARMO.

Figure 21 Propensity Score–Adjusted Incidence Rate Ratios for the Male Composite Cancer Outcome Stratified by Insulin Use at the Index Date, by Data Source



NE = not estimable.

10.4.5 Analysis of Time-Varying Variables

10.4.5.1 Diabetes Severity

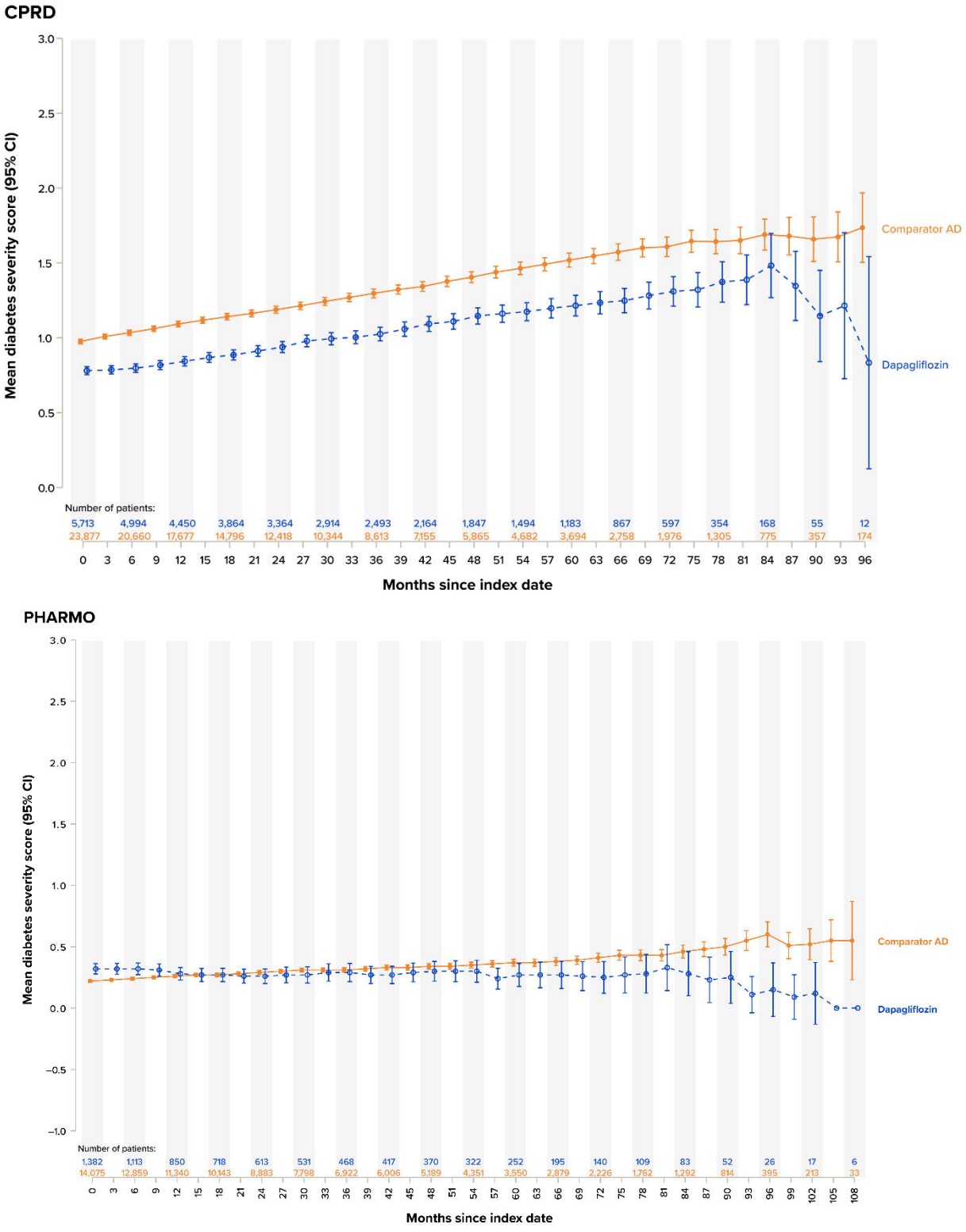
The results of the analyses assessing the changes in diabetes severity score during follow-up are displayed for the overall propensity score–trimmed female breast cancer cohorts and sex-combined bladder cancer cohorts, respectively, in Figure 22 and in Appendix J, BladderCa Figure 3.

10.4.5.1.1 Female Breast Cancer

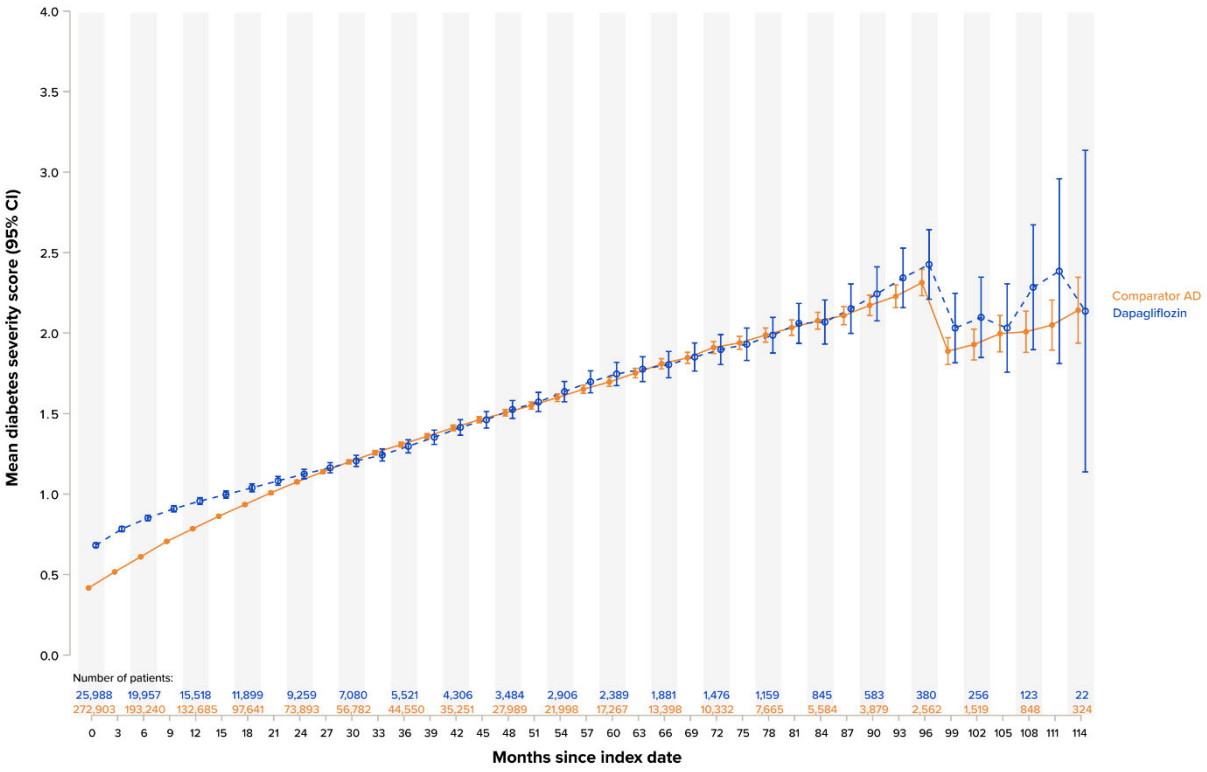
For the overall female breast cancer cohorts, the mean diabetes severity score generally increased steadily in both exposure groups throughout most of follow-up in CPRD, PHARMO, the HIRD, and Medicare; whereas, in PHARMO, the mean diabetes severity score remained stable for both exposure groups throughout most of the follow-up period (Figure 22). In the dapagliflozin group in CPRD, a decline in mean diabetes severity score was observed at approximately 84 months; a similar finding was observed in both exposure groups in the HIRD at 96 months. This may be due, in part, to the decline in sample size around this time, with wide CIs observed in the dapagliflozin group. In CPRD, mean diabetes severity score was slightly higher in the comparator AD group than in the dapagliflozin group across 96 months of follow-up. Similar findings were observed in Medicare between 6 and 90 months of follow-up, with comparable mean scores observed during other time intervals. Within the HIRD, mean diabetes severity scores were slightly higher in the dapagliflozin new users from index date to approximately 21 months, after which the mean diabetes severity scores were comparable in both exposure groups with overlapping CIs.

In PHARMO, indicators of diabetes severity were based on only hospital diagnoses, which do not accurately reflect the severity of conditions that are more likely to be treated in general practice or ambulatory health care settings instead of in a hospital setting. This limitation prevents a proper assessment of the changes in diabetes severity during follow-up in this data source.

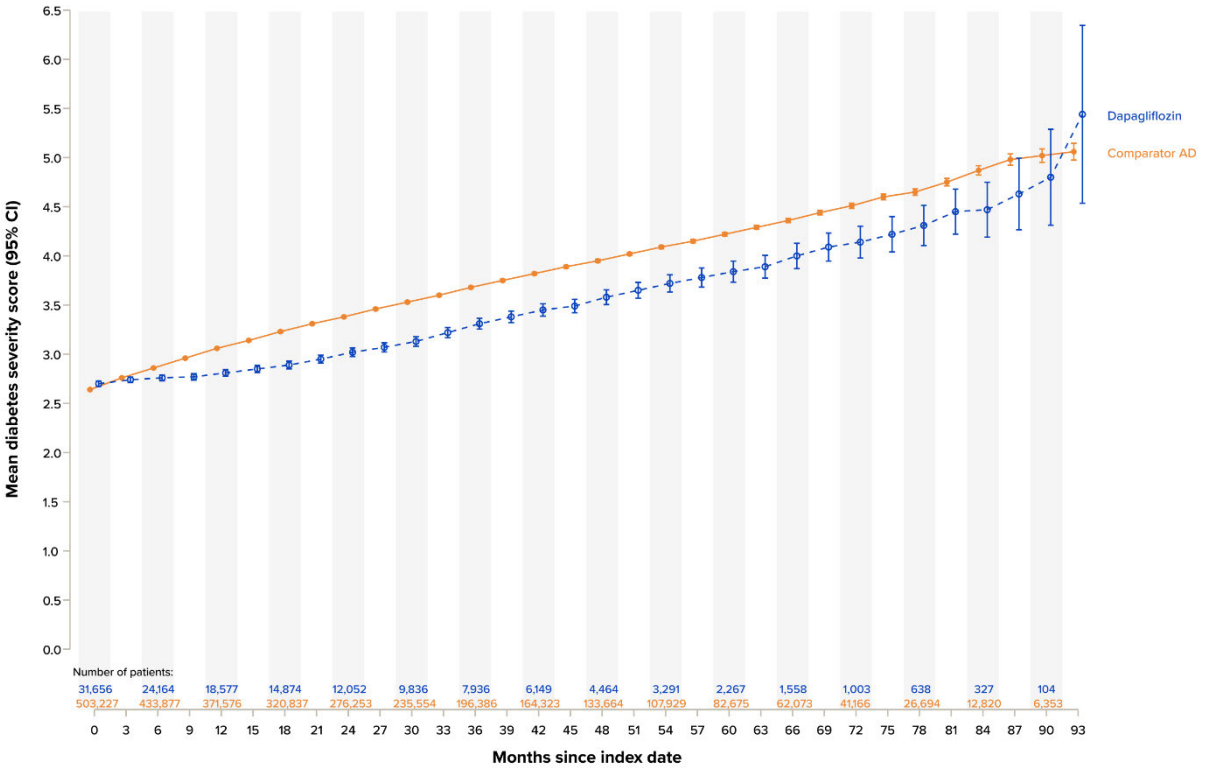
Figure 22 **Diabetes Severity Score Over Time in New Users of Dapagliflozin and of Comparator ADs in the Female Breast Cancer Cohort, by Data Source**



HIRD



Medicare



10.4.5.1.1 BLADDER CANCER

The changes in diabetes severity score during follow-up for the sex-combined bladder cancer cohorts were nearly identical to those for the female breast cancer cohorts across all data sources. Appendix J, BladderCa Figure 3, shows these results for each data source.

10.4.5.2 Treatment Changes

The results of the analyses assessing the change in intensity of AD treatments during follow-up are displayed in Figure 23 for the female breast cancer cohort for each data source and in Appendix J, BladderCa Figure 4, for the sex-combined bladder cancer cohorts for each data source.

10.4.5.2.1 FEMALE BREAST CANCER

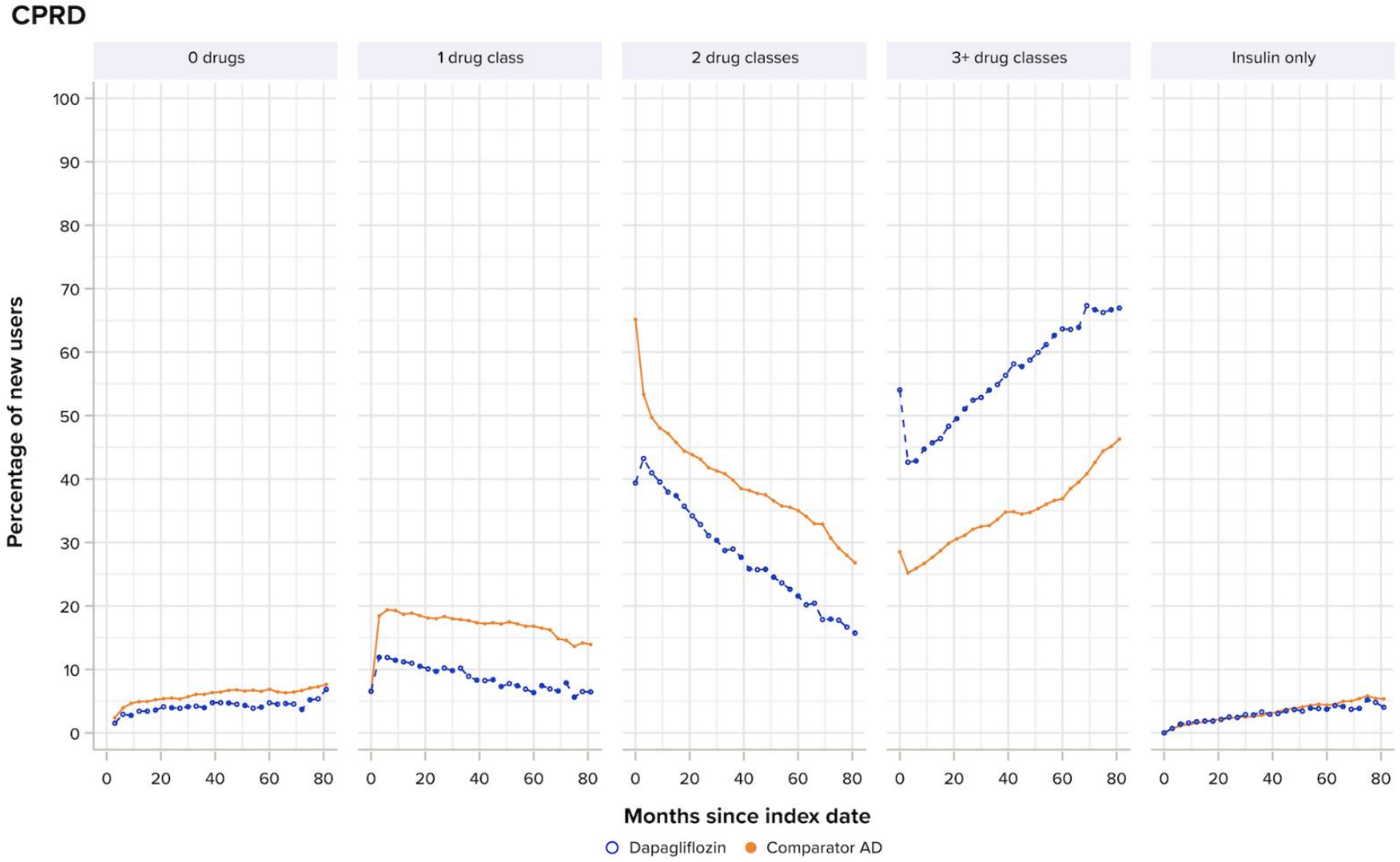
In the overall propensity score–trimmed female breast cancer cohort, the trends observed suggest a decrease in the proportion of new users of dapagliflozin and of comparator ADs taking one AD class or two AD classes across CPRD, PHARMO, the HIRD, and Medicare. Simultaneously, an increase was observed across all data sources in new users who were taking at least three AD classes.

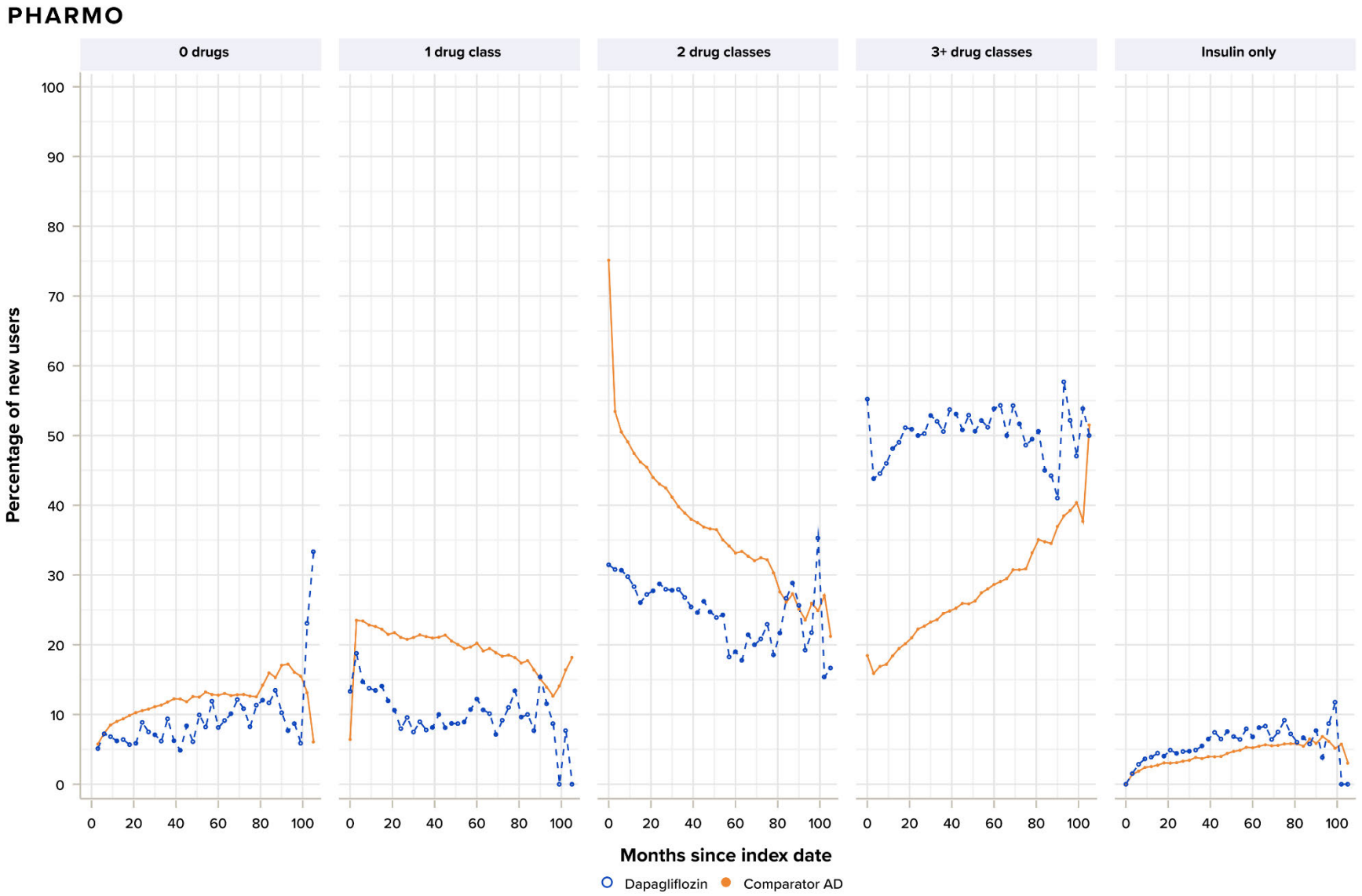
In CPRD, a steady decrease was observed in the proportion of new users taking one AD class, a sharp decline in those taking two AD classes, and an increase over the study period in new users taking at least three AD classes. Conversely, a slightly different pattern was observed in the HIRD, a US data source with a slightly younger population and more recent years of data (100 months of follow-up). In the HIRD, a decrease in the proportion of new users taking one AD was observed for most of the study period except for more recent time when a sharp increase was observed amongst new users of dapagliflozin. Around the same period, the steady increase in the proportion of new users taking at least three AD classes reversed, and a sharp decline was observed. In PHARMO, the proportion of new users of comparator ADs with two or fewer AD classes was higher than the dapagliflozin proportion over most of the follow-up period; however, the reverse was observed when assessing the proportions with at least three AD classes. The proportion of new users of comparator ADs with two AD classes declined steadily over the follow-up period, whilst the proportion with at least three AD classes increased steadily over the follow-up period. The patterns for number of AD classes in new users of dapagliflozin were relatively stable over the follow-up period, with no obvious shifts. In Medicare, steady declines were observed in the proportion of new users taking one and two AD classes, and a slight increase was observed over follow-up in new users taking at least three AD medications across both exposure groups.

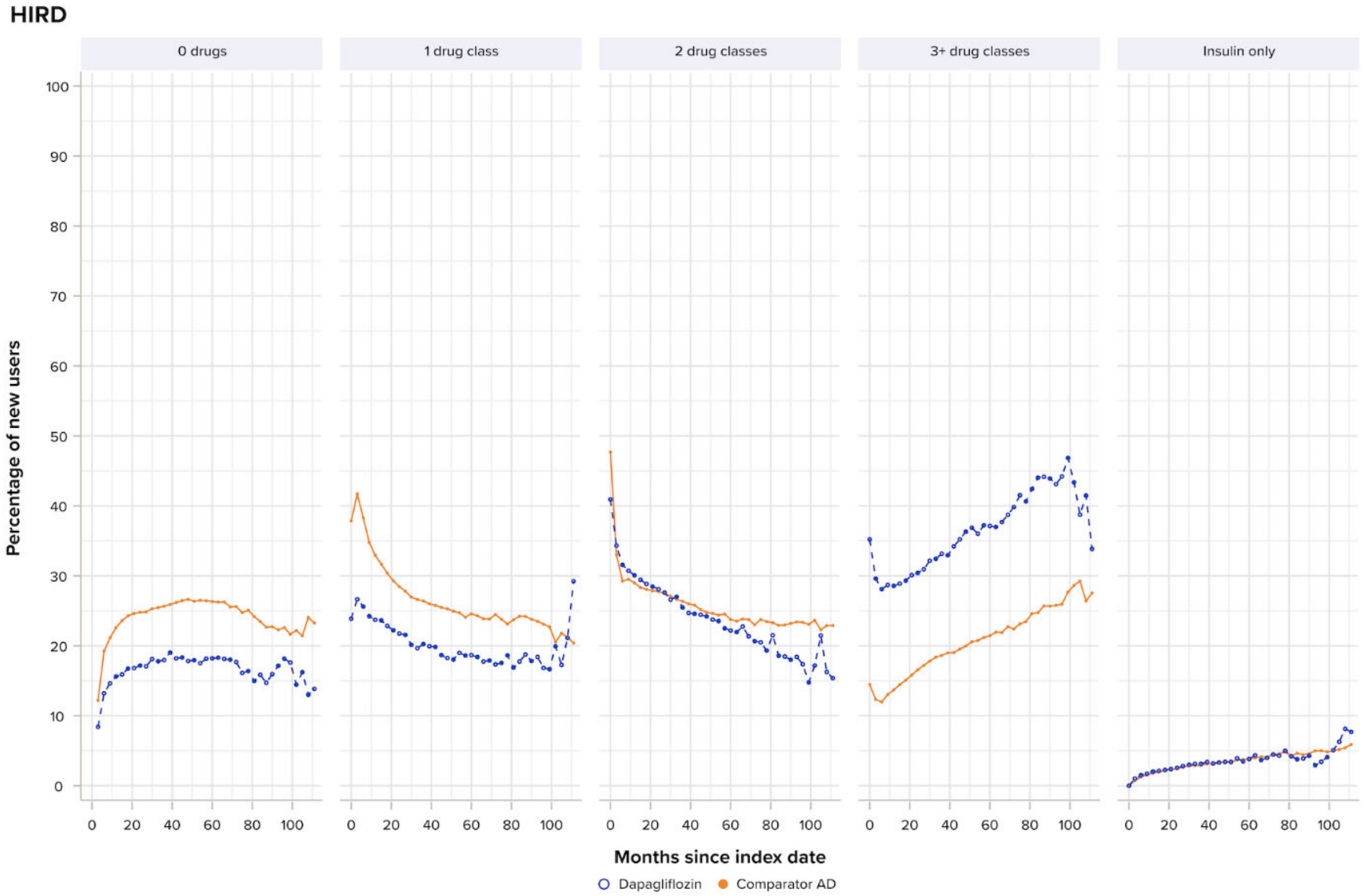
In general, the proportion of new users taking one or two AD classes was higher amongst comparator AD new users than amongst dapagliflozin new users; the opposite was observed

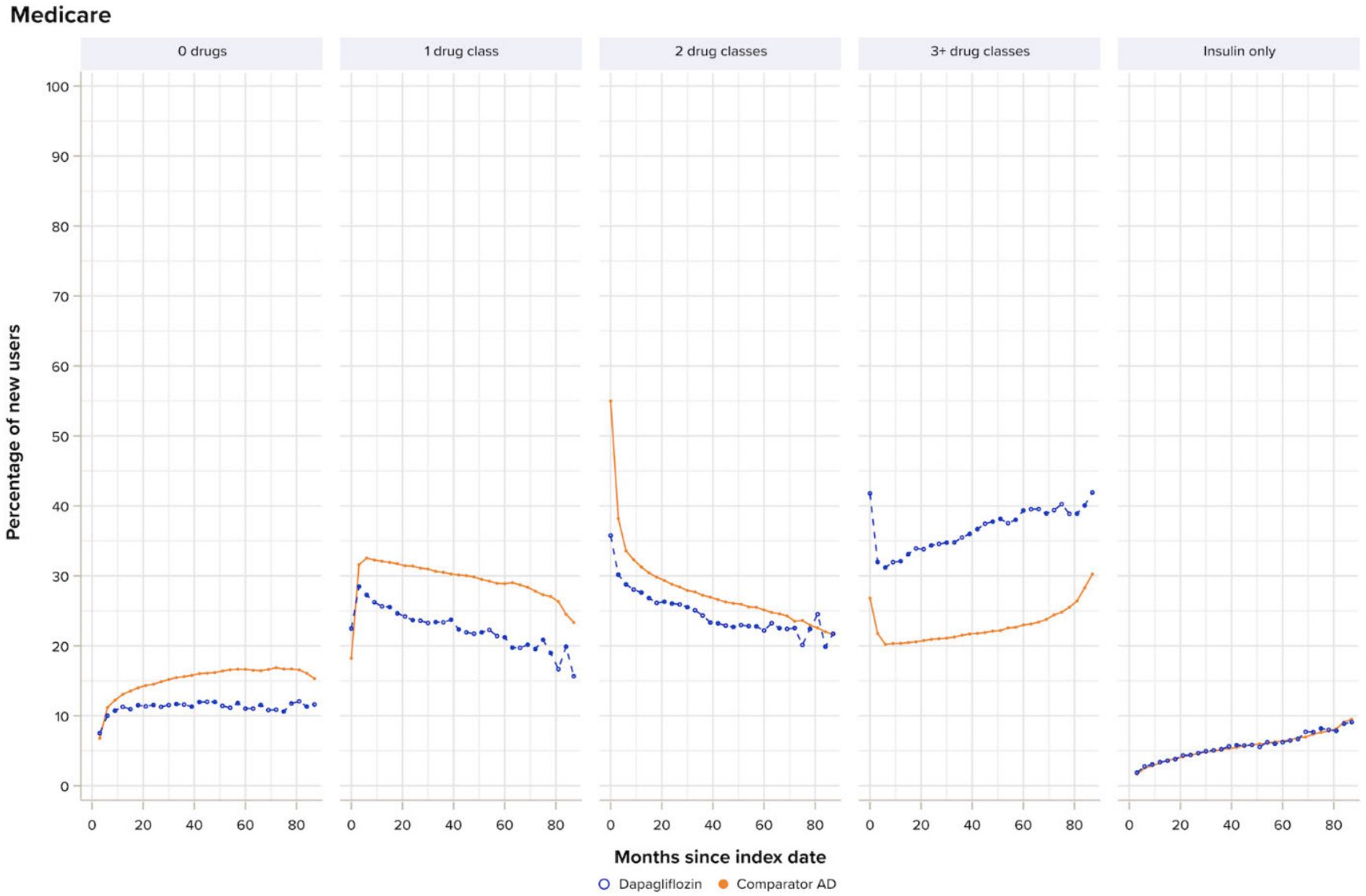
for individuals taking at least three AD classes during follow-up, with a higher proportion of dapagliflozin new users in this category (Figure 23).

Figure 23 **Antidiabetic Treatment Changes Over Time in New Users of Dapagliflozin and of Comparator ADs in the Female Breast Cancer Cohort, by Data Source. Balance of Covariates in the Female Breast Cancer Cohort, by Data Source**




PHARMO
0 drugs
1 drug class
2 drug classes
3+ drug classes
Insulin only
Percentage of new users
0
10
20
30
40
50
60
70
80
90
100
0
20
40
60
80
100
Months since index date
Dapagliflozin
Comparator AD


HIRD
0 drugs
1 drug class
2 drug classes
3+ drug classes
Insulin only
Percentage of new users
0
10
20
30
40
50
60
70
80
90
100
0 20 40 60 80 100
Months since index date
Dapagliflozin
Comparator AD


Medicare
0 drugs
1 drug class
2 drug classes
3+ drug classes
Insulin only
Percentage of new users
0
10
20
30
40
50
60
70
80
90
100
0
20
40
60
80
Months since index date
Dapagliflozin
Comparator AD

10.4.5.2.2 BLADDER CANCER

The results of the AD treatment changes during follow-up for the sex-combined bladder cancer cohorts were nearly identical to those for the female breast cancer cohorts across all data sources. In Appendix J, BladderCa Figure 4, shows these results for each data source.

10.5 Other Analyses

10.5.1 Potential Impact of Differential Outcome Misclassification

The results of the simulation analysis to assess the potential impact of differential outcome misclassification on the interpretation of the IRRs from the main analysis are presented in Table 23. For CPRD, the HIRD, and Medicare, the table displays the observed PPVs from the validation studies for female breast cancer and sex-combined bladder cancer, as reported in the fourth interim report, as well as the final 120-month analysis observed propensity score–adjusted IRR for each of these outcomes and the "corrected" propensity score–adjusted IRR point estimates as estimated under the hypothetical worst-case scenario of differential outcome misclassification (methods detailed in Section 9.9.8).

Table 23 Positive Predictive Values of Electronic Algorithms for Female Breast Cancer and Sex-Combined Bladder Cancer and Observed Propensity Score–Adjusted IRRs Compared With Potential IRRs Under the Worst-Case Scenario of Differential Outcome Misclassification (CPRD, the HIRD, and Medicare)

	PPV[a] % (95% CI)	Observed adjusted IRR[b] (95% CI)	Worst-case scenario "corrected" adjusted IRR point estimate[c]
Female breast cancer			
CPRD	99.0 (94.5-100.0)	1.05 (0.75-1.46)	1.06
HIRD	81.5 (74.5-88.5)	1.06 (0.90-1.24)	1.34
Medicare	84.6 (75.5-91.3)	0.98 (0.86-1.11)	1.17
Bladder cancer			
CPRD	96.6 (88.1-99.6)	0.74 (0.45-1.21)	0.77
HIRD	88.2 (80.9-95.4)	0.82 (0.55-1.24)	0.95
Medicare	97.8 (92.4-99.7)	0.74 (0.59-0.94)	0.76

[a] PPV: Observed PPVs represent the PPV 2 (post-review provisional cases [insufficient information to assign case status] were excluded from the numerator and denominator), as is defined in the interim 4 report. Each PPV was estimated using cases, identified by the respective case-finding algorithm, in the combined dapagliflozin and comparator AD exposure groups. Post-review provisional cases (insufficient information to assign case status) were excluded from the numerator and denominator. Detailed methods on outcome adjudication and estimation of the PPVs are presented in interim report 2 and interim report 4.

[b] Propensity score–adjusted IRR observed in the main analysis for each data source.

[c] Propensity score–adjusted IRR point estimate "corrected" under the hypothetical worst-case scenario of differential outcome misclassification [19].

PPV = positive predictive value.

10.5.1.1 Female Breast Cancer

In CPRD, for female breast cancer, an observed propensity score–adjusted IRR of 1.05 (95% CI, 0.75-1.46) and PPV of 99.0% were reported. To estimate the worst-case differential outcome misclassification scenario given the observed number of female breast cancer cases (51 observed cases in the dapagliflozin group and 210 observed cases in the comparator AD group), we assumed that if the dapagliflozin group had a PPV of 100% (all 51 cases confirmed to be true cases), then the comparator AD group would have a PPV of approximately 98.8% (207 of 210 confirmed to be true cases) in order for the observed PPV to be 99.0% (258 of 261 total cases confirmed), resulting in a hypothetical, worst-case corrected propensity score–adjusted IRR of 1.06.

In the HIRD, for female breast cancer, an observed propensity score–adjusted IRR of 1.06 (95% CI, 0.90-1.24) and PPV of 81.5% were reported. To estimate the worst-case differential outcome misclassification scenario given the observed number of female breast cancer cases (173 observed cases in the dapagliflozin group and 1,377 observed cases in the comparator AD group), we assumed that if the dapagliflozin group had a PPV of 100% (all 173 cases confirmed to be true cases), then the comparator AD group would have a PPV of approximately 79.2% (1,090 of 1,377 confirmed to be true cases) in order for the observed PPV to be 81.5% (1,263 of 1,550 total cases confirmed), resulting in a hypothetical, worst-case corrected propensity score–adjusted IRR of 1.34.

In Medicare, for female breast cancer, an observed propensity score–adjusted IRR of 0.98 (95% CI, 0.86-1.11) and PPV of 84.6% were reported. To estimate the worst-case differential outcome misclassification scenario given the observed number of female breast cancer cases (236 observed cases in the dapagliflozin group and 5,422 observed cases in the comparator AD group), we assumed that if the dapagliflozin group had a PPV of 100% (all 236 cases confirmed to be true cases), then the comparator AD group would have a PPV of approximately 83.9% (4,551 of 5,422 confirmed to be true cases) in order for the observed PPV to be 84.6% (4,787 of 5,658 total cases confirmed), resulting in a hypothetical, worst-case corrected propensity score–adjusted IRR of 1.17.

10.5.1.2 Sex-Combined Bladder Cancer

In CPRD, for sex-combined bladder cancer, an observed propensity score–adjusted IRR of 0.74 (95% CI, 0.45-1.21) and PPV of 96.6% were reported. To estimate the worst-case differential outcome misclassification scenario given the observed number of bladder cancer cases (19 observed cases in the dapagliflozin group and 106 observed cases in the comparator AD group), we assumed that if the dapagliflozin group had a PPV of 100% (all 19 cases confirmed to be true cases), then the comparator AD group would have a PPV of approximately 96.0% (102 of 106 confirmed to be true cases) in order for the observed PPV to be 96.6% (121 of 125 total cases confirmed), resulting in a hypothetical, worst-case corrected propensity score–adjusted IRR of 0.77.

In the HIRD, for sex-combined bladder cancer, an observed propensity score–adjusted IRR of 0.82 (95% CI, 0.55-1.24) and PPV of 88.2% were reported. To estimate the worst-case differential outcome misclassification scenario given the observed number of bladder cancer cases (27 observed cases in the dapagliflozin group and 212 observed cases in the comparator AD group), we assumed that if the dapagliflozin group had a PPV of 100% (all 27 cases confirmed to be true cases), then the comparator AD group would have a PPV of approximately 86.7% (184 of 212 confirmed to be true cases) in order for the observed PPV to be 88.2% (211 of 239 total cases confirmed), resulting in a hypothetical, worst-case corrected propensity score–adjusted IRR of 0.95.

In Medicare, for sex-combined bladder cancer, an observed propensity score–adjusted IRR of 0.74 (95% CI, 0.59-0.94) and PPV of 97.8% were reported. To estimate the worst-case differential outcome misclassification scenario given the observed number of bladder cancer cases (79 observed cases in the dapagliflozin group and 2,042 observed cases in the comparator AD group), we assumed that if the dapagliflozin group had a PPV of 100% (all 79 cases confirmed to be true cases), then the comparator AD group would have a PPV of approximately 97.7% (1,995 of 2,042 confirmed to be true cases) in order for the observed PPV to be 97.8% (2,074 of 2,121 total cases confirmed), resulting in a hypothetical, worst-case corrected propensity score–adjusted IRR of 0.76.

10.5.2 Sensitivity Analyses

10.5.2.1 Reasons for Censoring in the Primary Cancer Outcome Cohorts in the Main Analysis

The number of new users of dapagliflozin and comparator AD censored in the main analysis, by reason of censoring, is described for the female breast cancer cohorts in Table 24, the sex-combined bladder cancer cohorts in Table 25, and the female and male bladder cancer cohorts in Appendix J, BladderCa Table 14F and BladderCa Table 14M, respectively, for each data source.

In the overall propensity score–trimmed female breast cancer cohorts (Table 24), the most common reasons for censoring in the main analysis were end of the study period (ranging from 39.7% [CPRD] to 52.2% [Medicare]) and end of patient-specific data available in the data source (ranging from 19.6% [Medicare] to 33.2% [PHARMO]). After these, the most common reasons for censoring (that were common across all data sources), in order, were:

- Initiation of a non-dapagliflozin SGLT2 inhibitor (except for Medicare in which censoring due to death [9.4%] was similar to censoring due to initiating a non-dapagliflozin SGLT2 inhibitor [8.6%])
- Initiation of dapagliflozin (for the comparator AD group only)
- Death (for CPRD and PHARMO)

- Having a cancer diagnosis other than the cancer outcome of interest (in the HIRD, there were similar proportions of censoring due to death [0.5%] and censoring due to a cancer diagnosis other than the cancer outcome of interest [0.6%])

Of these, proportions were similar between new users of dapagliflozin and new users of comparator AD, with a few exceptions, including:

- In Medicare, the proportion censored due to having a cancer diagnosis other than the cancer outcome of interest was higher in new users of comparator AD (6.9%) than in users of dapagliflozin (4.6%).
- In CPRD, the proportion censored due to initiating a non-dapagliflozin SGLT2 inhibitor was higher in new users of comparator AD (13.5%) than in users of dapagliflozin (10.3%).
- In CPRD, PHARMO, and Medicare, the proportions censored due to death were higher in new users of comparator AD than in users of dapagliflozin (CPRD, 5.1% vs 2.3%; PHARMO, 9.0% vs 3.2%; Medicare, 9.8% vs 3.8%).

Similar patterns were observed in the female breast cancer cohorts stratified by insulin use at the index date.

In the sex-combined bladder cancer cohorts (Table 25), female bladder cancer cohorts (Appendix J, BladderCa Table 14F), and male bladder cancer cohorts (Appendix J, BladderCa Table 14M), similar patterns were observed as described for the overall female breast cancer cohorts.

The sensitivity analyses were conducted on the primary outcome cohorts, and, therefore, the analyses to describe the number of patients censored for each censoring criterion are presented only for the primary cancer outcomes. However, in the composite cancer outcome cohorts (secondary outcomes), analyses were conducted to describe the number of new users with follow-up censored due to initiation of a non-dapagliflozin SGLT2 inhibitor across the index exposure groups, and these results are presented for the overall cohort and insulin use–stratified cohorts in the propensity score–trimmed analysis samples for each data source in Appendix J, FemaleCompositeCa Table 14 and MaleCompositeCa Table 14. The proportions that were censored due to initiating a non-dapagliflozin SGLT2 inhibitor in each of the composite cancer cohorts were similar to the proportions censored for this reason in the female breast cancer and sex-combined bladder cancer cohorts described in Table 24 and Table 25.

Table 24 Number of Users Censored During Follow-up by Reason of Censoring in the Female Breast Cancer Cohorts, Overall and Stratified by Insulin Use at the Index Date, Propensity Score–Trimmed Analysis Samples, by Data Source

Reasons for censoring during follow-up,[a] n (%)	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Reached 65 years of age[b]	Insulin	NA	NA	NA	NA	270 (7.0)	1,633 (7.2)	NA	NA
	No insulin	NA	NA	NA	NA	1,381 (6.3)	12,728 (5.2)	NA	NA
	Overall	NA	NA	NA	NA	1,668 (6.4)	14,459 (5.3)	NA	NA
Cancer diagnosis other than cancer outcome of interest	Insulin	36 (3.3)	83 (3.3)	6 (3.5)	31 (2.4)	41 (1.1)	174 (0.8)	286 (4.6)	5,423 (6.9)
	No insulin	133 (2.9)	630 (2.9)	18 (1.6)	310 (2.4)	180 (0.8)	1,430 (0.6)	1,176 (4.6)	29,119 (6.8)
	Overall	169 (3.0)	704 (2.9)	28 (2.0)	340 (2.4)	215 (0.8)	1,621 (0.6)	1,459 (4.6)	34,557 (6.9)
Incomplete diagnosis information for cancer outcome of interest[c]	Insulin	NA	NA	NA	NA	NA	NA	20 (0.3)	242 (0.3)
	No insulin	NA	NA	NA	NA	NA	NA	46 (0.2)	1,151 (0.3)
	Overall	NA	NA	NA	NA	NA	NA	63 (0.2)	1,363 (0.3)
End of patient-specific data available in data source[d]	Insulin	374 (34.6)	770 (31.0)	53 (30.8)	445 (35.0)	1,530 (39.7)	9,420 (41.7)	1,124 (18.2)	16,703 (21.4)
	No insulin	1,400 (30.9)	6,494 (30.2)	364 (32.6)	4,209 (33.1)	7,425 (33.7)	79,096 (32.3)	3,637 (14.3)	83,103 (19.5)
	Overall	1,809 (31.7)	7,179 (30.1)	429 (31.0)	4,703 (33.4)	8,981 (34.6)	89,379 (32.8)	4,766 (15.1)	100,048 (19.9)
Initiated dapagliflozin (for comparator AD group only)	Insulin	NA	365 (14.7)	NA	51 (4.0)	NA	1,354 (6.0)	NA	2,347 (3.0)
	No insulin	NA	2,229 (10.4)	NA	582 (4.6)	NA	8,514 (3.5)	NA	10,788 (2.5)
	Overall	NA	2,581 (10.8)	NA	623 (4.4)	NA	9,862 (3.6)	NA	13,212 (2.6)
Initiated non-dapagliflozin SGLT2 inhibitor	Insulin	122 (11.3)	348 (14.0)	12 (7.0)	51 (4.0)	432 (11.2)	3,396 (15.0)	688 (11.1)	7,610 (9.7)
	No insulin	463 (10.2)	2,890 (13.4)	40 (3.6)	532 (4.2)	2,102 (9.5)	20,084 (8.2)	2,572 (10.1)	34,777 (8.2)
	Overall	589 (10.3)	3,222 (13.5)	56 (4.1)	571 (4.1)	2,561 (9.9)	23,566 (8.6)	3,283 (10.4)	42,655 (8.5)
Died	Insulin	41 (3.8)	176 (7.1)	7 (4.1)	159 (12.5)	42 (1.1)	288 (1.3)	342 (5.5)	9,917 (12.7)
	No insulin	83 (1.8)	1,034 (4.8)	32 (2.9)	1,033 (8.1)	122 (0.6)	923 (0.4)	884 (3.5)	40,099 (9.4)
	Overall	131 (2.3)	1,220 (5.1)	44 (3.2)	1,265 (9.0)	154 (0.6)	1,205 (0.4)	1,208 (3.8)	49,159 (9.8)

Reasons for censoring during follow-up,[a] n (%)	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
End of study period	Insulin	497 (45.9)	722 (29.1)	93 (54.1)	517 (40.7)	1,515 (39.3)	6,297 (27.9)	3,694 (59.7)	35,372 (45.3)
	No insulin	2,415 (53.3)	8,039 (37.4)	654 (58.6)	5,902 (46.4)	10,720 (48.7)	121,328 (49.5)	16,954 (66.7)	222,882 (52.4)
	Overall	2,964 (51.9)	8,774 (36.7)	815 (59.0)	6,416 (45.6)	12,291 (47.3)	131,937 (48.3)	20,704 (65.4)	258,507 (51.4)

[a] The reasons for censoring were are mutually exclusive. If a patient met the criteria for multiple censoring reasons on the same date, then they were counted towards each of the relevant censoring reasons.

[b] The HIRD is the only data source that censored follow-up time when the patient reached 65 years of age.

[c] In Medicare, a cancer outcome event was defined as having two diagnosis codes for the cancer outcome of interest that are ≤ 60 days apart. If a patient had a diagnosis code for the cancer outcome of interest without a second diagnosis code within 60 days, then, in the main analysis, the patient's follow-up was censored on the date of the single cancer diagnosis code.

[d] For CPRD, patient transferred out of a practice or the last date of data collection by the practice. For PHARMO, patient transferred out of pharmacy. For the HIRD and Medicare, patient's end of enrolment in health plan.

Note: The size of the overall sample may not equal the sum of the insulin use–stratified samples because propensity score estimation and trimming were performed separately in each sample.

NA = not applicable.

Table 25 Number of Users Censored During Follow-up by Reason of Censoring in the Sex-Combined Bladder Cancer Cohorts, Overall and Stratified by Insulin Use at the Index Date, Propensity Score–Trimmed Analysis Samples, by Data Source

Reasons for censoring during follow-up[a], n (%)	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
Reached 65 years of age[b]	Insulin	NA	NA	NA	NA	625 (7.1)	3,505 (7.4)	NA	NA
	No insulin	NA	NA	NA	NA	3,481 (6.4)	27,164 (5.6)	NA	NA
	Overall	NA	NA	NA	NA	4,123 (6.5)	30,682 (5.8)	NA	NA
Cancer diagnosis other than cancer outcome of interest	Insulin	89 (3.8)	227 (4.0)	17 (3.4)	81 (3.5)	114 (1.3)	567 (1.2)	780 (6.5)	11,843 (8.6)
	No insulin	414 (3.7)	1,940 (3.8)	67 (2.3)	1,034 (3.7)	611 (1.1)	4,621 (1.0)	2,959 (5.9)	64,605 (8.7)
	Overall	508 (3.6)	2,173 (3.8)	82 (2.5)	1,099 (3.6)	720 (1.1)	5,172 (1.0)	3,770 (6.0)	76,984 (8.7)
Incomplete diagnosis information for cancer outcome of interest[c]	Insulin	NA	NA	NA	NA	NA	NA	1-10	139 (0.1)
	No insulin	NA	NA	NA	NA	NA	NA	34 (0.1)	778 (0.1)
	Overall	NA	NA	NA	NA	NA	NA	39 (0.1)	928 (0.1)
End of patient-specific data available in data source[d]	Insulin	809 (34.5)	1,784 (31.7)	137 (27.5)	780 (33.2)	3,406 (38.7)	19,811 (41.6)	1,907 (15.9)	26,717 (19.5)
	No insulin	3,465 (30.8)	15,195 (29.8)	889 (30.6)	9,027 (32.3)	17,940 (33.1)	160,519 (33.4)	6,734 (13.4)	135,346 (18.2)
	Overall	4,397 (31.3)	16,931 (29.9)	1,022 (31.0)	9,921 (32.6)	21,476 (33.8)	180,288 (34.0)	8,711 (13.9)	163,222 (18.5)
Initiated dapagliflozin (for comparator AD group only)	Insulin	NA	780 (13.8)	NA	145 (6.2)	NA	2,946 (6.2)	NA	4,295 (3.1)
	No insulin	NA	5,318 (10.4)	NA	1,603 (5.7)	NA	19,409 (4.0)	NA	20,281 (2.7)
	Overall	NA	6,124 (10.8)	NA	1,729 (5.7)	NA	22,323 (4.2)	NA	24,687 (2.8)
Initiated non-dapagliflozin SGLT2 inhibitor	Insulin	280 (11.9)	822 (14.6)	28 (5.6)	125 (5.3)	997 (11.3)	7,634 (16.0)	1,372 (11.4)	14,801 (10.8)
	No insulin	1,061 (9.4)	7,213 (14.2)	107 (3.7)	1,444 (5.2)	5,227 (9.6)	47,531 (9.9)	5,294 (10.5)	67,516 (9.1)
	Overall	1,383 (9.8)	8,021 (14.2)	136 (4.1)	1,605 (5.3)	6,268 (9.9)	55,063 (10.4)	6,730 (10.7)	82,935 (9.4)
Died	Insulin	106 (4.5)	452 (8.0)	26 (5.2)	268 (11.4)	126 (1.4)	714 (1.5)	685 (5.7)	17,542 (12.8)
	No insulin	256 (2.3)	2,402 (4.7)	88 (3.0)	2,268 (8.1)	372 (0.7)	2,498 (0.5)	1,985 (3.9)	68,504 (9.2)
	Overall	369 (2.6)	2,814 (5.0)	121 (3.7)	2,558 (8.4)	507 (0.8)	3,196 (0.6)	2,681 (4.3)	85,978 (9.7)

Reasons for censoring during follow-up[a], n (%)	Insulin use at the index date	CPRD		PHARMO		HIRD		Medicare	
		Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD	Dapagliflozin	Comparator AD
End of study period	Insulin	1,055 (45.0)	1,558 (27.7)	289 (57.9)	940 (40.1)	3,568 (40.5)	12,597 (26.5)	7,279 (60.6)	62,082 (45.3)
	No insulin	6,055 (53.8)	18,808 (36.9)	1,748 (60.1)	12,452 (44.5)	26,709 (49.2)	220,181 (45.8)	33,352 (66.3)	386,783 (52.0)
	Overall	7,380 (52.5)	20,504 (36.2)	1,926 (58.4)	13,351 (43.9)	30,580 (48.1)	234,307 (44.2)	40,980 (65.2)	450,542 (50.9)

[a] The reasons for censoring were are mutually exclusive. If a patient met the criteria for multiple censoring reasons on the same date, then they were counted towards each of the relevant censoring reasons.

[b] The HIRD is the only data source that censored follow-up time when the patient reached 65 years of age.

[c] In Medicare, a cancer outcome event was defined as having two diagnosis codes for the cancer outcome of interest that are ≤ 60 days apart. If a patient had a diagnosis code for the cancer outcome of interest without a second diagnosis code within 60 days, then, in the main analysis, the patient's follow-up was censored on the date of the single cancer diagnosis code.

[d] For CPRD, patient transferred out of a practice or the last date of data collection by the practice. For PHARMO, patient transferred out of pharmacy. For the HIRD and Medicare, patient's end of enrolment in health plan.

Note: The size of the overall sample may not equal the sum of the insulin use–stratified samples, as propensity score estimation and trimming were performed separately in each sample.

NA = not applicable.

10.5.2.2 Sensitivity Analysis: Cumulative Incidence (Time-to-Event Analysis)

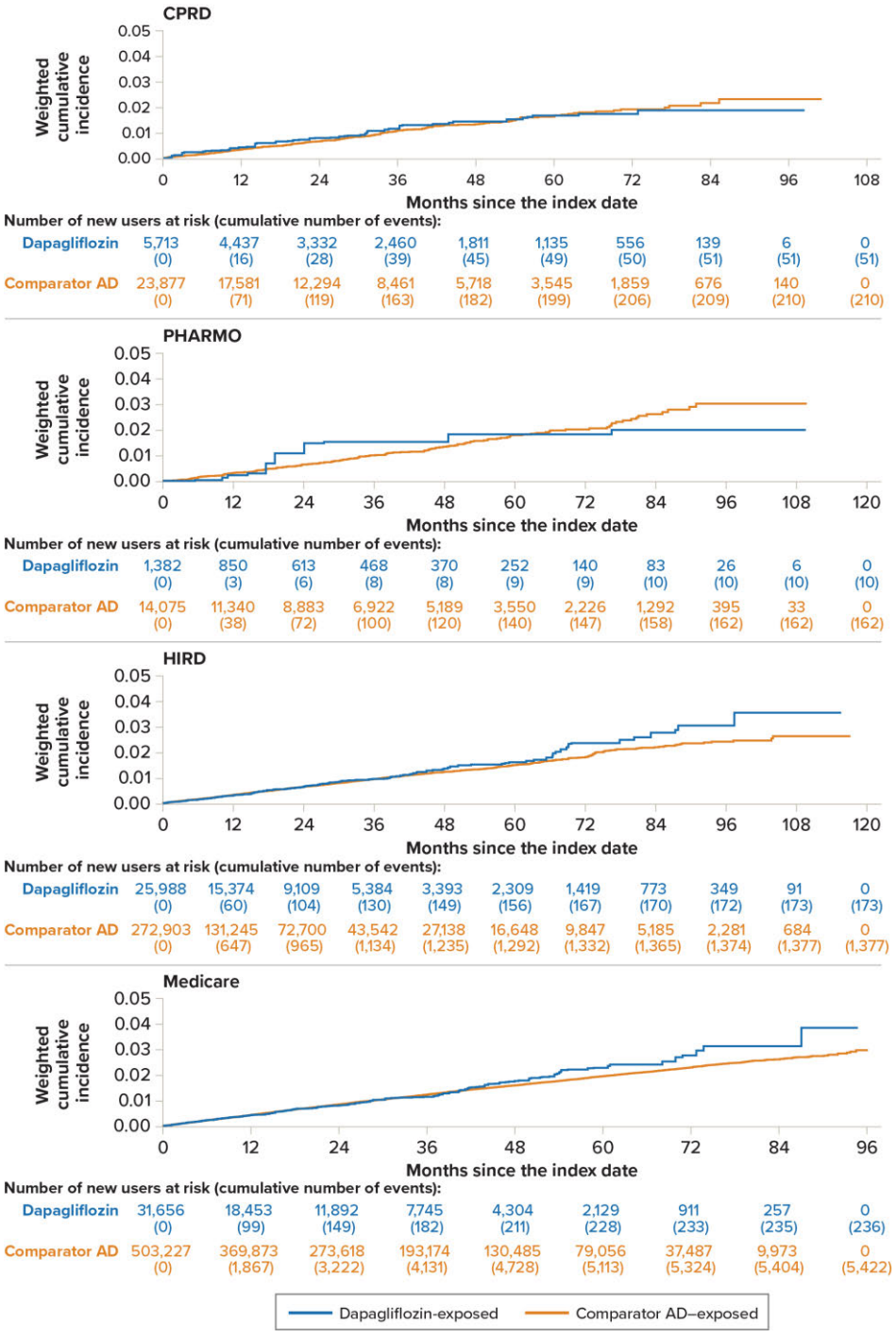
The weighted cumulative incidence of cancer over follow-up in the overall cohorts is presented for each data source for female breast cancer in Figure 24 and for bladder cancer (sex-combined and sex-specific) in Figure 25.

For female breast cancer, the weighted cumulative incidence of cancer events was < 3% throughout most or all of the follow-up period for CPRD, PHARMO, and the HIRD and was < 4% throughout most or all of the follow-up period for Medicare (Figure 24). The weighted cumulative incidence of breast cancer events was similar in new users of dapagliflozin and new users of comparator AD through at least 72 months of follow-up in CPRD and PHARMO, through at least 60 months of follow-up in the HIRD, and through at least 48 months of follow-up in Medicare, after which, in each of the respective data sources, a separation of the curves is observed, likely due to the weighted cumulative incidence estimates becoming less stable due to the lower numbers of new users at risk and new cancer events, particularly in the dapagliflozin groups. Given the similar weighted cumulative incidence rate in the dapagliflozin and comparator AD groups in the early years following initiation of therapy, the weighted cumulative incidence curves for all data sources suggest that there is no evidence of surveillance bias for female breast cancer.

For bladder cancer, the weighted cumulative incidence of cancer events was < 1% across all or most of the follow-up period for the sex-combined cohorts and the female and male cohorts for all data sources (Figure 25). The weighted cumulative incidence of bladder cancer events was similar for new users of dapagliflozin and new users of comparator AD throughout the follow-up period for all data sources (approximately 96 months of follow-up in CPRD and Medicare and approximately 108 months of follow-up in PHARMO and the HIRD), with lower stability in the cumulative incidence estimates in the dapagliflozin groups in the sex-specific cohorts due to few bladder cancer events occurring over the follow-up period (for females, the number of events ranged up to 13 total events [Medicare]; for males, the number of total events ranged from 10 [PHARMO] to 67 [Medicare]). The weighted cumulative incidence curves for all data sources appear to show no evidence of surveillance bias for bladder cancer in the sex-combined cohorts as well as the sex-specific cohorts, given similar cumulative incidence rates in both exposure groups following initiation of each therapy.

The weighted cumulative incidence of female breast cancer and bladder cancer (sex-combined and sex-specific) over the follow-up period was similar in the cohorts stratified by insulin use at the index date across all data sources (Appendix J, BreastCa Figure 5, BladderCa Figures 5, 5F, and 5M).

Figure 24 Sensitivity Analysis: Cumulative Incidence of Female Breast Cancer by Time Since Dapagliflozin Initiation or Comparator AD Initiation, Overall Cohort, by Data Source



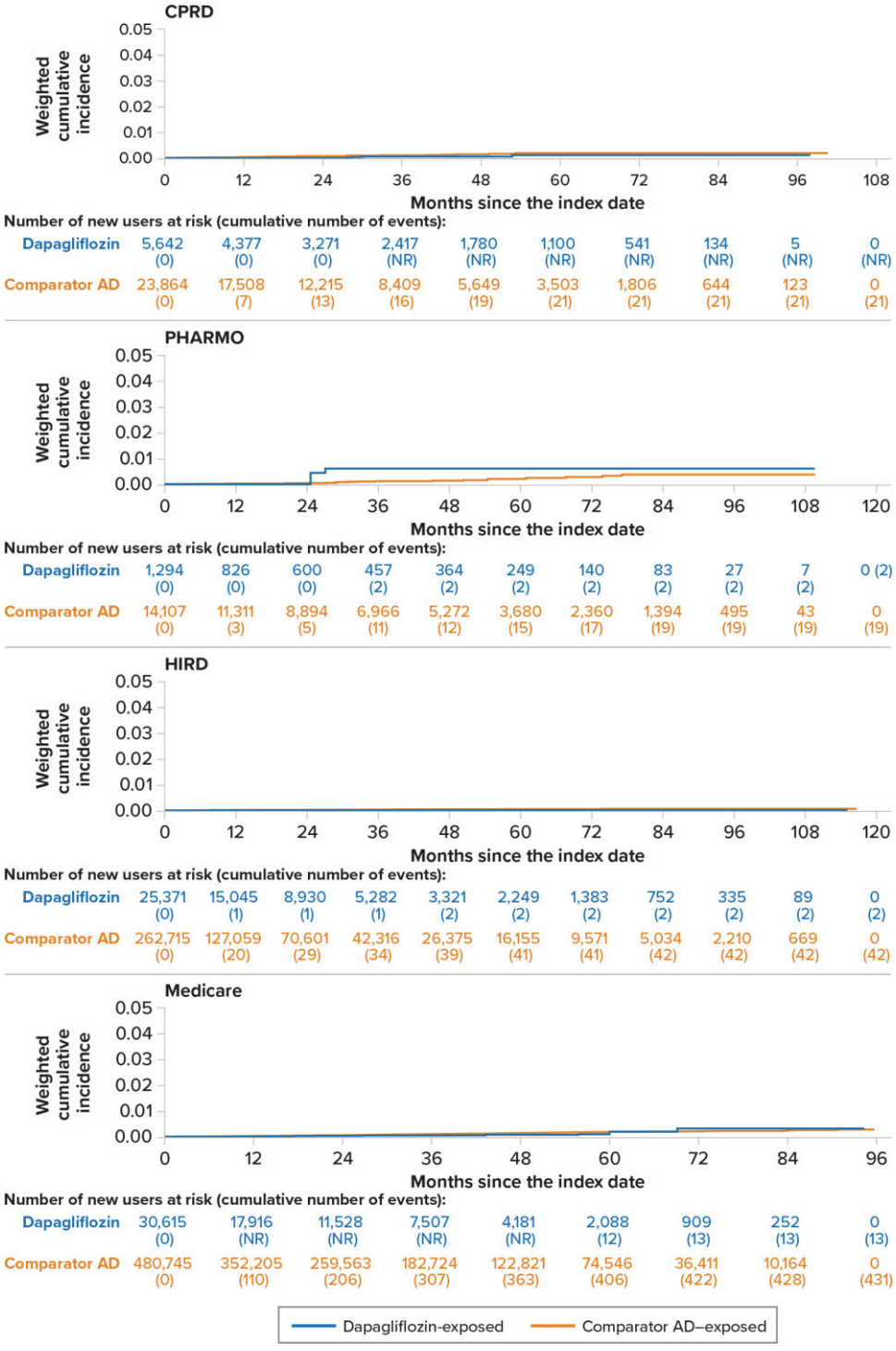
Note: Cancer risk was estimated using the KM estimator, which included stabilised IPTW (generated from the propensity score) for confounding adjustment. Weighted cumulative incidence was calculated using 1 − KM estimates. Methods are described in Section 9.9.5.1.

IPTW = inverse probability of treatment weights.

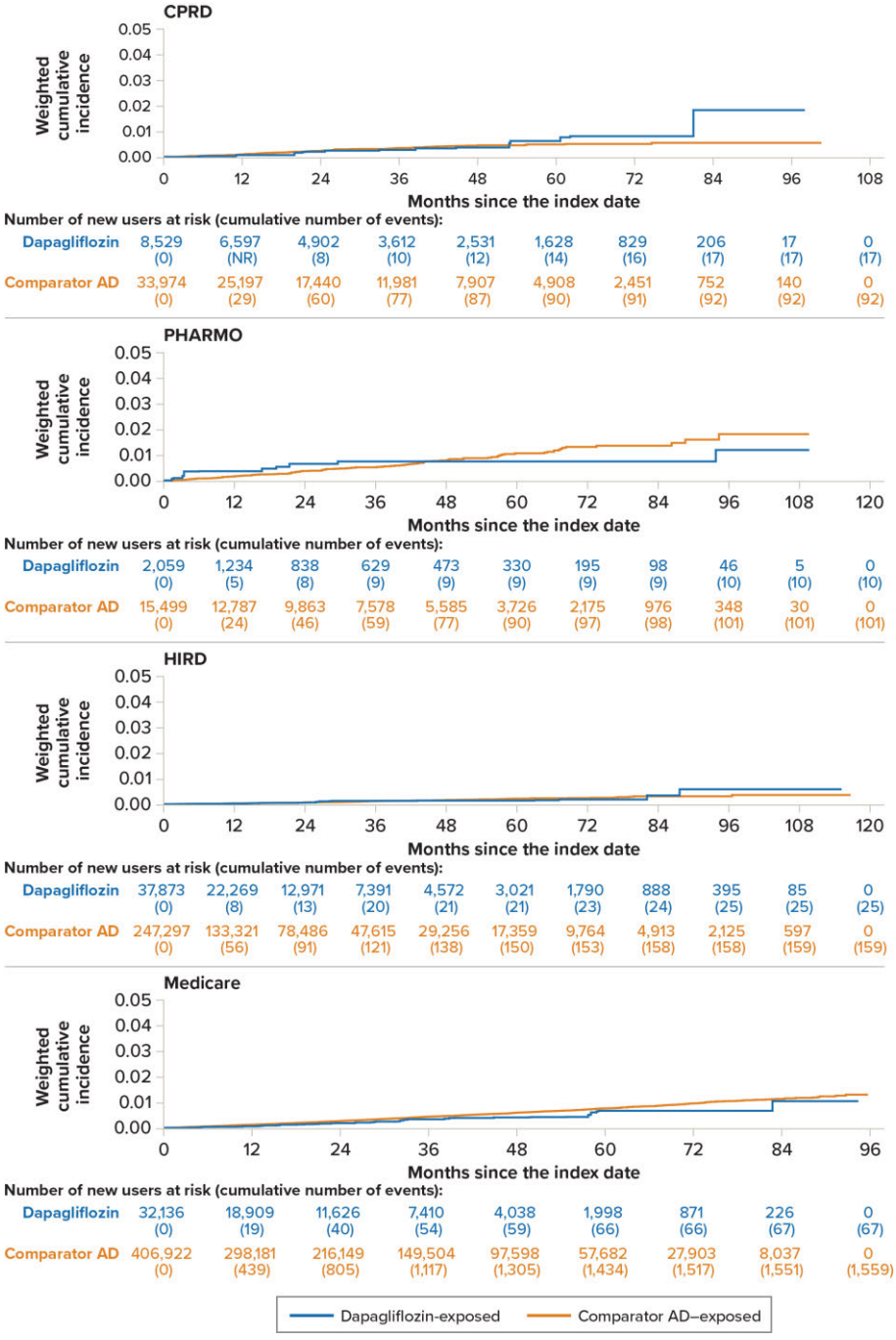
Figure 25 Sensitivity Analysis: Cumulative Incidence of Bladder Cancer by Time Since Dapagliflozin Initiation or Comparator AD Initiation, Overall Cohort, by Data Source

Sex-Combined Cohort

Female Cohort



Male Cohort



Note: For CPRD data, any cell with a value of 1-4, or any cell that allows a value of 1-4 to be derived from other reported cells or information, cannot be reported. For Medicare data, any cell with a value of 1-10, or any cell that allows a value of 1-10 to be derived from other reported cells or information, cannot be reported. To comply with these reporting requirements, values that could be used to derive small count sizes in other cells are masked as NR.

Note: Cancer risk was estimated using the KM estimator, which included stabilised IPTW (generated from the propensity score) for confounding adjustment. Weighted cumulative incidence was calculated using 1 − KM estimates. Methods are described in Section 9.9.5.1.

IPTW = inverse probability of treatment weights; NR = not reportable.

10.5.2.3 Sensitivity Analysis: Potential Impact of Unmeasured Confounding

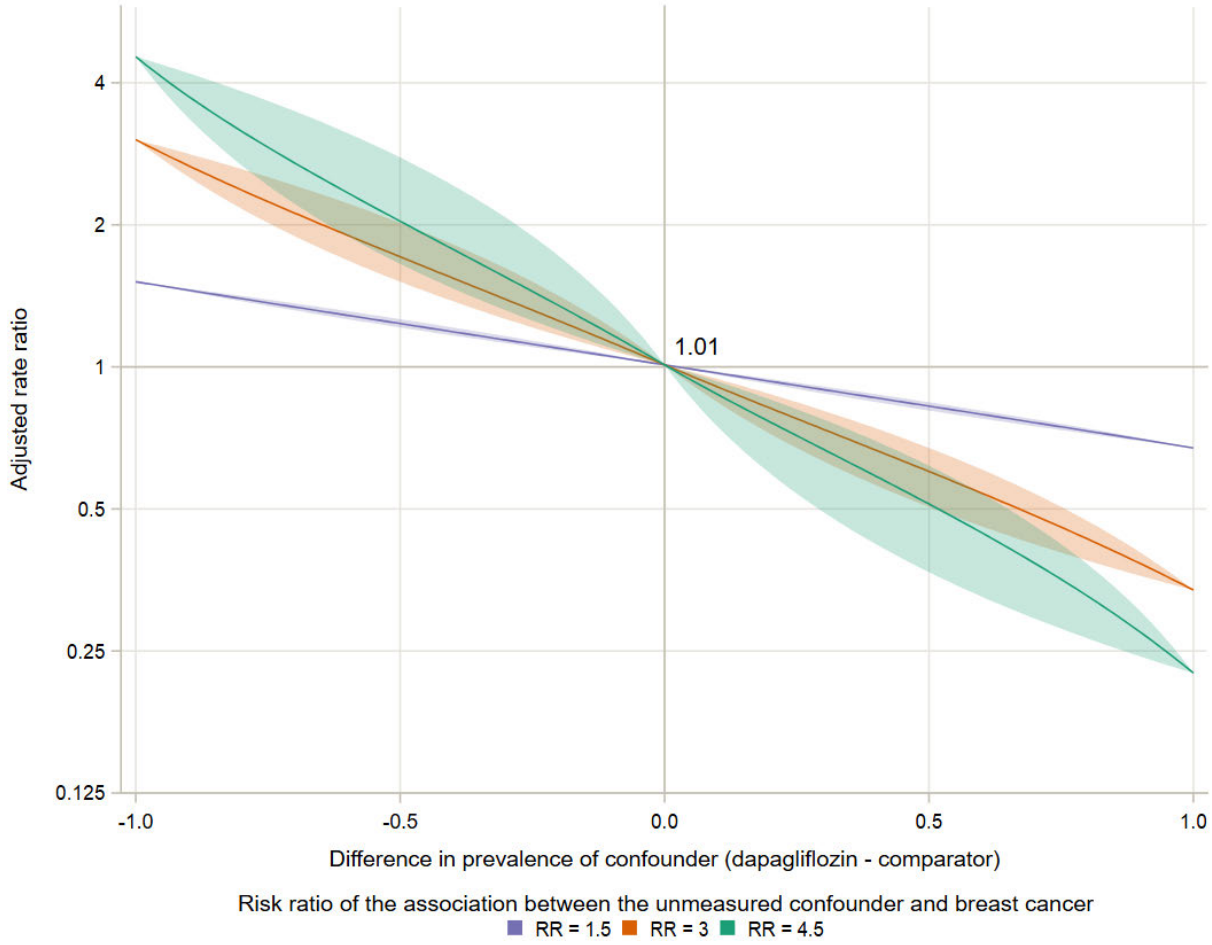
This sensitivity analysis evaluated the potential impact of an unmeasured confounder on the estimated pooled adjusted IRRs for the primary cancer outcomes, and the results are displayed in Figure 26 (female breast cancer) and Figure 27 (sex-combined and sex-specific bladder cancer). Because no specific unmeasured confounder was being evaluated, three scenarios of the association of hypothetical unmeasured confounders with each cancer outcome of interest were evaluated: RR = 1.5 (moderate association, in blue), RR = 3.0 (strong association, in red), and RR = 4.5 (very strong association, in green). An underlying assumption is that these confounder relationships were independent and not associated with measured confounders. For a factor to have been a confounder of the relationship between dapagliflozin and the cancer outcome of interest, it must have been imbalanced between the two index treatment groups. For each hypothetical confounder scenario, we calculated IRRs corrected for the unmeasured confounder with imbalances ranging from −100% (present in every comparator AD patient and not present in any dapagliflozin patient) to 100% (present in every dapagliflozin patient and not present in any comparator AD patient). In the figure for each cancer outcome of interest, the IRR values at 0% imbalance in each scenario are the observed pooled adjusted IRR estimated in the overall cohort, as it assumed no imbalance of the unmeasured confounder and thus no confounding. For each level of covariate imbalance, an adjusted IRR corrected for the hypothetical unmeasured confounder was calculated assuming the strength of the confounder in that scenario and the confounder prevalences in the two treatment groups.

A given covariate imbalance may be achieved through numerous combinations of confounder levels (eg, an imbalance of 10% can result from confounder prevalences of 10% in dapagliflozin patients and 0% in comparator AD patients, or 90% in dapagliflozin patients and 80% in comparator AD patients). The coloured bands for each confounding scenario represent the minimum and maximum possible corrected adjusted IRRs at each level of imbalance, and the solid line represents the mean corrected adjusted IRR at each imbalance level.

As illustrated in Figure 26 for female breast cancer, in the worst-case scenario of having a hypothetical moderate confounder (RR = 1.5) in which groups would be completely imbalanced (0% prevalence in the dapagliflozin group and 100% prevalence in the comparator AD group), the maximum true IRR would be 1.52; any imbalance less extreme would result in IRRs lower than 1.52. A hypothetical moderate confounder (RR = 1.5, blue line in Figure 26) would require an imbalance with a higher prevalence amongst comparator AD patients to mask a true adjusted IRR greater than 1.00; conversely, a hypothetical moderate confounder with an imbalance with a higher prevalence amongst dapagliflozin patients would yield a corrected adjusted IRR below 1.00. If the hypothetical unmeasured confounder had a stronger independent relationship with the outcome, RR = 3.0 or 4.5, then smaller imbalances (with a higher prevalence in the comparator AD group) would be required to mask stronger positive associations between dapagliflozin and female breast cancer. It is not anticipated that a

common, moderate confounder would be unmeasured, imbalanced with a higher prevalence amongst comparator AD users, and uncorrelated with measured covariates (that were included in the analysis) to mask a true harmful association of female breast cancer with dapagliflozin.

Figure 26 **Sensitivity Analysis: Adjusted Incidence Rate Ratios for Female Breast Cancer Under Varying Assumptions of Unmeasured Confounding Compared With the Observed Adjusted Incidence Rate Ratio Estimate From the Pooled Analysis, Without Stratification by Insulin Use at the Index Date**



Note 1: The estimated pooled adjusted IRR for female breast cancer in the overall cohort was 1.01 (95% CI, 0.91-1.11).

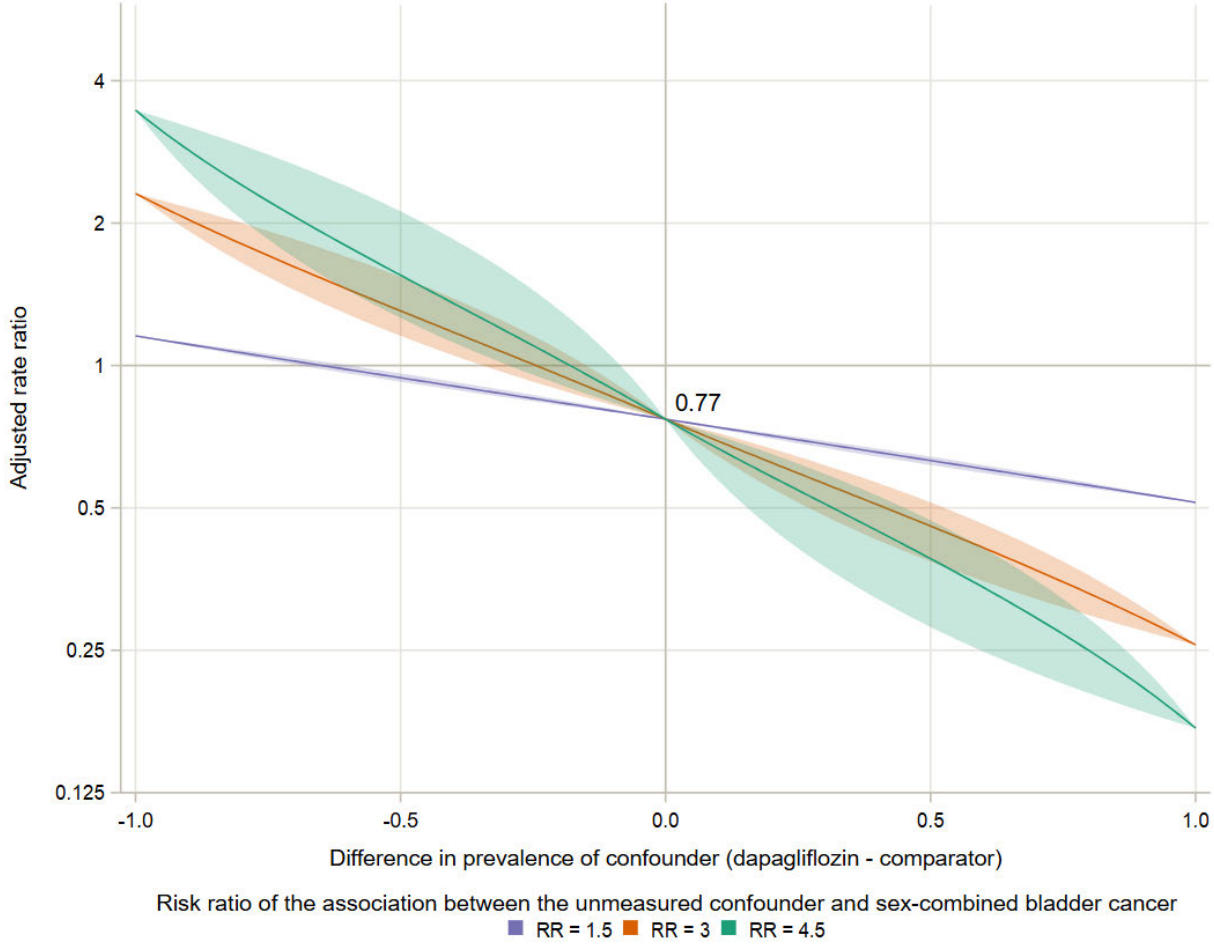
Note 2: The "Adjusted rate ratio" refers to the corrected IRR estimate under varying scenarios of a potential unmeasured confounder. The "risk ratio" (or RR) refers to the association between a hypothetical unmeasured confounder and the cancer outcome.

As illustrated in Figure 27 for sex-combined bladder cancer, in the worst-case scenario of having a hypothetical moderate confounder (RR = 1.5) in which groups would be completely imbalanced (0% prevalence in the dapagliflozin group and 100% prevalence in the comparator AD group), the maximum true IRR would be 1.16; any imbalance less extreme would result in

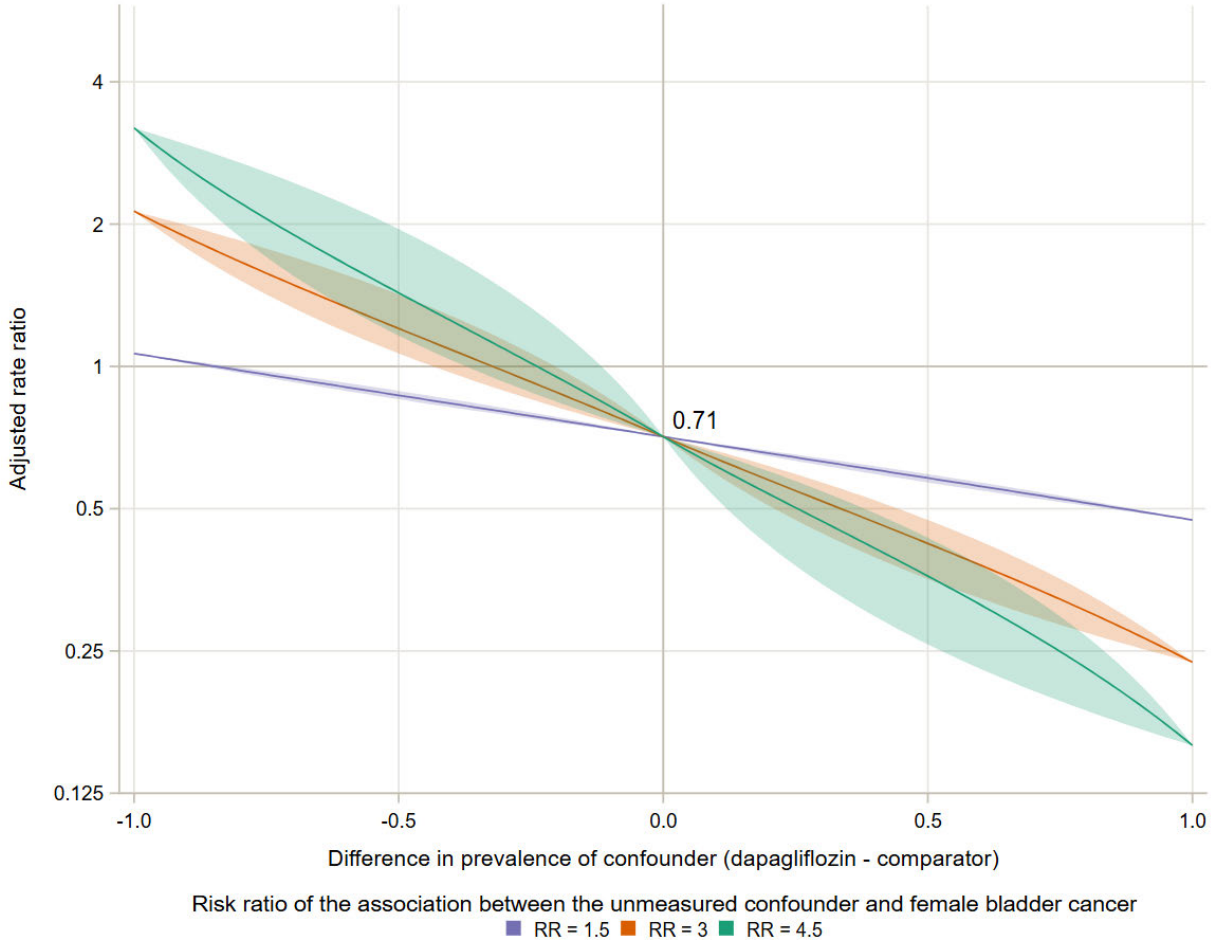
IRRs lower than 1.16. A hypothetical moderate confounder (RR = 1.5) would require an imbalance of at least approximately −60% (ie, higher prevalence in the comparator AD patients) to mask a true IRR greater than 1.00. If the hypothetical unmeasured confounder had a stronger independent relationship with the outcome, RR = 3.0 or 4.5, then a smaller imbalance would be required. When assessing the potential impact of unmeasured confounding for assessing bladder cancer amongst females and males, similar patterns were observed as for the bladder cancer in the sex-combined cohort. It is not anticipated that a common, moderate confounder would be unmeasured, imbalanced enough with a higher prevalence amongst comparator AD users, and uncorrelated with measured covariates (that were included in the analysis) to mask a true harmful association of bladder cancer with dapagliflozin.

Figure 27 **Sensitivity Analysis: Adjusted Incidence Rate Ratios for Bladder Cancer Under Varying Assumptions of Unmeasured Confounding Compared With the Observed Adjusted Incidence Rate Ratio Estimate From the Pooled Analysis, Without Stratification by Insulin Use at the Index Date**

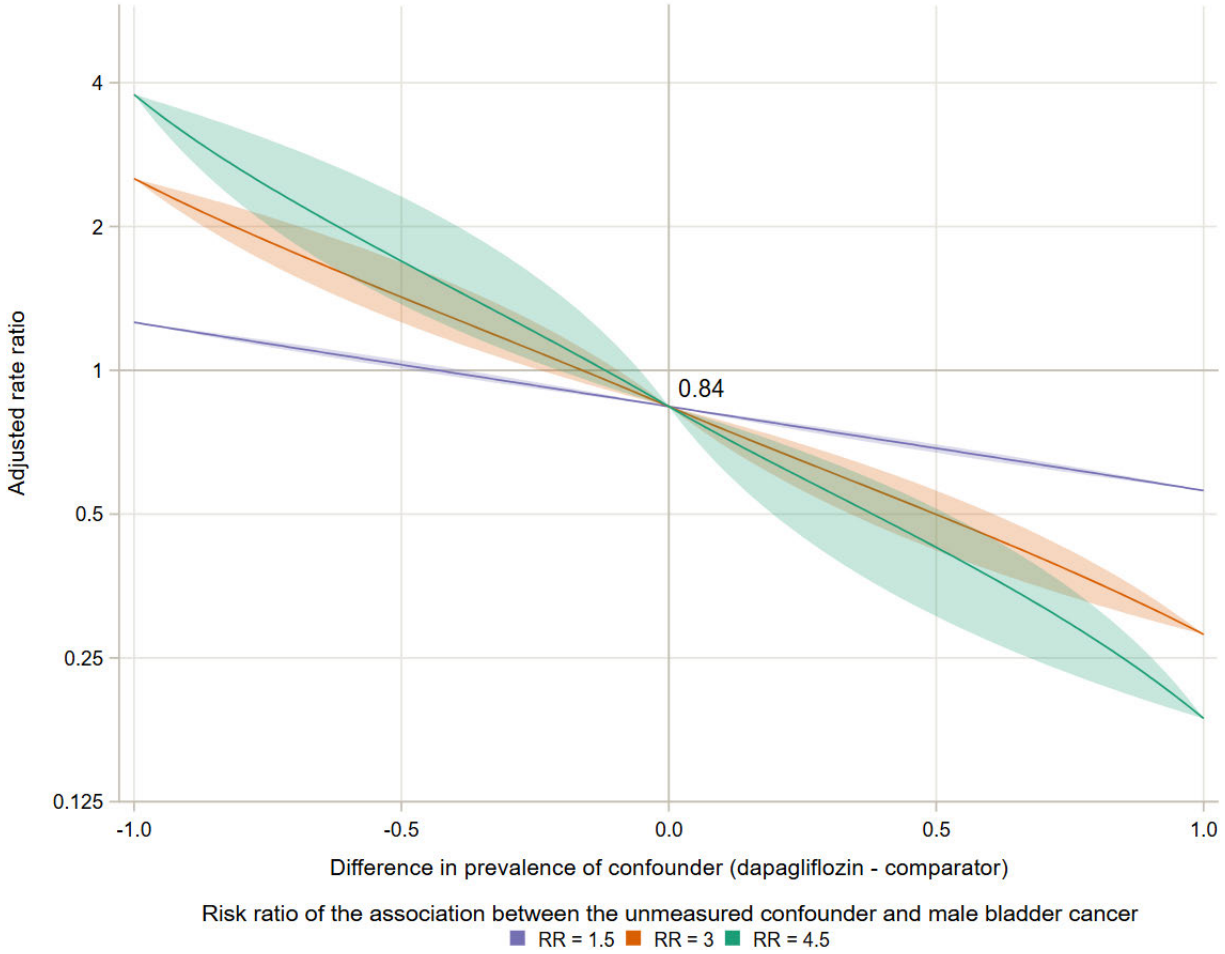
Sex-Combined Cohort



Female Cohort



Male Cohort



Note 1: The estimated pooled adjusted IRR in the overall cohort for sex-combined bladder cancer was 0.77 (95% CI, 0.64-0.92), for female bladder cancer was 0.71 (95% CI, 0.42-1.18), and for male bladder cancer was 0.84 (95% CI, 0.69-1.02).

Note 2: The "Adjusted rate ratio" refers to the corrected IRR estimate under varying scenarios of a potential unmeasured confounder. The "risk ratio" (or RR) refers to the association between a hypothetical unmeasured confounder and the cancer outcome.

10.5.2.4 Sensitivity Analysis: Removal of Various Censoring Criteria

Propensity score–adjusted incidence rates and IRRs for the primary cancer outcomes (female breast cancer, sex-combined bladder cancer, female bladder cancer, and male bladder cancer) after removing each of two censoring criteria ([1] censoring at the first occurrence of any cancer other than the cancer outcome of interest [ie, labelled as 'first occurrence of each type of malignancy'] and [2] censoring at the initiation of a non-dapagliflozin SGLT2 inhibitor) are presented for the overall cohorts and the cohorts stratified by insulin use at the index in Appendix J, BreastCa Table 25 and BladderCa Tables 25, 25F, and 25M for each data source. The propensity score–adjusted IRRs from the main analysis and the sensitivity analyses to remove the various censoring criteria are also displayed in forest plots in Figure 28 (female breast cancer) and Figure 29 (bladder cancer, sex-combined and sex-specific). These

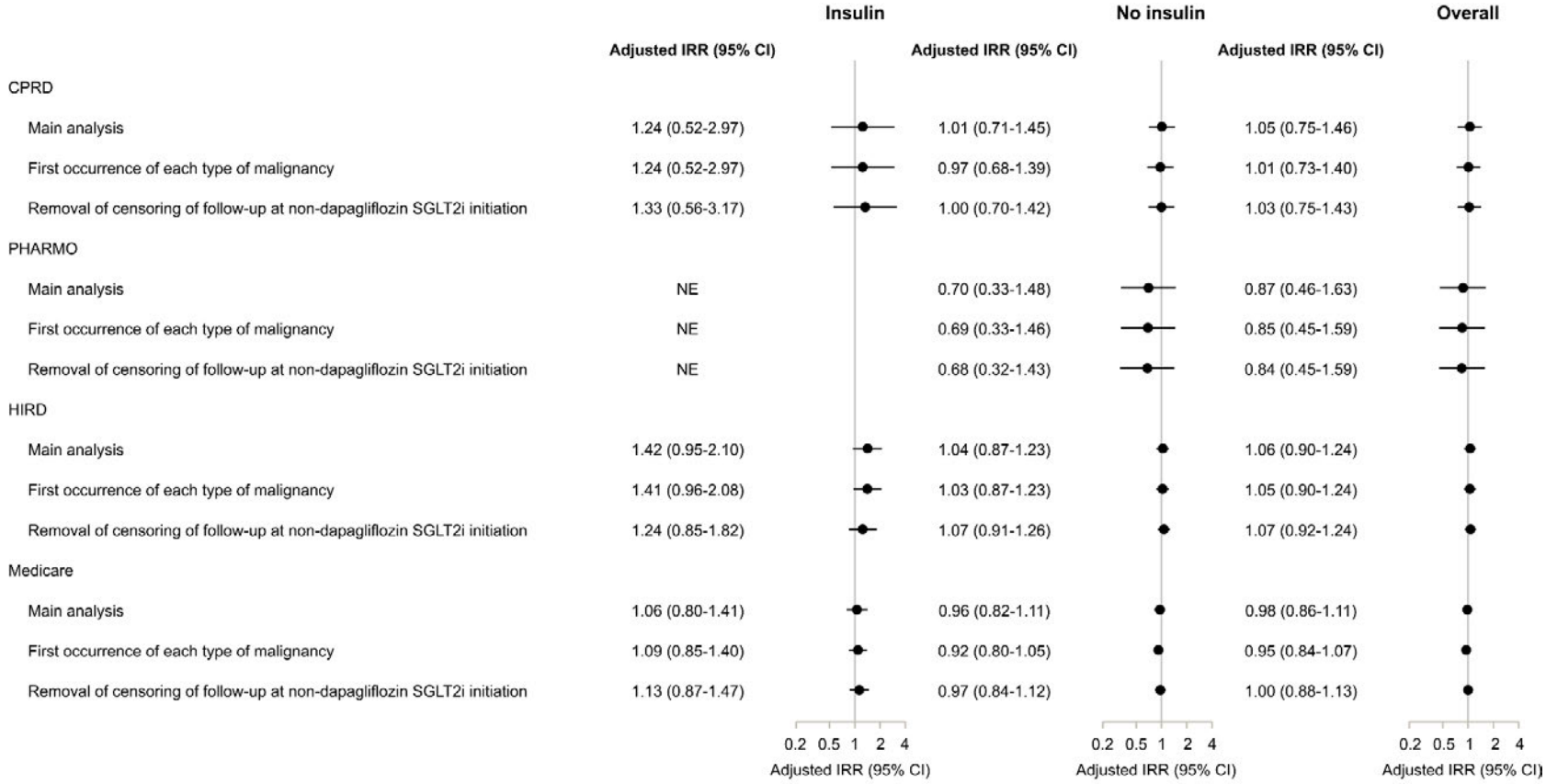
sensitivity analyses were only conducted in cohorts for which a propensity score–adjusted IRR was estimable in the main analysis. Given that the propensity score–adjusted IRRs for the primary cancer outcomes in each of the data sources in the main analysis did not suggest an increased risk of cancer for new users of dapagliflozin, the condition for conducting the sensitivity analysis to assess IRRs after the removal of censoring at the time of dapagliflozin initiation for new users of comparator AD was not conducted in any cohort (details described in Section 9.9.5.3.3).

For female breast cancer, in the overall cohorts and the cohorts stratified by insulin use at the index date, the propensity score–adjusted incidence rates in the dapagliflozin groups and comparator AD groups were similar after the removal of each of the two censoring criteria compared with the propensity score–adjusted incidence rates from the main analysis (Appendix J, BreastCa Table 25 for each data source), with a few exceptions: (1) in the HIRD, amongst insulin users, where the propensity score–adjusted incidence rate in the dapagliflozin group after the removal of censoring at the initiation of a non-dapagliflozin SGLT2 inhibitor was lower (41.1 per 10,000 person-years; 95% CI, 28.5-57.4) than that from the main analysis (45.6 per 10,000 person-years; 95% CI, 31.4-64.0), and (2) in Medicare, in the overall cohort and the cohorts stratified by insulin use at the index date, the propensity score–adjusted incidence rates were higher in both the dapagliflozin and comparator AD groups after the removal of censoring for cancers other than the cancer of interest compared with the adjusted incidence rates from the main analysis. The propensity score–adjusted IRRs for female breast cancer after the removal of each of the two censoring criteria were similar to the propensity score–adjusted IRRs from the main analysis (Figure 28). All propensity score–adjusted IRRs from the sensitivity analyses varied by < 0.1 from the propensity score–adjusted IRR main analysis, except for the HIRD. Amongst insulin users in the HIRD, the propensity score–adjusted IRR was closer to 1.0 (1.24; 95% CI, 0.85-1.82) after the removal of censoring at the initiation of a non-dapagliflozin SGLT2 inhibitor compared with the propensity score–adjusted IRR from the main analysis (1.42; 95% CI, 0.95-2.10).

For bladder cancer (sex-combined and sex-specific), in the overall cohorts and the cohorts stratified by insulin use at the index date, the propensity score–adjusted incidence rates in the dapagliflozin groups and comparator AD groups were similar after the removal of each of the two censoring criteria compared with the propensity score–adjusted incidence rates from the main analysis (Appendix J, BladderCa Table 25, BladderCa Table 25F, and BladderCa Table 25M, for each data source), with the exception of Medicare. In Medicare, in the overall cohort and the cohorts stratified by insulin use at the index date, the propensity score–adjusted incidence rates for both sex-combined bladder cancer and bladder cancer in males were higher in both the dapagliflozin and comparator AD groups after the removal of censoring for cancers other than the cancer of interest compared with the propensity score–adjusted incidence rates from the main analysis. All propensity score–adjusted IRRs from the sensitivity analyses in the sex-combined bladder cancer cohorts and the sex-specific bladder

cancer cohorts varied by ≤ 0.1 from the propensity score–adjusted IRR in the main analysis, with the exception of CPRD insulin users (Figure 29). Amongst insulin users in CPRD, the removal of censoring at the initiation of a non-dapagliflozin SGLT2 inhibitor yielded propensity score–adjusted IRRs of 1.30 (95% CI, 0.45-3.71) for sex-combined bladder cancer (vs 1.01 [95% CI, 0.33-3.09] in the main analysis), 2.16 (95% CI, 0.18-26.16) for bladder cancer in females (vs 2.00 [95% CI, 0.16-25.73] in the main analysis), and 1.17 (95% CI, 0.39-3.48) for bladder cancer in males (vs 0.85 [95% CI, 0.27-2.75] in the main analysis); however, all of these estimates are imprecise with wide CIs due to small sample sizes and few bladder cancer events in the insulin use cohorts. For PHARMO and the HIRD, propensity score–adjusted IRRs for bladder cancer in females were not estimable in the main analysis for the overall cohorts and the cohorts stratified by insulin use at the index date, and, therefore, the sensitivity analyses to remove censoring criteria were not conducted in the female bladder cancer cohorts in these data sources.

Figure 28 **Sensitivity Analysis: Propensity Score–Adjusted Incidence Rate Ratios for Female Breast Cancer When Removing Various Censoring Criteria, Overall and Stratified by Insulin Use at the Index Date, by Data Source**



Note: Adjusted IRR estimates below 1.0 (ie, on the left side of 1 in the plots) favour dapagliflozin, and estimates above 1.0 (ie, on the right side of 1 in the plots) favour comparator AD.

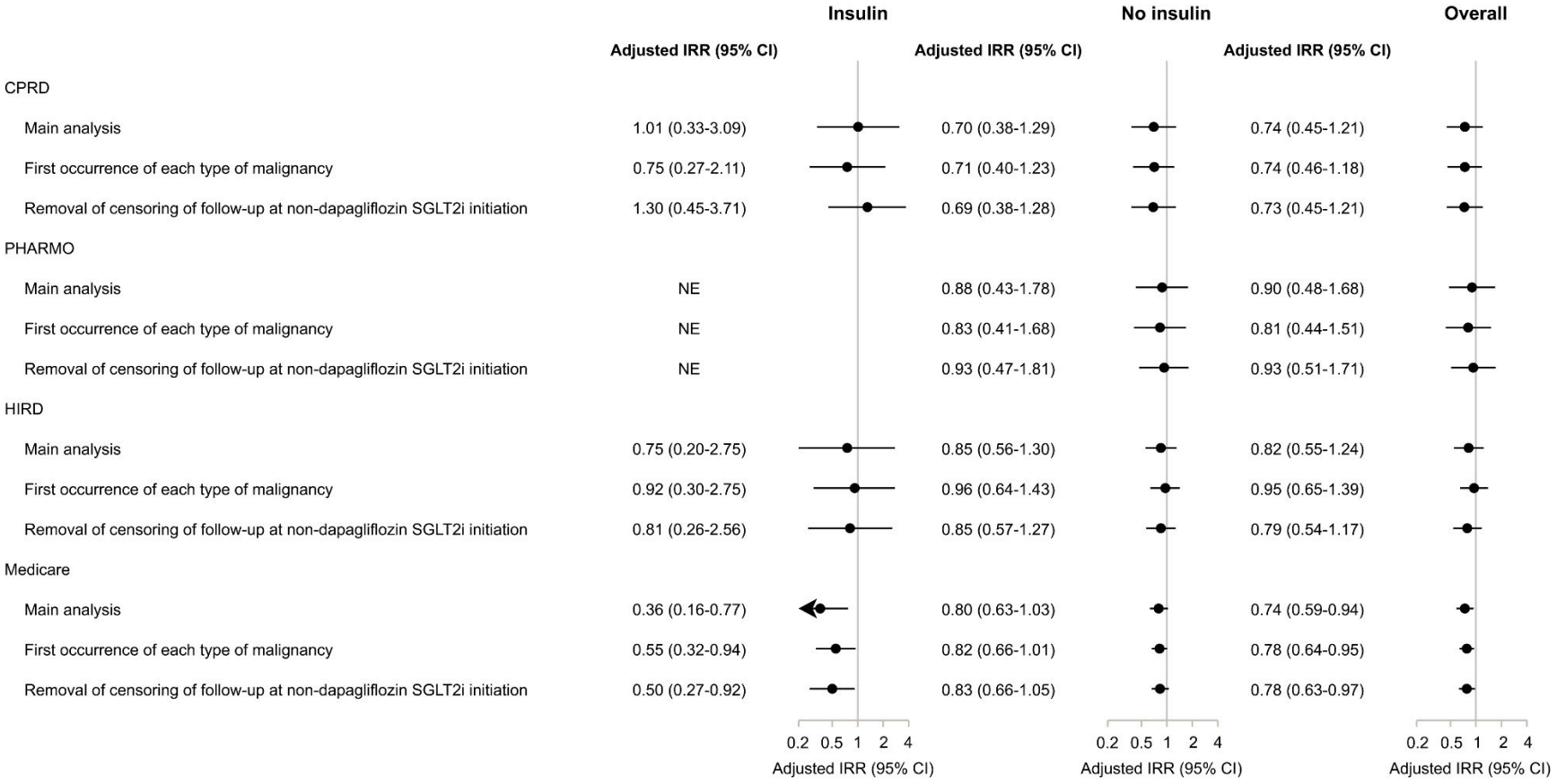
Note 1: The sensitivity analysis assessing the first occurrence of each type of malignancy estimated the propensity score–adjusted IRRs after the removal of censoring at the occurrence of a cancer other than the cancer of interest.

Note 2: The sensitivity analysis to assess the impact of censoring at the initiation of dapagliflozin amongst new users of comparator AD was not conducted in any cohort, across all data sources, as there was no increased risk of any of the primary cancer outcomes observed amongst dapagliflozin users in the main analysis (see Section 9.9.5.3.3 for details).

NE = not estimable.

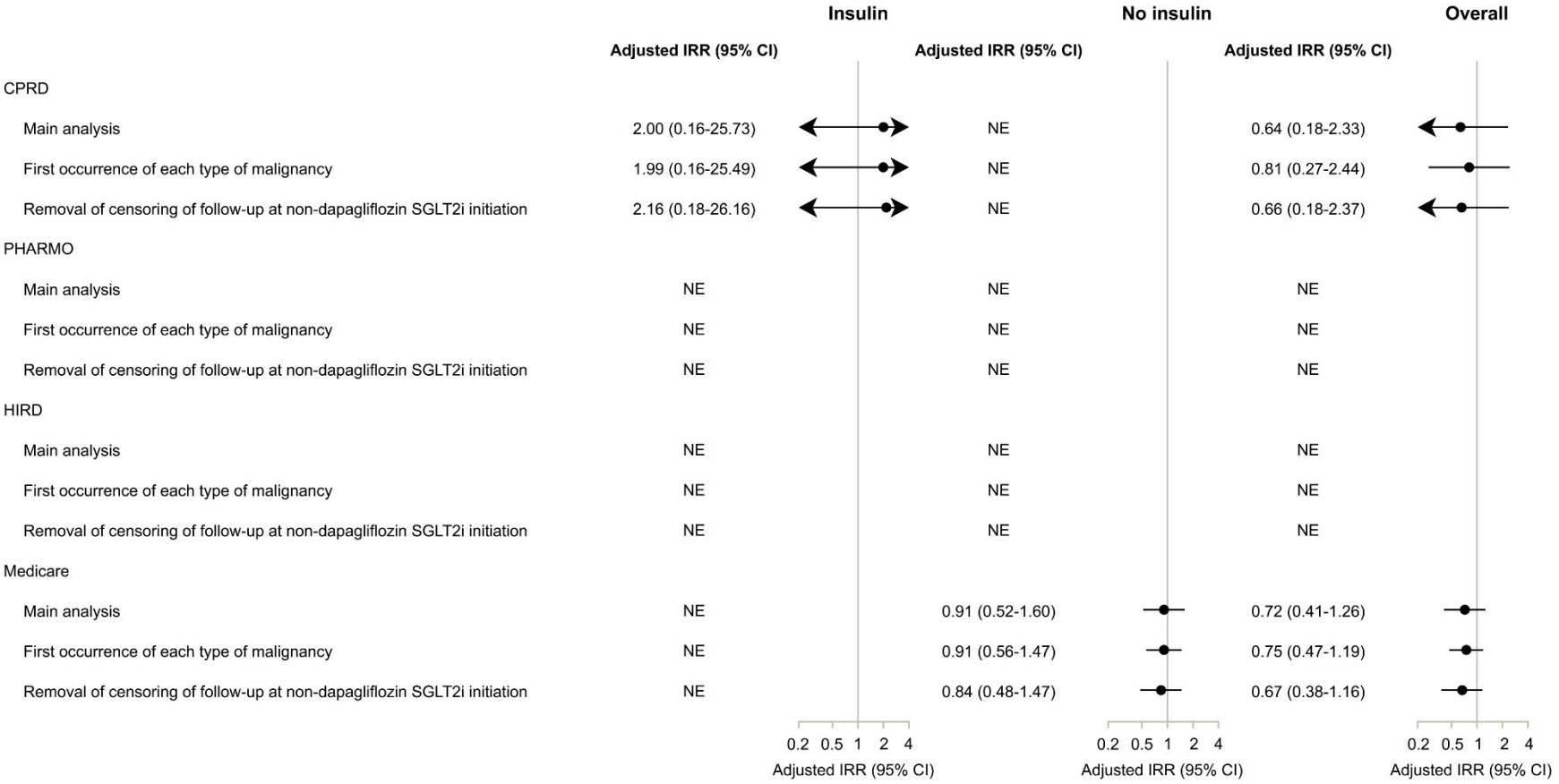
Figure 29 Sensitivity Analysis: Propensity Score–Adjusted Incidence Rate Ratios for Bladder Cancer When Removing Various Censoring Criteria, Overall and Stratified by Insulin Use at the Index Date, by Data Source

Sex-Combined Cohort



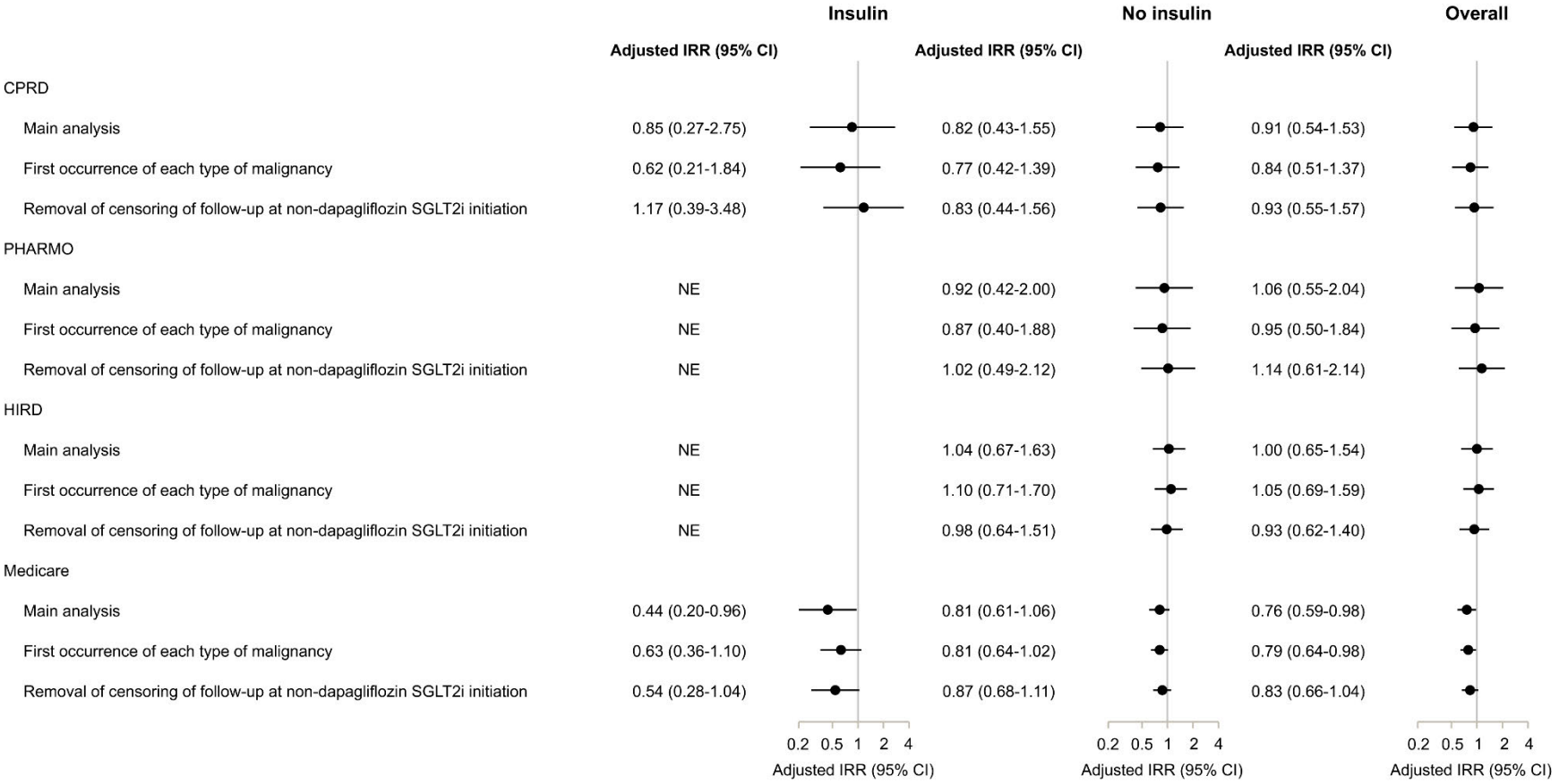
Note: Adjusted IRR estimates below 1.0 (ie, on the left side of 1 in the plots) favour dapagliflozin, and estimates above 1.0 (ie, on the right side of 1 in the plots) favour comparator AD.

Female Cohort



Note: Adjusted IRR estimates below 1.0 (ie, on the left side of 1 in the plots) favour dapagliflozin and estimates above 1.0 (ie, on the right side of 1 in the plots) favour comparator AD.

Male Cohort



Note: Adjusted IRR estimates below 1.0 (ie, on the left side of 1 in the plots) favour dapagliflozin, and estimates above 1.0 (ie, on the right side of 1 in the plots) favour comparator AD.

Note 1: The sensitivity analysis assessing the first occurrence of each type of malignancy estimated the propensity score–adjusted IRRs after the removal of censoring at the occurrence of a cancer other than the cancer of interest.

Note 2: The sensitivity analysis to assess the impact of censoring at the initiation of dapagliflozin amongst new users of comparator AD was not conducted in any cohort, across all data sources, as there was no increased risk of any of the primary cancer outcomes observed amongst dapagliflozin users in the main analysis (see Section 9.9.5.3.3 for details).

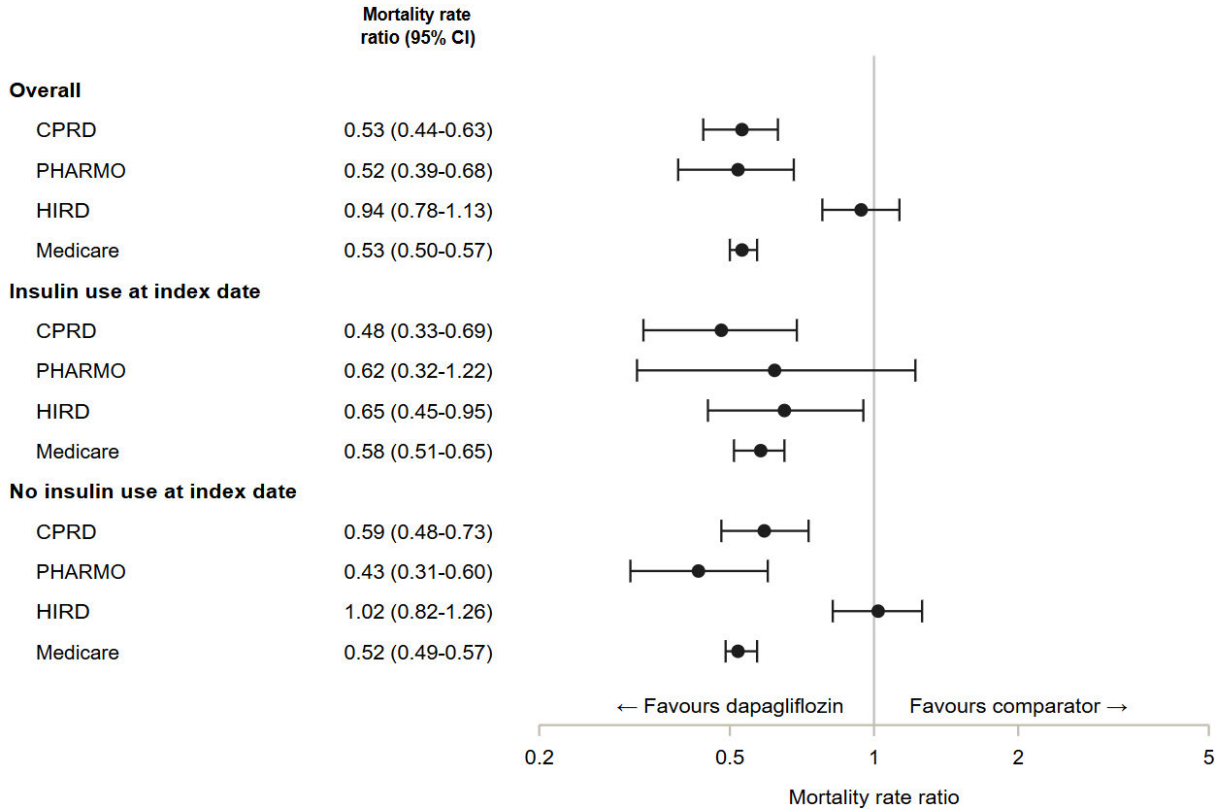
NE = not estimable.

10.5.2.5 Sensitivity Analysis: Evaluation of Deaths (All-Cause Mortality)

All-cause mortality rates are presented for each primary cancer outcome cohort (female breast cancer, sex-combined bladder cancer, female bladder cancer, and male bladder cancer) in Appendix J, BreastCa Table 25, BladderCa Table 25, BladderCa Table 25F, and BladderCa Table 25M, for each data source. All-cause mortality rate ratios are displayed for the female breast cancer cohorts in Figure 30 and for the bladder cancer cohorts (sex-combined and sex-specific) in Figure 31.

In the female breast cancer cohorts, all-cause mortality rates (per 10,000 person-years) in the overall cohorts ranged from 26.9 (HIRD) to 238.4 (Medicare) in the dapagliflozin group and ranged from 28.8 (HIRD) to 449.2 (Medicare) in the comparator AD group (Appendix J, BreastCa Table 25). Mortality rates were higher amongst insulin users than insulin non-users in both the dapagliflozin group and comparator AD group across all data sources. When comparing all-cause mortality in new users of dapagliflozin with new users of comparator AD in the female breast cancer cohorts, the mortality rate ratios were near or below 1.0 across the overall cohorts and the cohorts stratified by insulin use at the index date (Figure 30).

Figure 30 Sensitivity Analysis: Propensity Score–Adjusted All-Cause Mortality Rate Ratios for the Female Breast Cancer Cohorts, Overall and Stratified by Insulin Use at the Index Date, by Data Source

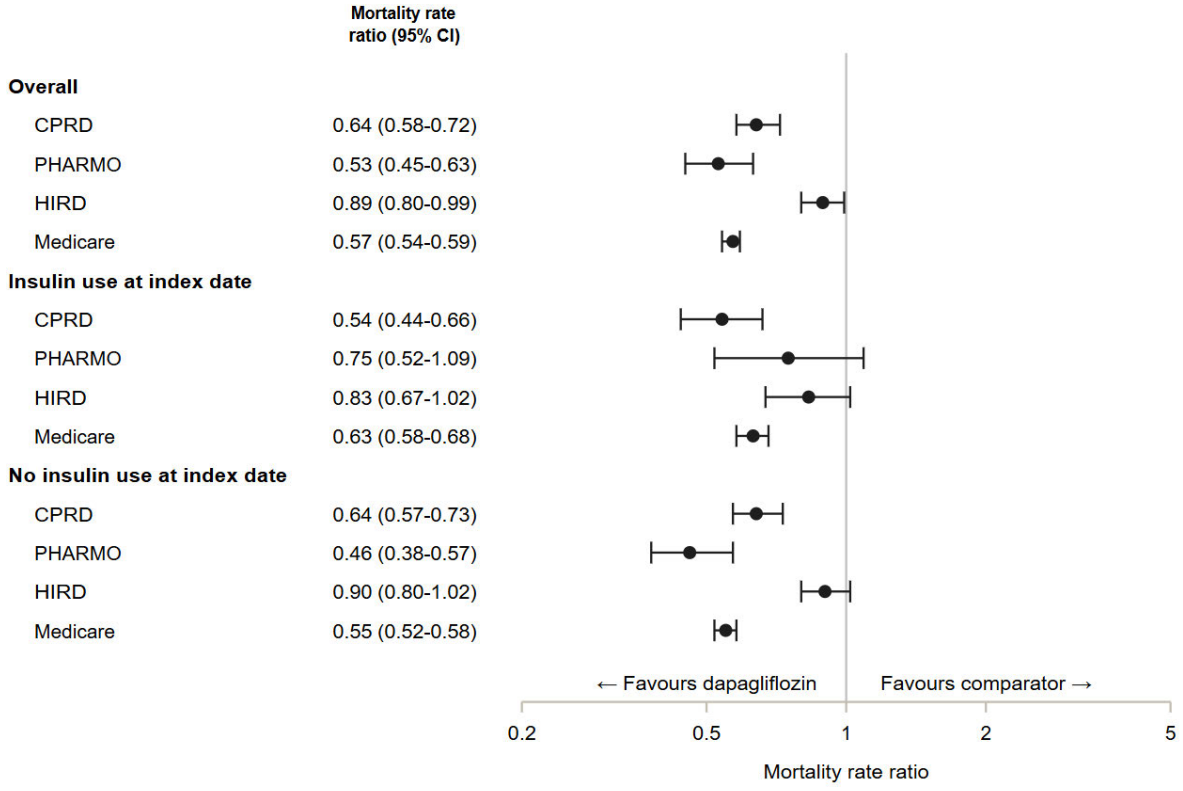


In the sex-combined bladder cancer cohorts, all-cause mortality rates (per 10,000 person-years) in the overall cohorts ranged from 34.2 (HIRD) to 261.6 (Medicare) in the dapagliflozin group and ranged from 38.5 (HIRD) to 460.9 (Medicare) in the comparator AD group. Like the female breast cancer cohort, mortality rates in the sex-combined bladder cancer cohorts were higher amongst insulin users than insulin non-users in both the dapagliflozin group and comparator AD group, for all data sources. When comparing all-cause mortality in new users of dapagliflozin with new users of comparator AD in the sex-combined bladder cancer cohorts, the mortality rate ratios were below 1.0 across the overall cohorts and the cohorts stratified by insulin use at the index date (Figure 31). The all-cause mortality rate and mortality rate ratio patterns observed in the female bladder cancer and male bladder cancer cohorts were similar to those observed in the sex-combined bladder cancer cohorts.

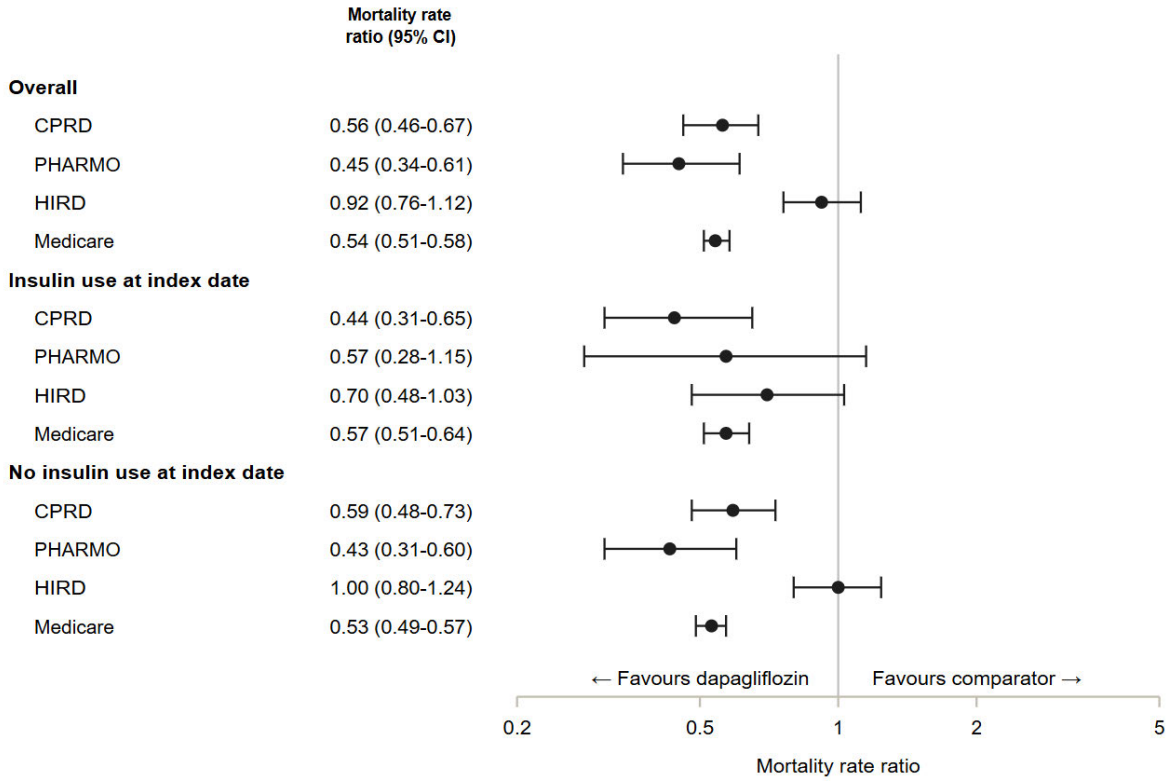
As the results across all primary cancer cohorts in all data sources do not suggest an increased mortality risk amongst new users of dapagliflozin compared with new users of comparator AD, additional competing risk analysis was not needed and was not conducted in any of the data sources (details described in Section 9.9.5.4).

Figure 31 Sensitivity Analysis: Propensity Score–Adjusted All-Cause Mortality Rate Ratios for the Bladder Cancer Cohorts, Overall and Stratified by Insulin Use at the Index Date, by Data Source

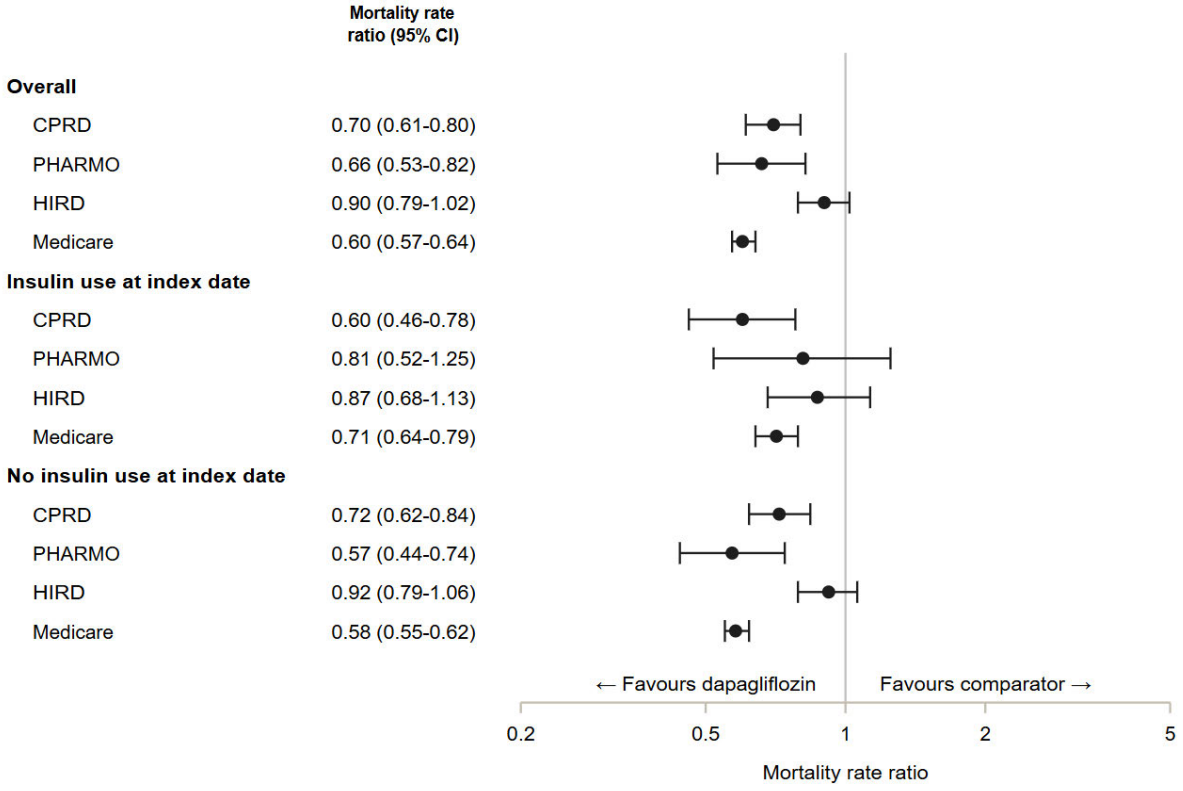
Sex-Combined Cohort



Female Cohort



Male Cohort



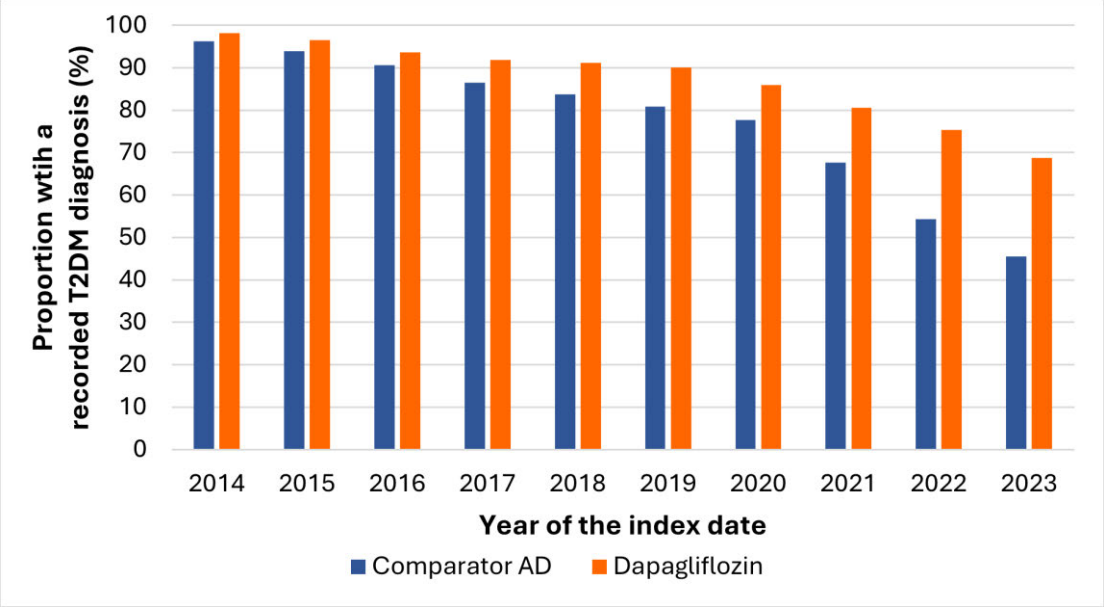
10.5.2.6 Post Hoc Sensitivity Analysis (the HIRD)

The results of the post hoc sensitivity analysis in the HIRD to explore the impact of potential confounding due to differences in the proportion of patients with a previous T2DM diagnoses in the female breast cancer cohorts (as described in Section 9.9.6) are presented in Figure 32 and Figure 33 (display prevalences of T2DM and obesity diagnoses in the HIRD female composite cancer cohort) as well as in Table 26 (displays the propensity score–adjusted incidence rate and propensity score–adjusted IRRs for female breast cancer when including the T2DM variable in the propensity score models).

10.5.2.6.1 Prevalence of T2DM and Obesity by Year of the Index Date Amongst Females in the HIRD

In this final (120-month) analysis, it was observed that, amongst females in the composite cancer cohort in the HIRD, there was a lower prevalence of having a recorded T2DM diagnosis before or on the index date amongst those in the comparator AD group (63%) than in the dapagliflozin group (85%). This finding persisted in the HIRD in the propensity score–trimmed samples (in which T2DM was not adjusted for in the propensity score model) in the overall cohort. In the propensity score–trimmed composite cancer cohorts, the proportion of patients with a T2DM diagnosis before or on the index date decreased over time. In 2014, amongst the female composite cancer cohort, 96% of comparator AD new users and 98% of dapagliflozin new users had a T2DM diagnosis before or on the index date. In 2023, 46% of comparator AD new users and 69% of dapagliflozin new users had a recorded T2DM diagnosis before or on the index date (Figure 32).

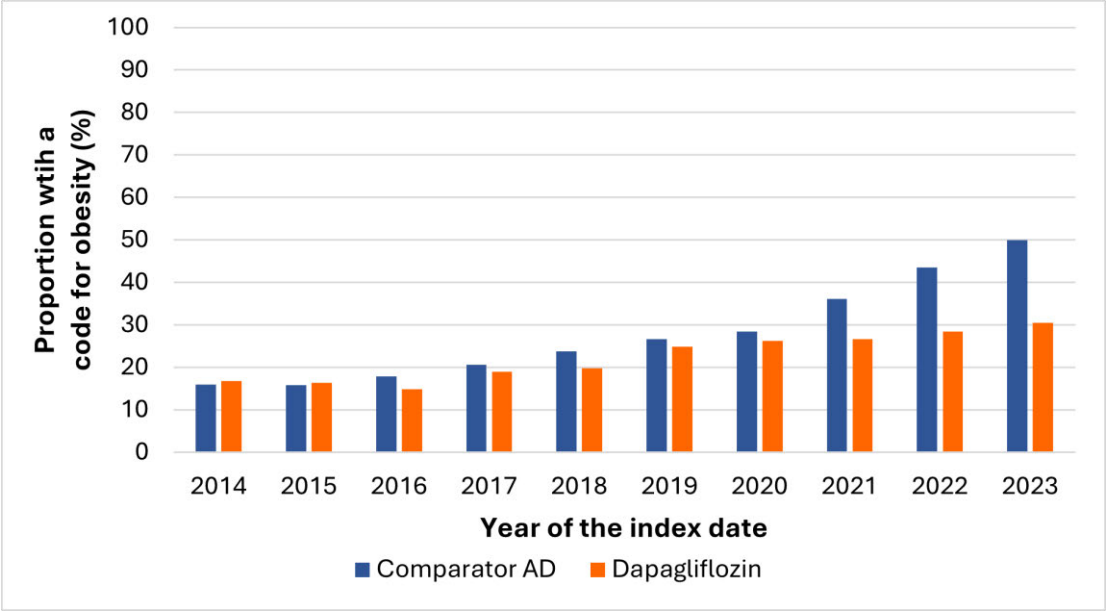
Figure 32 **Sensitivity Analysis (the HIRD): Prevalence of a Recorded T2DM Diagnosis Before or on the Index Date Over the Study Period in the Female Composite Cancer Cohort, Overall Cohort, Propensity Score–Trimmed Analysis Sample**



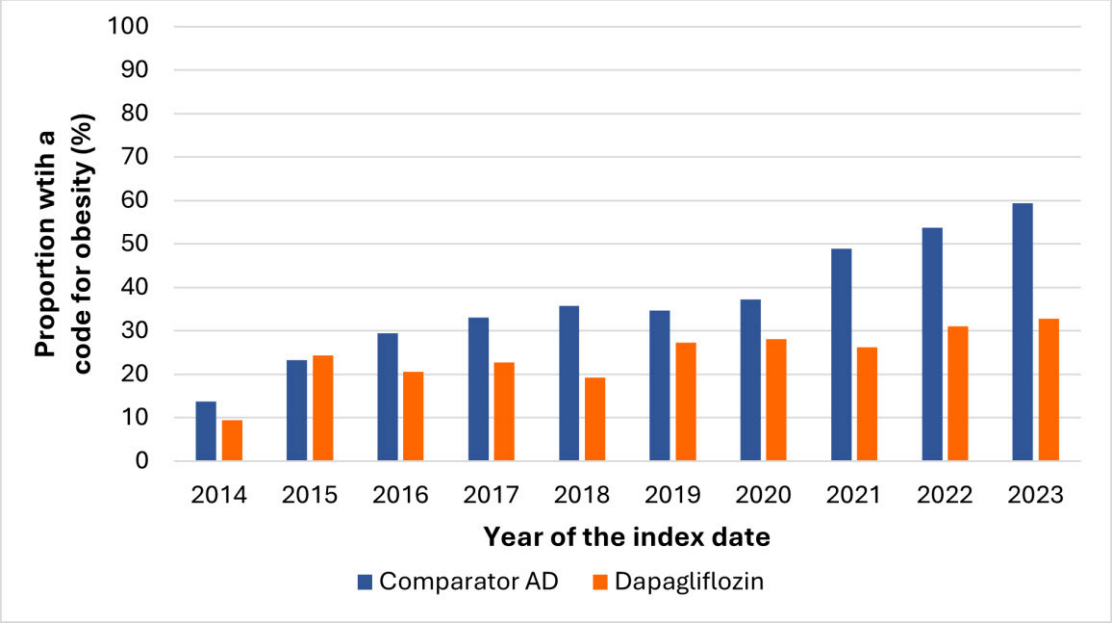
In contrast, the proportion of new users of comparator AD with a code indicating an obesity diagnosis before or on the index date increased over the study period (Figure 33). In 2014, amongst the female composite cancer cohort, 16% of comparator AD new users and 17% of dapagliflozin new users had a code for obesity before or on the index date. By 2023, these proportions increased to 50% and 31% in new users of comparator AD and dapagliflozin, respectively. Amongst females without a recorded T2DM diagnosis before or on the index date, 14% of comparator AD new users and 9% of dapagliflozin new users had a code for obesity in 2014; by 2023, these proportions increased to 59% of comparator AD new users and 33% of dapagliflozin new users.

Figure 33 Sensitivity Analysis (the HIRD): Prevalence of Obesity Before or on the Index Date Over the Study Period in the Female Composite Cancer Cohort, Overall Cohort

Females with or without a recorded T2DM diagnosis



Females without a recorded T2DM diagnosis



10.5.2.6.2 Female Breast Cancer Incidence Rates and Comparative Analysis When Including T2DM in the Propensity Score Models in the HIRD

After including T2DM in the propensity score model for the female breast cancer cohort, the propensity score–adjusted incidence rates of breast cancer amongst new users of dapagliflozin

and comparator AD in the overall cohort and the cohorts stratified by insulin use at the index date were similar to the incidence rates in the main analysis (Table 26). The propensity score–adjusted IRR (95% CI) in the overall cohort was 1.07 (0.91-1.25), which was consistent with the results of the main analysis (1.06 [0.90-1.24]). Likewise, the propensity score–adjusted IRRs in the cohorts stratified by insulin use at the index date were similar to those estimated in the main analysis.

Table 26 Sensitivity Analysis (the HIRD): Propensity Score–Adjusted Incidence Rates and Incidence Rate Ratios for Female Breast Cancer, Overall and Stratified by Insulin Use at the Index Date: Results From Main Analysis and Sensitivity Analysis When Including T2DM in the Propensity Score Models

	Insulin use at the index date		No insulin use at the index date		Overall	
	Dapagliflozin	**Comparator AD**	**Dapagliflozin**	**Comparator AD**	**Dapagliflozin**	**Comparator AD**
Main analysis						
Adjusted incidence rate[a] (95% CI)	45.58 (31.37-64.01)	32.40 (25.92-39.84)	34.38 (29.03-40.43)	33.23 (31.17-35.37)	34.62 (29.65-40.18)	32.98 (30.99-35.06)
Adjusted IRR[b] (95% CI)	1.42 (0.95-2.10)	Reference	1.04 (0.87-1.23)	Reference	1.06 (0.90-1.24)	Reference
Sensitivity analysis[c]						
Adjusted incidence rate[a] (95% CI)	45.68 (31.44-64.15)	31.77 (25.49-39.00)	33.50 (28.23-39.46)	33.33 (31.26-35.49)	34.74 (29.77-40.30)	32.61 (30.64-34.67)
Adjusted IRR[b] (95% CI)	1.45 (0.98-2.15)	Reference	1.01 (0.85-1.21)	Reference	1.07 (0.91-1.25)	Reference

[a] Propensity score–adjusted incidence rates were standardised to the person-time of the dapagliflozin cohort using the propensity score strata.

[b] Propensity score–adjusted IRRs were calculated using the stratified Mantel-Haenszel approach.

[c] Sensitivity included the inclusion of T2DM in the propensity score models before propensity score trimming and stratification.

10.6 Adverse Events/Adverse Reactions

In alignment with guidance from the International Society for Pharmacoepidemiology's *Guidelines for Good Pharmacoepidemiology Practices (GPP)* [69] and the EMA's *Guideline on Good Pharmacovigilance Practices (GVP), Module VI – Collection, Management, and Submission of Reports of Suspected Adverse Reactions to Medicinal Products* [38] and *Module VIII – Post-Authorisation Safety Studies* [39], individual case study reports of suspected adverse reactions are not required for non-interventional, post-authorisation safety studies based in secondary health databases.

11. DISCUSSION

11.1 Key Results

The primary study objectives were to evaluate the risk of female breast cancer and sex-combined and sex-specific in situ and invasive bladder cancer amongst new users of dapagliflozin compared with new users of other ADs and to evaluate risk stratified by insulin use at the index date and, for sex-combined bladder cancer, stratified by pioglitazone use at the index date. The secondary study objectives were to compare baseline characteristics of new users of dapagliflozin with new users of comparator AD, assess the potential for surveillance bias to explain observed findings, and evaluate the risk of composite cancers across exposure groups in males and females.

In this final analysis, based on data from November 2012 through September 2023, the study population included 186,881 dapagliflozin new users (80,966 females and 105,915 males) and 1,896,899 comparator AD new users (1,002,460 females and 894,439 males) before propensity score trimming. After applying outcome-specific exclusion criteria, the study population included 80,707 dapagliflozin new users and 999,103 comparator AD new users for the female breast cancer cohort and 182,050 dapagliflozin new users (79,025 females and 103,025 males) and 1,839,218 comparator AD new users (974,493 females and 864,725 males) for the sex-combined bladder cancer cohort.

11.1.1 Primary Objective

11.1.1.1 Female Breast Cancer

Incidence analyses for female breast cancer were based on 130,358 person-years of exposure to dapagliflozin across the four data sources. The mean duration of follow-up in dapagliflozin new users ranged from 1.9 years (the HIRD and Medicare) to 2.9 years (CPRD) and in comparator AD new users ranged from 1.6 years (the HIRD) to 3.3 years (PHARMO). Across all data sources, the propensity score–adjusted incidence rate of female breast cancer, in the overall cohorts, was comparable amongst dapagliflozin new users (propensity score–adjusted incidence rate, range, 28.7-39.1 per 10,000 person-years) and comparator AD new users (propensity score–adjusted incidence rate, range, 28.7-40.3 per 10,000 person-years). The highest breast cancer incidence rates were observed amongst older women, including in the subset of females in CPRD that were aged ≥ 65 years and in Medicare (a data source that included women who were only aged ≥ 65 years). The propensity score–adjusted incidence rate estimates observed in CPRD, PHARMO, the HIRD, and Medicare in this study were similar to the age-adjusted incidence rates of breast cancer amongst women aged 40 years or older in the general population in the UK and the Netherlands and amongst women aged 40 to 64 years and 65 years of age or older in the US general population [34; 36; 86].

The propensity score–adjusted IRRs for the overall female breast cancer cohorts were close to or below 1 across all data sources: CPRD (1.05; 95% CI, 0.75-1.46), PHARMO (0.87;

95% CI, 0.46-1.63), the HIRD (1.06; 95% CI, 0.90-1.24), and Medicare (0.98; 95% CI, 0.86-1.11). The 95% CI was widest for the PHARMO estimate, indicating less precision due to fewer patients and breast cancer cases than the other data sources. The overall pooled estimate for the adjusted IRR for female breast cancer was 1.01 (95% CI, 0.92-1.11). Amongst insulin users, the propensity score–adjusted IRRs (although imprecise) were higher than estimates in the overall cohorts, with the highest propensity score–adjusted IRR observed in the HIRD (1.42; 95% CI, 0.95-2.10). In PHARMO, the propensity score–adjusted IRR was not estimable in the insulin users cohort due to the low number of breast cancer cases (ie, a single case of breast cancer was observed amongst dapagliflozin users).

The female breast cancer IRR results observed in the present study had higher precision than the results from a 2017 meta-analysis of randomised clinical trials that reported elevated, although imprecise, effect estimates for breast cancer risk with dapagliflozin (odds ratio [OR], 2.50; 95% CI, 0.88-7.09) and SGLT2 inhibitors (OR, 1.68; 95% CI, 0.87-3.22) compared with other active blood glucose–lowering treatments; the meta-analysis included a limited number of patients and breast cancer cases and a short mean trial duration (61 weeks) [121]. In addition, the female breast cancer IRR estimates in the present study are similar to the relative risk estimates from a 2023 meta-analysis of 76 randomised controlled trials with a minimum follow-up of 48 weeks (range, 48 to 295.9 weeks) that reported no increased risk of breast cancer with dapagliflozin (RR, 0.98; 95% CI, 0.66-1.47) or SGLT2 inhibitors (RR, 1.01; 95% CI, 0.77-1.32) compared with placebo, active interventions, or no intervention [116]. The female breast cancer IRR estimates in the current study are also consistent with results from an observational study that used a new-user active-comparator approach to evaluate the association of SGLT2 inhibitors with incident cancer in individuals with diabetes, using data from a nationwide Japanese database with a mean duration of follow-up of 2 years [120]. In the primary analysis, the HR of female breast cancer for SGLT2 inhibitors compared with DPP-4 inhibitors was 0.72 (95% CI, 0.42-1.24). In sensitivity analyses restricted to individuals with T2DM, the HR of female breast cancer was 0.69 (95% CI, 0.38-1.24); applying IPTW, the HR was 0.85 (95% CI, 0.49-1.46); and after adjusting for lifestyle risk factors (smoking, alcohol consumption), the HR was 0.58 (95% CI, 0.30-1.14). Analysis of individual SGLT2 inhibitors did not show differences in risk of total cancer [120]. Although multiple studies have assessed the potential effect of insulin on the risk of breast cancer (with inconsistent results due to the different study design approaches, time-related biases, and duration of follow-up of the studies) [17; 21; 137], and other studies have assessed the risk of breast cancer in association with SGLT2 inhibitor use [116; 120; 121], to our knowledge, no studies have assessed SGLT2 inhibitor use and breast cancer risk separately amongst insulin users. Given this, the present study is the first observational study to evaluate the relationship between use of an SGLT2 inhibitor (ie, dapagliflozin) and breast cancer risk specifically amongst insulin users and resulted in a pooled adjusted IRR of 1.17 (95% CI, 0.94-1.46); the pooled IRR estimate was based on data from CPRD, the HIRD, and Medicare, which, due to the small proportions of the study samples that were insulin users (≤ 15% across all data

sources), yielded imprecise data source−specific IRRs with elevated point estimates in some data sources (ie, CPRD and the HIRD).

The electronic algorithms used in the current study to identify female breast cancer cases had high validity; PPVs ranged from 81.5% to 99.0% in the CPRD, the HIRD, and Medicare. Simulation analyses conducted to assess the impact of potential outcome misclassification indicated that it is unlikely that an increased risk of breast cancer associated with dapagliflozin is being masked by differential outcome misclassification.

Results from the sensitivity analyses that removed various censoring criteria yielded propensity score–adjusted IRRs that were similar to the propensity score–adjusted IRRs from the main analysis for the female breast cancer cohorts, across all data sources. The analysis of all-cause mortality did not indicate an increased risk of mortality in the dapagliflozin group, therefore, eliminating the concern that mortality may be masking a potential increased risk of cancer in the dapagliflozin group.

11.1.1.2 Bladder Cancer

Incidence analyses for sex-combined bladder cancer were based on 286,950 person-years of exposure to dapagliflozin across the four data sources. The numbers of bladder cancer events amongst dapagliflozin new users were low in PHARMO and the HIRD, as well as in the insulin use cohorts across all data sources. The mean duration of follow-up in dapagliflozin new users ranged from 1.9 years (the HIRD) to 2.9 years (CPRD) and in comparator AD new users ranged from 1.6 years (the HIRD) to 3.4 years (PHARMO).

In all data sources, the propensity score–adjusted incidence rate of sex-combined bladder cancer in the overall cohorts was somewhat lower in dapagliflozin new users (range, 2.3-13.3 per 10,000 person-years) than in comparator AD new users (range, 2.7-14.9 per 10,000 person-years). Propensity score–adjusted incidence rates in the overall cohorts were higher amongst new users aged at least 65 years than amongst new users aged younger than 65 years. The adjusted bladder cancer incidence rates were lower amongst females than amongst males, where propensity score–adjusted incidence rates were estimable. The propensity score–adjusted incidence rates of bladder cancer observed for both exposure groups in CPRD in this study were similar to the age-adjusted incidence rate of bladder cancer (in situ and invasive) amongst the UK general populations, respectively, aged 40 years or older [33; 35]. In the HIRD, the propensity score–adjusted incidence rates of bladder cancer were also higher (incidence rate per 10,000 person-years: dapagliflozin, 2.3; comparator AD, 2.7) than the bladder cancer age-adjusted incidence rates amongst the US general population aged 40 to 64 years (1.2 per 10,000 person-years) [87]. Finally, the propensity score–adjusted incidence rates in Medicare in this study (incidence rate per 10,000 person-years: dapagliflozin, 6.6; comparator AD, 8.8) were lower than the bladder cancer age-adjusted incidence rates in the US general population aged ≥ 65 years (11.2 per 10,000) [87]. The higher bladder cancer

incidence rates observed in the HIRD may be linked to differences in health care–seeking behaviour and potentially to differences in lifestyle factors between individuals that have medical care coverage mostly through employer-sponsored health insurance and the general population [124]. In PHARMO's sample, the mean age in the bladder cancer propensity score–trimmed sample was 64 years for the dapagliflozin group and 65 years for the comparator AD group. The crude incidence rate of bladder cancer in the Dutch general population aged 65 to 69 years is approximately 10 per 10,000 person-years, which is similar to the incidence rates observed in PHARMO in the current study [90]. In addition, the postulated relationship of diabetes with an elevated risk of bladder cancer [2] and the close monitoring of patients with T2DM in the Netherlands may also explain the higher incidence rates of bladder cancer observed in the PHARMO sample than the rates reported in the general population [97]. The lower bladder cancer incidence rates observed in Medicare may be due to the exclusion of individuals with a prior history of cancer in the current study and also to delays in cancer diagnoses and more specialised cancer screening (such as screening for bladder cancer) experienced during the COVID-19 pandemic beginning in March 2020; this may particularly impact Medicare, for which the study period ended on 31 December 2021, a time when the pandemic was still impacting health care systems and health care–seeking behaviour [46; 77; 82].

In all data sources, the overall propensity score–adjusted IRR point estimates for sex-combined bladder cancer were below 1: CPRD, 0.74 (95% CI, 0.45-1.21); PHARMO, 0.90 (95% CI, 0.48-1.68); the HIRD, 0.82 (95% CI, 0.55-1.24), and Medicare 0.74 (95% CI, 0.59-0.94); the precision of estimates was limited due to small sample sizes and low numbers of bladder cancer cases in all data sources. The overall pooled adjusted IRR estimate for sex-combined bladder cancer was 0.77 (95% CI, 0.64-0.92). Amongst insulin users, the propensity score–adjusted IRRs were slightly higher than the overall cohorts in CPRD (1.01; 95% CI, 0.33-3.09) and lower than the overall cohorts in the HIRD (0.75; 95% CI, 0.20-2.75) and Medicare (0.36; 95% CI, 0.16-0.77). In PHARMO, the propensity score–adjusted IRR was not estimable in the insulin users cohort due to the low number of bladder cancer cases, particularly amongst dapagliflozin users. The current study's results are consistent with results from an observational study conducted by Abrahami et al [1], with a design similar to the current study and that also included data from CPRD and Medicare, amongst other data sources. In the Abrahami et al [1] study, the risk of bladder cancer amongst adults with T2DM who were initiators of SGLT2 inhibitors was lower than the risk for initiators of GLP-1 receptor agonists (HR, 0.90; 95% CI, 0.81-1.00) and similar to the risk for DPP-4 inhibitor initiators (HR, 0.99; 95% CI, 0.91-1.09) [1]. The bladder cancer IRR estimates in the current study are also similar to the relative risk (RR) estimates from a 2023 meta-analysis of 76 randomised controlled trials with a minimum follow-up of 48 weeks (range, 48 to 295.9 weeks) that reported no increased risk of bladder cancer with dapagliflozin (RR, 0.63; 95% CIs, 0.42-0.95) or SGLT2 inhibitors (RR, 0.93; 95% CI, 0.71-1.21) compared with placebo, active interventions, or no intervention [116]. Furthermore, the current study's results

are consistent with results from an observational study evaluating the association of SGLT2 inhibitors with incident cancer in individuals with diabetes using data from a nationwide Japanese database [120]. In the primary analysis, the HR of bladder cancer for SGLT2 inhibitors compared with DPP-4 inhibitors was 0.84 (95% CI, 0.47-1.51). In sensitivity analyses restricted to individuals with T2DM, the HR of bladder cancer was 1.35 (95% CI, 0.69-2.66); applying IPTW, the HR was 0.77 (95% CI, 0.42-1.42); and after adjusting for lifestyle risk factors (smoking, alcohol consumption), the HR was 0.45 (95% CI, 0.21-0.97). Analysis of individual SGLT2 inhibitors did not show differences in risk of total cancer amongst individual SGLT2 inhibitors [120].

In this final 120-month analyses, adjusted incidence and comparative analyses for female bladder cancer could not be conducted in any of the overall or insulin use–stratified cohorts in PHARMO and the HIRD, in the insulin non-users cohort in CPRD, and in the insulin user cohort in Medicare, due to the small number of bladder cancer cases, particularly in dapagliflozin new users. The overall propensity score–adjusted IRR point estimates for bladder cancer in females were less than 1 in CPRD and Medicare; however, the estimates were imprecise because of small numbers. The overall propensity score–adjusted IRR point estimates for bladder cancer in males were below 1 in CPRD (0.91; 95% CI, 0.54-1.53) and Medicare (0.76; 95% CI, 0.59-0.98), equal to 1 in the HIRD (1.00; 95% CI, 0.65-1.54), and above 1 in PHARMO (1.06; 95% CI, 0.55-2.04); however, the precision of the IRR estimates was limited due to small numbers.

The electronic algorithms used to identify bladder cancer cases had high validity; PPVs ranged from 88.2% to 97.8% in CPRD, the HIRD, and Medicare. Simulation analyses conducted to assess impact of potential outcome misclassification indicated that it is unlikely that an increased risk of bladder cancer associated with dapagliflozin is being masked by differential misclassification.

Results from the sensitivity analyses that removed various censoring criteria yielded propensity score–adjusted IRRs for the sex-combined bladder cancer, which were similar to the propensity score–adjusted IRRs from the main analysis, across all data sources. The analysis of all-cause mortality did not indicate an increased risk of mortality in the dapagliflozin group, and, therefore, eliminated the concern that mortality may be masking a potential increased risk of cancer amongst dapagliflozin new users.

11.1.1.2.1 Bladder Cancer Stratified by Pioglitazone Use

In the sex-combined bladder cancer cohort, the pioglitazone use–stratified analyses in the propensity score–trimmed cohorts were limited due to the infrequent use of pioglitazone in both exposure cohorts (CPRD, 6.7%; the HIRD, 5.1%; and Medicare, 11.1%), particularly amongst dapagliflozin new users (CPRD, 4.4%; the HIRD, 3.8%; and Medicare, 5.9%). In PHARMO, propensity score modelling was unable to be conducted in the pioglitazone use

group due to the low number bladder cancer cases amongst the pioglitazone users (dapagliflozin, n = 0; comparator AD, n = 1); therefore, the propensity score–trimmed samples in the pioglitazone use–stratified cohorts were not generated, and the pioglitazone use–stratified incidence and comparative analyses were not conducted in PHARMO.

Amongst pioglitazone users, the propensity score–adjusted incidence rate per 10,000 person-years for sex-combined bladder cancer was estimable only amongst comparator AD new users in CPRD (11.0; 95% CI, 2.3-24.8) and Medicare (11.8; 95% CI, 9.6-14.2) with very wide 95% CIs due to the low number of events, particularly amongst the dapagliflozin group. The propensity score–adjusted incidence rate was not estimable in either exposure group in the HIRD due to the low number of bladder cancer cases. In the pioglitazone non-user cohort, the propensity score–adjusted incidence rate per 10,000 person-years amongst dapagliflozin new users was estimable only in the HIRD (2.2; 95% CI, 1.4-3.3) and was lower in the HIRD (2.6; 95% CI, 2.2-3.1) than in Medicare (8.4; 95% CI, 7.9-9.0) amongst comparator AD new users.

Amongst pioglitazone users, the propensity score–adjusted IRR (95% CI) estimate for sex-combined bladder cancer was 1.41 (0.35-5.60) in CPRD and 0.82 (0.41-1.67) in Medicare. The 95% CIs of the propensity score–adjusted IRR estimates were wide due to the small size of the pioglitazone user cohorts and low number of bladder cancer events.

For the pioglitazone non-user cohort, the propensity score–adjusted IRR (95% CI) estimate for sex-combined bladder cancer was 0.76 (0.44-1.32) in CPRD, 0.84 (055-1.29) in the HIRD, and 0.71 (0.55-0.91) in Medicare, similar to the propensity score–adjusted IRRs observed in the overall cohorts.

11.1.2 Secondary Objectives

11.1.2.1 Baseline Characteristics

Across all cancer outcome cohorts, before propensity score trimming, in the overall and in the insulin use–stratified cohorts, the baseline prevalences of the most common indicators of diabetes severity (coronary heart disease in PHARMO and peripheral vascular disease and retinopathy in the HIRD and Medicare) were higher in dapagliflozin new users than in comparator AD new users, whereas in CPRD, the baseline prevalence of retinopathy was similar in both exposure groups. Across all cancer outcome cohorts in all data sources, the baseline prevalences of the diabetes severity indicators were higher in the insulin users cohorts than either the overall cohorts or the insulin non-users cohorts. The baseline prevalence of heart failure was higher amongst dapagliflozin new users than comparator AD new users in PHARMO and the HIRD, and similar across both exposure groups in CPRD and Medicare. The baseline prevalence of chronic kidney disease was higher in comparator AD new users than dapagliflozin new users in CPRD, higher in dapagliflozin new users than comparator AD new users in the HIRD, and was similar in both exposure groups in PHARMO and Medicare. In CPRD, for the female breast cancer cohort, across the overall and insulin use–stratified

cohorts, dapagliflozin new users were more likely to be severely obese than comparator AD new users, whereas in PHARMO, the proportions for each BMI category were similar across the exposure groups. In CPRD and PHARMO, the proportions of current smokers or former smokers and alcohol users with low-moderate, heavy, or very heavy alcohol intake were similar in dapagliflozin new users and comparator AD new users for the primary outcome cohorts. Across all cancer outcome cohorts in all data sources, before propensity score trimming, dapagliflozin new users were more likely than comparator AD new users to have had at least three AD classes prescribed or dispensed during various time periods before the index date. The distribution of other cancer outcome–specific relevant comorbidities was similar between exposure groups for each respective cancer outcome cohort across all data sources.

Across all cancer outcome cohorts and data sources, the prevalence of concomitant insulin use at the index date amongst dapagliflozin new users was approximately equal to or somewhat lower than the prevalence anticipated in the study protocol (20%). Amongst comparator AD new users, the prevalence of insulin use at the index date was approximately 10% in CPRD, PHARMO, and the HIRD and 15% in Medicare.

Propensity scores were generated to improve comparability of the exposure groups. When assessing the propensity score distributions for dapagliflozin new users and comparator AD new users, there was adequate overlap of the propensity score distributions for all cancer outcomes in CPRD, the HIRD, and Medicare. In PHARMO, the propensity score distribution was skewed to the right across all cancer outcome cohorts with limited area of support for dapagliflozin-treated patients (see Section 11.2 for additional details). After propensity score trimming and stratification, baseline characteristics were generally adequately balanced between dapagliflozin and comparator AD in most propensity score strata across all study outcome cohorts in all data sources. For the variables that remained imbalanced after propensity score trimming, the differences between dapagliflozin new users and comparator new users were small (StDiff values < 0.30).

11.1.2.2 Surveillance Bias

The sensitivity analysis that evaluated the weighted cumulative incidence of female breast cancer and bladder cancer over follow-up did not indicate any potential for a surveillance bias associated to dapagliflozin exposure or differences between the exposure groups in the time to occurrence of the primary outcomes during follow-up.

For the female breast cancer and bladder cancer cohorts, there were no relevant differences in HCRU measures between dapagliflozin users and comparator AD users throughout the yearly study periods assessed, across all data sources, suggesting that there were no differences in the frequency of HCRU between the exposure groups.

11.1.2.3 Female Composite Cancer

Incidence analyses for female composite cancers were based on 130,787 person-years of exposure to dapagliflozin across the four data sources. For the overall female composite cancer cohorts (not stratified by insulin use), the frequency of the selected composite cancers in CPRD, the HIRD, and Medicare, was similar amongst dapagliflozin new users (propensity score–adjusted incidence rate, range, 42.6-44.4 per 10,000 person-years) and comparator AD new users (propensity score–adjusted incidence rate, range, 38.9-43.8 per 10,000 person-years). In PHARMO, the propensity score–adjusted incidence rates were higher than in the other data sources and were similar in dapagliflozin new users (66.7 per 10,000 person-years) and comparator AD new users (72.1 per 10,000 person-years). Although the incidence rates observed for the female composite cancer outcome were higher in PHARMO than in the other data sources, the rates are comparable to those reported in the general Dutch population. From 2014 to 2022 in the general Dutch population, the age-specific crude incidence rate for the cancer types included in the female composite cancer outcome ranged from 21 to 24 per 10,000 person-years for women aged 45 to 59 years, 56 to 66 per 10,000 person-years for those aged 60 to 74 years, and 71 to 78 per 10,000 person-years for the women aged ≥ 75 years [91].

The propensity score–adjusted IRRs across all data sources did not suggest an increased risk of female composite cancers associated with dapagliflozin exposure. Propensity score–adjusted IRRs were 0.99 (95% CI, 0.75-1.30) in CPRD, 0.92 (95% CI, 0.59-1.41) in PHARMO, 1.11 (95% CI, 0.96-1.28) in the HIRD, and 1.01 (95% CI, 0.89-1.15) in Medicare. Propensity score–adjusted IRRs were higher in younger (< 65 years) than older (≥ 65 years) females in CPRD and PHARMO.

11.1.2.4 Male Composite Cancer

Incidence analyses for male composite cancers were based on 165,308 person-years of exposure to dapagliflozin across the four data sources. For the overall cohorts (not stratified by insulin use), the selected male composite cancers were less frequent amongst dapagliflozin new users (propensity score–adjusted incidence rate, range, 49.2-100.9 per 10,000 person-years) than comparator AD new users (propensity score–adjusted incidence rate, range, 55.9-125.9 per 10,000 person-years) across all data sources. Propensity score–adjusted incidence rates (per 10,000 person-years) of the male composite cancer outcome in dapagliflozin and comparator AD new users were lowest in the HIRD (49.2 and 55.9, respectively; study population aged 40-64 years) and highest in Medicare (100.9 and 102.4, respectively; study population aged ≥ 65 years) and PHARMO (176.3 and 189.1, respectively; study population aged ≥ 65 years). The crude incidence rate of the male composite cancer outcome in the general Dutch population aged ≥ 65 years was approximately 160 per 10,000 person-years during the period of 2014-2022 [92] and, in the general US population aged ≥ 65 years, was 144 per 10,000 persons-years during the period of 2014-2021 [88]. The incidence rates observed for the male composite cancer outcome in PHARMO in this study are comparable to

or higher than those reported in the general Dutch population likely due to the older age of the study samples. In Medicare, the observed incidence rates were lower than the incidence rates reported in the US general population aged ≥ 65 years.

Propensity score–adjusted IRRs across all data sources did not suggest an increased risk of male composite cancers associated with dapagliflozin exposure. Propensity score–adjusted IRRs were 0.89 (95% CI, 0.74-1.06) in CPRD, 0.74 (95% CI, 0.54-1.01) in PHARMO, 0.88 (95% CI, 0.79-0.99) in the HIRD, and 0.98 (95% CI, 0.90-1.06) in Medicare. Propensity score–adjusted IRRs were higher in younger (aged < 65 years) than older (aged ≥ 65 years) males in CPRD, whereas the opposite was found in PHARMO.

11.2 Limitations

The findings of the study should be interpreted in the context of some limitations.

11.2.1 Potential Selection Bias

Detection bias (including surveillance bias) is a form of selection bias that can occur when there is a difference in the measurement or detection of an outcome [135]. This can manifest itself in multiple ways in epidemiological observational studies of cancer outcomes, including differences in timing, and/or screening or testing across exposure groups, making it more or less likely to detect a given cancer in one exposure group than the other. In the case of the present study, if an early imbalance in cancer cases had been detected for breast and bladder cancers amongst patients taking dapagliflozin [20], it might indicate that clinicians may be more likely to screen patients on dapagliflozin. Alternatively, noted side effects listed on the product label [52], such as urinary tract infections, might result in more frequent testing and increased likelihood of early detection of bladder cancer. However, the final (120-month) analysis results did not suggest an increased risk of breast or bladder cancer in new users of dapagliflozin compared with new users of the comparator AD. Additionally, as part of sensitivity analyses, weighted cumulative incidence plots were generated to visualise the estimated probability of each primary cancer outcome, individually in patients exposed to dapagliflozin and to the comparator AD over time. For the primary cancer outcomes (female breast cancer and bladder cancer), cumulative incidence curves were similar in the dapagliflozin and comparator AD groups. If surveillance bias had been at play, one might have expected to see separated curves with a higher incidence in the dapagliflozin group soon after medication initiation, where increased testing and/or screening may have resulted in more preclinical cancer cases being diagnosed. Lastly, a descriptive analysis of HCRU found no notable differences in receipt of screening and tests between the dapagliflozin and comparator AD groups over time.

11.2.2 Potential Confounding Bias

Overall, to control for confounding, propensity score models were constructed using all measured confounders specified in the protocol and SAP (see list of stand-alone documents in Appendix A). Tertiles to deciles of the propensity score were created with differences in characteristics assessed within these propensity score strata and then used to compute the adjusted incidence rates and IRRs. Separate propensity score models were constructed for the subgroup analyses (ie, insulin and pioglitazone use at index date) and for the final (120-month) analysis, for male and female bladder cancer, separately. This likely mitigated the occurrence of imbalances in baseline characteristics that may arise within subgroups when the propensity score is generated in the overall sample, especially if the process by which treatment factors are assigned varies across subgroups. Simulation studies have demonstrated that this is especially true when sample size is smaller or outcomes rare [58]. However, the role of residual confounding attributed to how potential confounders are measured cannot be ruled out.

Lifestyle variables such as BMI, smoking status, and alcohol consumption that could be important confounding variables in the study were available in CPRD for the entire study sample and in a subset of patients (40%) with linked GP data in PHARMO and were not available in the HIRD or Medicare claims data. Therefore, these variables could not be used in propensity score models in the HIRD or with Medicare claims data. In CPRD, these variables were assessed in the entire medical history of the patient, using the values that were closest in time before the index date. Variables assessed using values recorded closer to the index date can be expected to be more relevant than those recorded farther away in time from the index date, thereby mitigating residual confounding originating from these variables not being measured appropriately. Although data on smoking status and BMI were available for almost all patients in CPRD, these data were not available in PHARMO for approximately 60% of the patients (ie, patients without availability of GP data or with data not known or not recorded by GPs). In the CPRD, the completeness of BMI data has been shown to vary by age and sex but has been documented to be complete in 97% of patients with a record of T2DM [14].

Comorbid conditions (eg, chronic kidney disease, cerebrovascular disease, and peripheral vascular disease) were considered for inclusion in propensity score models as potential confounders. In PHARMO, diagnoses for conditions related to severity of diabetes and for other specific medical conditions of interest were captured from hospital admission diagnoses, which would not accurately capture diagnoses of conditions likely to be diagnosed and treated in general practice or ambulatory health care settings rather than in a hospital (ie, retinopathy, peripheral vascular disease), likely resulting in incomplete capture of diagnoses of these conditions. However, data missing due to lack of linkage to GP data is expected to be missing at random across exposure groups; therefore, the impact of such missing data is likely to bias estimates towards the null (M. Meulendijk, Unpublished data. 2023). Additionally, in the final

analysis, for variables that might contain missing information attributed to the lack of linkage to GP data (eg, BMI, smoking status, retinopathy, peripheral vascular disease, urinary tract infections), the propensity score models contained an interaction term between the indicator variable for the covariate of interest and the inverse of the indicator variable for whether GP data were available. This interaction term was to account for missing data attributed to lack of linkage to GP data within the effect estimates and to distinguish this from data for a variable missing due to other reasons (eg, due to information not being entered by a health care provider).

The eligibility criterion for identifying new users of dapagliflozin in this study was defined at a time when dapagliflozin's single approved indication in the US and the EU was for the treatment of T2DM. The observation periods in the final analysis overlapped with the approval periods of the additional indications for heart failure in 2020 and chronic kidney disease in 2021. The additional subsequently approved indications for dapagliflozin for the treatment of heart failure and chronic kidney disease presented a challenge in the current study as patients may have been prescribed dapagliflozin for the treatment of these non-T2DM indications, as opposed to the initial indication of interest in this study. In PHARMO and the HIRD, the two data sources for which the study period ended after the approval of new indications for dapagliflozin (December 2022 for PHARMO and September 2023 for the HIRD), the prevalence of heart failure and chronic kidney disease was higher in dapagliflozin new users than in comparator AD new users. In CPRD and Medicare, the end of the study period coincided with or was closer in time to the time of approvals, which limited the interpretation of the frequency of heart failure and chronic kidney disease in these data sources. In the final analysis, although the initial criteria for inclusion were retained, covariates for heart failure, and chronic kidney disease were considered for inclusion in the propensity score models, and in cases where the variable met the criteria, variables were included in the propensity score model as covariates. Additionally, the presence of a diagnosis of T2DM in the data source was not a required criterion for inclusion into the study cohort, as at the time of study commencement, T2DM was the only indication for receipt of dapagliflozin and comparator AD medications. However, since 2021, drug classes included in the comparator AD group, such as GLP-1 receptor agonists, have been approved for indications other than diabetes, such as for treatment of obesity [45; 95]. A lower prevalence of T2DM was noted in the comparator AD group in the HIRD, which contains more recent data (through September 2023) and may reduce the comparability of the two exposure groups. However, in a post hoc sensitivity analysis conducted in the breast cancer cohort in the HIRD, no notable difference was observed in the incidence of breast cancer when T2DM was included as a covariate in the propensity score model compared with the main analysis, which did not include T2DM as a covariate in the propensity score model. This suggests that not considering T2DM in the propensity score models in the main analysis was unlikely to have biased the IRR estimates.

In the present study, adequate overlap ("common support") of the propensity score distributions for dapagliflozin new users and comparator AD new users in CPRD, the HIRD, and Medicare were observed for all cancer cohorts, in the overall and insulin use–stratified groups. In PHARMO, a right-skewed distribution was observed across all cohorts with limited area of common support for dapagliflozin-exposed patients. The lack of common support across exposure groups is likely attributed to changes in the treatment guidelines in the Netherlands in recent years that have favoured dapagliflozin as a first-line therapy for very high-risk patients with T2DM. In 2021, T2DM treatment guidelines in the Netherlands were updated to provide broader recommendations for people with T2DM at high risk of cardiorenal disease based on results of the cardiovascular and kidney outcomes trials involving SGLT2 inhibitors such as dapagliflozin [50; 94]. In this final (120-month) analysis, the majority of new users of dapagliflozin in the PHARMO cohorts, including 61% of new users of dapagliflozin in the female breast cancer cohort and 60% in the sex-combined bladder cancer cohort (before propensity score trimming), initiated the drug during the period of March 2021 through December 2022, after the release of the updated treatment guidelines. In light of this, the broadening of the T2DM treatment guidelines in the Netherlands in 2021, in combination with the additional indications of dapagliflozin for heart failure and chronic kidney disease, may be underlying the differences in this current study in the observed baseline characteristics of new users of dapagliflozin initiating treatment during recent years compared with new users who initiated dapagliflozin in the past, explaining some of the limited overlap observed in propensity score distributions for new users of dapagliflozin and new users of comparator ADs.

As part of the final (120-month) analysis, a quantitative bias analysis was conducted to assess the potential impact of unmeasured confounding. The findings of this analysis indicated that for the IRR comparing breast or bladder cancer between new users of dapagliflozin and new users of comparator AD, the association between a potential unmeasured confounder and the outcome would have to be large, as would the difference in the prevalence of the confounder across exposure groups, for the conclusions to change in a notable way. Some lifestyle characteristics are well-established risk factors for some cancers, and, whilst information on some lifestyle factors was available in our study for CPRD and for a subset of the PHARMO sample, potential unmeasured confounding may persist in our study including obesity for female breast cancer and smoking for bladder cancer, particularly in the data sources where this information was completely unavailable (the HIRD and Medicare). The impact of obesity on breast cancer risk appears to be specific to postmenopausal women in whom a moderate association with invasive breast cancer overall (meta-analysis RR = 1.18 [95% CI, 1.12-1.25]) has been observed, with stronger associations observed with some breast cancer phenotypes such as ER-positive tumours (meta-analysis RR = 1.39 [95% CI, 1.14-1.70]) [84]. Smoking is recognised as a strong risk factor for bladder cancer amongst both women and men, with studies reporting an increased bladder cancer risk of more than two or three times higher amongst smokers compared with nonsmokers, with the risk increasing with smoking intensity

and duration [71]. Although obesity and smoking are important potential confounders for breast cancer and bladder cancer, respectively, the likelihood of a large imbalance in the prevalence of these characteristics across the exposure groups is low and therefore unlikely to be masking a potential positive association between dapagliflozin and cancer risk. This suggests that there is likely minimal impact of unmeasured confounding in the present study.

11.2.3 Potential Outcome Misclassification

Potential misclassification of the outcome may occur if a primary cancer diagnosis is recorded as part of the initial diagnostic workup in those without confirmed cancers. If the presence of these diagnostic codes are then used to identify the occurrence of cancer outcomes, it may result in the inclusion of "suspected" or "rule out" diagnoses as cases in the study. In PHARMO, data were linked to the cancer registry, which serves as the "gold standard" for ascertainment of cancer cases, making outcome misclassification unlikely, although outcome misclassification might have occurred due to the probabilistic linkage. In CPRD, the HIRD, and Medicare, algorithms using primary care and hospital diagnosis data and claims data were used to identify each cancer outcome and were termed provisional cancer cases. In these data sources, to address the potential for outcome misclassification, a validation study of the primary outcomes was conducted during the 48-month and 96-month interim analyses. Findings from the case validation noted high PPVs (ranging from 81.5% in the HIRD to 99.0% in CPRD for female breast cancer and 88.2% in the HIRD to 97.8% in Medicare for sex-combined bladder cancer), suggesting that outcome misclassification was minimal (assuming non-differential sensitivity). However, validation was not done individually across exposure group. If misclassification is more likely to occur in one exposure group than another (eg, if clinicians were more likely to code for rule out diagnoses in the dapagliflozin-exposed patients than in the comparator AD–exposed patients), it might lead to differential misclassification of the outcome, which may make the therapy spuriously appear as if it is associated with an elevated cancer risk. The occurrence and magnitude are lessened when the PPVs are very high, as was noted in this study, as a high PPV suggests that most instances of a cancer outcome identified through coding algorithms were in fact true cases. Nonetheless, for the final analysis, a differential outcome misclassification simulation analysis was conducted to assess the worst-case scenario IRRs under different misclassification scenarios for the primary cancer outcomes (female breast cancer and bladder cancer) and data sources for which PPVs were estimated (CPRD, the HIRD, and Medicare). Using the observed PPVs and IRRs, this analysis suggested that under most misclassification scenarios (100% PPVs in the dapagliflozin group), the degree of misclassification would be minimal, and it appears unlikely in CPRD, the HIRD, and Medicare that a precise, meaningful elevated IRR could be masked due to differential misclassification of the breast cancer and bladder cancer outcome given the observed IRR and PPV estimates in this study.

11.2.4 Limited Precision for Some Subgroups

For some subgroups and data sources, in particular for the insulin user cohorts and pioglitazone user cohorts, and within the PHARMO data source, there was limited precision of the IRR point estimates, as indicated by wide 95% CIs, to make valid conclusions. This was attributed to the limited number of cancer cases and small sample sizes in these subgroups. Findings were interpreted in the context of these wide CIs, where applicable. Also, analyses to pool the propensity score–adjusted IRRs across the data sources were conducted for female breast cancer, sex-combined bladder cancer, and sex-specific bladder cancer to provide more precise estimates.

For the CPRD data source, the delay by the CPRD administration in releasing updated HES data to link to the latest available CPRD GOLD data did not allow for the final analysis to be performed in an updated data cut. Therefore, the same CPRD GOLD data and linked HES data set that were extracted for the fourth interim (96-month) analysis (data through March 2021) were used for the final (120-month) analysis.

11.3 Interpretation

Incidence rates of sex-combined bladder cancer in PHARMO were higher than in the general Dutch population aged ≥ 40 years, with the difference more pronounced for the comparator AD group. In the HIRD, incidence rates were approximately two-fold higher than in the general US population aged 40 to 64 years in both the dapagliflozin and comparator AD groups in this study. Differences in health care–seeking behaviour and potentially in lifestyle factors of relevance for bladder cancer—ie, smoking, alcohol use—between individuals that have medical care coverage (mostly through employer-sponsored health insurance) and the general population may have played a role in the higher rates of bladder cancer observed in the HIRD. The older mean age of the PHARMO study sample (age 64 and 65 years for dapagliflozin and AD comparator new users, respectively) and the potential higher risk of bladder cancer amongst patients with T2DM may have contributed to the higher bladder cancer incidence rates in the study population compared with the Dutch general population. In the sex-combined bladder cancer cohort in Medicare, rates were comparable in the dapagliflozin and comparator AD groups but were almost 50% lower than would be expected in the general US population aged ≥ 65 years. This may be attributed to the exclusion of patients with prior history of any cancer in the current study, which may have resulted in a study population that was healthier than the general population. In CPRD, rates of bladder cancer in the dapagliflozin and comparator AD groups were comparable to the UK general population.

Similarly, when comparing the incidence rates of new-use dapagliflozin compared with new-use comparator AD, unstratified and stratified by concomitant insulin use at the index date, the results also do not suggest an increased risk of any of the cancer outcomes assessed in this

study amongst dapagliflozin users, although some estimates are imprecise due to the small number of cancer events. These findings were confirmed in the pooled analysis and in sensitivity analyses where the varying censoring criteria were removed, highlighting the robustness of the findings in the main analyses. Additionally, the findings in the present study are consistent with other observational studies of dapagliflozin exposure and cancer risk [25; 120; 132]. A recent study using Optum's Clinformatics Data Mart Database, a large commercial claims database in the US, reported a HR of 0.96 (95% CI, 0.87-1.06) for breast cancer amongst women with T2DM initiating an SGLT2 inhibitor compared with those initiating a DPP-4 inhibitor [132]. In another population-based cohort study in Hong Kong, users of SGLT2 inhibitors had a lower risk of breast cancer (HR, 0.51; 95% CI, 0.32-0.80) compared with DPP-4 inhibitors; and, likewise, there was a similar finding when assessing dapagliflozin separately (HR, 0.48; 95% CI, 0.27-0.83) [25]. Another study, conducted in Japanese commercial insurance claims, reported a HR of 0.72 (95% CI, 0.42-1.24) for breast cancer when comparing users of an SGLT2 inhibitor to users of a DPP-4 inhibitor [120]. Similarly, a systematic literature review and meta-analysis of randomised controlled trials, which included > 100,000 patients, also noted no increased risk of breast cancer (RR, 0.99; 95% CI, 0.85-1.16) amongst users of SGLT2 inhibitors, including dapagliflozin [116].

The results of the current study did not find an increased risk of bladder cancer amongst patients initiating dapagliflozin compared with comparator ADs. This is similar to findings from other observational studies. For example, a cohort study from April 2013 to December 2018 conducted in data sets from three Scandinavian countries observed no increased risk of bladder cancer in new users of SGLT2 inhibitors compared with GLP-1 receptor agonists (adjusted HR, 0.88; 95% CI, 0.59-1.31) [127]. Similar findings were noted in a systematic literature review of randomised controlled trials, which reported no increased risk of bladder cancer for users of SGLT2 inhibitors (RR, 0.93; 95% CI, 0.71-1.21) nor for users of dapagliflozin (RR, 0.63; 95% CIs, 0.42-0.95) compared with placebo, active intervention, or no intervention [116].

For the bladder cancer cohorts, the analyses stratified by pioglitazone use at the index date did not suggest an increased risk of bladder cancer with dapagliflozin compared with comparator AD treatments amongst either pioglitazone users or pioglitazone non-users; however, the analyses conducted in the pioglitazone use cohorts are based on a small number of bladder cancer events. In PHARMO, these analyses could not be conducted because the prespecified criteria for number of outcome events was not met. Findings in the literature are inconsistent with regards to pioglitazone exposure and bladder cancer [37; 56; 74; 78]; however, in the current study, the assessment of pioglitazone use was to address the potential modifying effect of pioglitazone, which was inconclusive given the small proportions of the samples with concomitant pioglitazone use resulting in imprecise IRR estimates. In addition to conducting the pioglitazone use–stratified analyses, pioglitazone was included in the propensity score

models for all cohorts in which bladder cancer was evaluated as an outcome, to account for pioglitazone use as a potential important confounder.

Since the previous/fourth (96-month) analysis, the number of dapagliflozin new users in the full samples increased in the female breast cancer cohort and the sex-combined bladder cancer cohort, respectively, by approximately 28% and 59% in PHARMO, 59% and 63% in the HIRD, and 101% and 113% in Medicare. Relevant increases in dapagliflozin new users observed in PHARMO, the HIRD, and Medicare in this final 120-month analysis are likely linked to updates in the treatment guidelines for T2DM introduced in 2021 that recommend SGLT2 inhibitors, such as dapagliflozin, for patients with T2DM with established atherosclerotic cardiovascular disease or indicators of high risk, established kidney disease or heart failure [3; 50]. The distributions of the key baseline characteristics between the dapagliflozin new users and comparator AD new users in the propensity score–trimmed analysis populations across all study outcome cohorts showed adequate balance of covariates in CPRD, the HIRD, and Medicare despite the low number of cases identified in some of the cancer outcome cohorts. In PHARMO, small imbalances remained in a few select variables (eg, year of index date and dapagliflozin index monotherapy treatment). These are likely attributed to changes in diabetes management guidelines in the Netherlands that have shifted towards recommendation of SGLT2 inhibitors such as dapagliflozin as monotherapy as opposed to add-on third-line therapy, in particular for patients with established atherosclerotic cardiovascular disease or indicators of high risk, established kidney disease or heart failure [50]. This has resulted in an increased number of new users of dapagliflozin in more recent years of the study (with a majority of new use of the comparator AD medications in the earlier years of the study) and is consistent with trends observed in the general population in the Netherlands where SGLT2 inhibitor prescriptions increased from ~2 million in 2019 to 30 million in 2023 [93]. This difference in new users by calendar year of the index date likely explains the limited overlap of the propensity score distributions for dapagliflozin new users and comparator AD new users observed in PHARMO, where approximately 60% of dapagliflozin new users in the primary outcome cohorts initiated treatment from September 2021 to December 2022. In PHARMO, in the full samples, dapagliflozin new users identified in the last year of the study had different characteristics from new users identified in the early years of the study, which contributed to the relevant proportions of new users trimmed out of the propensity score−trimmed analysis samples. As a result, a lower average number of months of dapagliflozin exposure was observed in PHARMO across all cancer outcome cohorts.

Differences in HCRU or cancer screening patterns can introduce detection bias, whereby differences in timing of cancer diagnoses may exist between exposure groups, resulting in an imbalance in cancer cases that are not reflective of a true association between the exposure and the outcome. Therefore, the secondary objectives of the current study planned to analyse HCRU amongst new users of dapagliflozin and of comparator AD and to assess whether

HCRU may be a mediator in the association between initiation of dapagliflozin or comparator AD and cancer risk. The study noted that HCRU was similar across both exposure groups over the course of follow-up, and the need did not arise to conduct this analysis as the criteria were not met for any of the primary cancer outcomes in any data source and no findings in this study suggested an increased risk of cancer amongst dapagliflozin new users.

11.4 Generalisability

The CPRD population represented 4.5% of the UK population as a whole in January 2021, and the age and sex distribution was similar to that of the UK census population in 2011 [27]. The geographic distribution and size of practices in the CPRD may not be representative of all UK practices; therefore, results would not be generalisable to patient groups that are not contained in the CPRD, such as private patients, prisoners, the homeless, and patients in some residential homes [62].

The PHARMO Data Network combines primary (GP) (for a 40% subset of the population) and secondary (hospital) health care settings and covers approximately 20% of the entire population of the Netherlands, with median follow-up of approximately 10 years [96]. The age and sex distribution in the population included in the PHARMO Database Network is considered to be representative of the Dutch population, although PHARMO contains a larger population aged > 80 years than has been reported for the general Dutch population [75].

The HIRD comprises a geographically diverse spectrum of longitudinal claims data from health plan members from all regions of the US. As of July 2024, more than 91 million private US commercially insured individuals (largely employer-based insurance) with medical and pharmacy coverage were included in the HIRD [13]. Compared with the general population, coverage of the HIRD is skewed towards younger ages and those with a higher socioeconomic status. Moreover, the results may not be as generalisable to patient groups that are not contained in the HIRD data used for this study, such as patients aged 65 years or older (nearly all of whom are covered by Medicare), patients from other publicly insured programmes (ie, Medicaid), and uninsured populations. In particular, the current analysis is restricted to those aged 40 to 64 years; thus, the results from this data source cannot be generalised to older (or younger) populations.

Medicare provides near-universal health insurance coverage for individuals aged at least 65 years in the US, although the current analysis is restricted to those with Medicare fee-for-service plans, for which individual-level claims would be available, and those with Medicare Part D prescription coverage. Approximately 60% of Medicare enrolees opt for Part D prescription coverage [53]. The current study population did not include the uninsured, those who did not opt for Part D coverage, or those in Medicare Advantage managed care plans, which represents about half of all Medicare beneficiaries as of 2022 [72]) and are similar in

age, race/ethnicity, income level, and number of chronic conditions to those in Medicare fee-for-service [123].

Findings in this study related to patterns of use, patient characteristics, and relative risk estimates apply to patient populations initiating selected study AD medications in primary care practices in the UK, newly dispensed study AD medications in community pharmacies in the Netherlands, and patients covered by similar health insurance plans or through Medicare in the US.

12. OTHER INFORMATION

Not applicable.

13. CONCLUSION

This observational cohort study included 147,969 dapagliflozin new users and 1,545,926 comparator AD new users in the UK, the Netherlands, and the US. Results suggest that use of dapagliflozin does not appear to be associated with an increased risk of female breast cancer or bladder cancer compared with use of comparator AD. The sensitivity analyses were aimed at assessing the impact of various study design elements on the estimates for female breast cancer and bladder cancer and yielded results similar to those observed in the main analysis. After propensity score trimming and stratification, the dapagliflozin and comparator AD groups were comparable with regards to important sociodemographic and clinical variables. No notable difference in health care utilisation was observed following initiation of either dapagliflozin or comparator AD, with findings suggestive of no surveillance bias. Similar to the primary cancer outcomes, there did not appear to be an increased risk of the composite cancer outcome for females or males.

Overall, the results were consistent across the data sources and when stratified by insulin use or, in the bladder cancer cohort, when stratified by pioglitazone use. However, some analyses, particularly those amongst insulin and pioglitazone users, involved a small number of cancer cases in one or both exposure groups, and the precision of the effect estimates was low for some data sources.

To our knowledge, this is the first observational study evaluating dapagliflozin initiation and incidence of female breast cancer and bladder cancer, covering a study period of up to 10 years and a mean duration of follow-up amongst dapagliflozin new users ranging from 1.7 to 2.9 years across the primary cancer outcome cohorts. The findings from this study indicate that, compared with initiators of comparator ADs, there is no increased risk of female breast cancer or bladder cancer (sex-specific or sex-combined) in patients initiating dapagliflozin in real-world settings, including in patients initiating dapagliflozin with concomitant insulin use.

14. REFERENCES

1. Abrahami D, Tesfaye H, Yin H, Vine S, Hicks B, Yu OHY, et al. Sodium-glucose cotransporter 2 inhibitors and the short-term risk of bladder cancer: an international multisite cohort study. Diabetes Care 2022;45(12):2907-17.

2. Ahmadinezhad M, Arshadi M, Hesari E, Sharafoddin M, Azizi H, Khodamoradi F. The relationship between metabolic syndrome and its components with bladder cancer: a systematic review and meta-analysis of cohort studies. Epidemiol Health 2022;44(e2022050).

3. American Diabetes Association. Pharmacologic approaches to glycemic treatment: standards of medical care in diabetes—2021. Diabetes Care 2021;44(Suppl 1):S111–S24.

4. Ammann EM, Kalsekar I, Yoo A, Johnston SS. Validation of body mass index (BMI)-related ICD-9-CM and ICD-10-CM administrative diagnosis codes recorded in US claims data. Pharmacoepidemiol Drug Saf 2018;27(10):1092-100.

5. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature 2019;567(7748):305-7.

6. Arnold M, Freisling H, Stolzenberg-Solomon R, Kee F, O'Doherty MG, Ordonez-Mena JM, et al. Overweight duration in older adults and cancer risk: a study of cohorts in Europe and the United States. Eur J Epidemiol 2016;31(9):893-904.

7. AstraZeneca. Farxiga approved in the US for the treatment of heart failure in patients with heart failure with reduced ejection fraction. US Food and Drug Administration; 6 May 2020. Available at: https://www.astrazeneca.com/media-centre/press-releases/2020/farxiga-approved-in-the-us-for-the-treatment-of-heart-failure-in-patients-with-heart-failure-with-reduced-ejection-fraction.html#. Accessed 29 August 2025.

8. AstraZeneca. Farxiga approved in the US for the treatment of chronic kidney disease in patients at risk of progression with and without type-2 diabetes. 30 April 2021. Available at: https://www.astrazeneca.com/media-centre/press-releases/2021/farxiga-approved-in-the-us-for-ckd.html#. Accessed 29 August 2025.

9. AstraZeneca Pharmaceuticals LP. Observational study protocol MB102-118 ST/D1690R00007: Comparison of the risk of cancer between patients with type 2 diabetes exposed to dapagliflozin and those exposed to other antidiabetic treatments, version 3.1. 06 June 2023. Available at: https://catalogues.ema.europa.eu/sites/default/files/document_files/Dapa%20Cancer%20Protocol%20MB102118__v3.1_06Jun2023_final_signed_Redacted.pdf. Accessed 28 August 2025.

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

11. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011;46(3):399-424.

12. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34(28):3661-79.

13. Barron JJ, Willey VJ, Doherty BT, Tunceli O, Waltz CR, Grabner M, et al. The Healthcare Integrated Research Database (HIRD) as a real-world data source for pharmacoepidemiologic research. Pharmacoepidemiol Drug Saf 2025;34(2):e70110.

14. Bhaskaran K, Forbes HJ, Douglas I, Leon DA, Smeeth L. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). BMJ Open 2013;3(9):e003389.

15. Bogumil D, Cortessis VK, Wilkens LR, Le Marchand L, Haiman CA, Maskarinec G, et al. Interethnic differences in bladder cancer incidence and the association between type 2 diabetes and bladder cancer in the Multiethnic Cohort Study. Cancer Res Commun 2023;3(5):755-62.

16. Boyle P, Boniol M, Koechlin A, Robertson C, Valentini F, Coppens K, et al. Diabetes and breast cancer risk: a meta-analysis. Br J Cancer 2012;107(9):1608-17.

17. Bradley MC, Chillarige Y, Lee H, Wu X, Parulekar S, Wernecke M, et al. Similar breast cancer risk in women older than 65 years initiating glargine, detemir, and NPH insulins. Diabetes Care 2020;43(4):785-92.

18. Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: analytic strategies using propensity scores. Ann Intern Med 2002;137(8):693-5.

19. Brenner H, Gefeller O. Use of the positive predictive value to correct for disease misclassification in epidemiologic studies. Am J Epidemiol 1993;138(11):1007-15.

20. Bristol-Myers Squibb and AstraZeneca. Background document: dapagliflozin, BMS-512148, NDA 202293. Advisory Committee briefing materials. 4 November 2013. Available at: https://web.archive.org/web/20161024141328/http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM378079.pdf. Accessed 28 August 2025.

21. Bronsveld HK, ter Braak B, Karlstad O, Vestergaard P, Starup-Linde J, Bazelier MT, et al. Treatment with insulin (analogues) and breast cancer risk in diabetics; a systematic review and meta-analysis of in vitro, animal and human evidence. Breast Cancer Res 2015;17:100.

22. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. Am J Epidemiol 2006;163(12):1149-56.

23. CDC. National diabetes statistics report, 2020. Estimates of diabetes and its burden in the United States. Atlanta, GA: Centers for Disease Control and Prevention; 2020. Available at: https://diabetesresearch.org/wp-content/uploads/2022/05/national-diabetes-statistics-report-2020.pdf. Accessed 28 August 2025.

24. Chang HY, Weiner JP, Richards TM, Bleich SN, Segal JB. Validating the adapted Diabetes Complications Severity Index in claims data. Am J Manag Care 2012;18(11):721-6.

25. Chung CT, Lakhani I, Chou OHI, Lee TTL, Dee EC, Ng K, et al. Sodium-glucose cotransporter 2 inhibitors versus dipeptidyl peptidase 4 inhibitors on new-onset overall cancer in type 2 diabetes mellitus: a population-based study. Cancer Med 2023;12(11):12299-315.

26. CMS. CMS fast facts, March 2023 version. Centers for Medicare and Medicaid Services; 2023. Available at: https://data.cms.gov/sites/default/files/2023-03/CMSFastFactsMar2023.pdf. Accessed 28 August 2025.

27. CPRD. Release notes: CPRD GOLD January 2021. UK Medicines & Healthcare products Regulatory Agency; 2021. Available at: https://www.cprd.com/sites/default/files/2022-02/2021%-01%20CPRD%20GOLD%20Release%20Notes.pdf. Accessed 28 August 2025.

28. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17(19):2265-81.

29. De Pergola G, Silvestris F. Obesity as a major risk factor for cancer. J Obes 2013;2013:291546.

30. Deeks JJ, Higgins JPT, Altman DG, McKenzie JE, Veroniki AA, editors. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane handbook for systematic reviews of interventions, version 6.4. 2023. Available at: https://training.cochrane.org/handbook/current/chapter-10. Accessed 28 August 2025.

31. Dobson AJ, Kuulasmaa K, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. Stat Med 1991;10(3):457-62.

32. Dutch College of General Practitioners (NHG). Diagnostic conditions table. Dutch coding system. December 2024. Available at: https://bepalingen.nhg.org/labcodes/determinations. Accessed 28 August 2025.

33. ECIS. Incidence and mortality statistics, by registry and cancer site: Netherlands, bladder cancer. European Cancer Information System, European Commission; 2025. Available at: https://ecis.jrc.ec.europa.eu/. Accessed 28 August 2025.

34. ECIS. Incidence and mortality statistics, by registry and cancer site: Netherlands, female breast cancer. European Cancer Information System, European Commission; 2025. Available at: https://ecis.jrc.ec.europa.eu/. Accessed 28 August 2025.

35. ECIS. Incidence and mortality statistics, by registry and cancer site: United Kingdom, bladder cancer. European Cancer Information System, European Commission; 2025. Available at: https://ecis.jrc.ec.europa.eu/. Accessed 28 August 2025.

36. ECIS. Incidence and mortality statistics, by registry and cancer site: United Kingdom, female breast cancer. European Commission, European Cancer Information System; 2025. Available at: https://ecis.jrc.ec.europa.eu/. Accessed 28 August 2025.

37. Elshama SS, El-Kenawy Ael M, Osman HE. Toxicological evaluation of subchronic use of pioglitazone in mice. Iran J Basic Med Sci 2016;19(7):712-9.

38. EMA. Guideline on good pharmacovigilance practices (GVP). Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). European Medicines Agency; 28 July 2017. Available at: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf. Accessed 28 August 2025.

39. EMA. Guideline on good pharmacovigilance practices (GVP). Module VIII – Post-authorisation safety studies (EMA/813938/2011 Rev 3). European Medicines Agency; 09 October 2017. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf. Accessed 28 August 2025.

40. EMA. Qtrilmet (metformin hydrochloride/saxagliptin/dapagliflozin) [authorisation has been withdrawn]. European Medicines Agency; 2020. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/qtrilmet. Accessed 28 August 2025.

41. EMA. Summary of opinion (post authorisation): Forxiga (dapagliflozin) [EMA/CHMP/535224/2020]. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP); 15 October 2020. Available at: https://www.ema.europa.eu/en/documents/smop/chmp-post-authorisation-summary-positive-opinion-forxiga-ws-1737_en.pdf. Accessed 28 August 2025.

42. EMA. Summary of opinion (post authorisation): Forxiga (dapagliflozin) [EMA/CHMP/347888/2021]. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP); 24 June 2021. Available at: https://www.ema.europa.eu/en/documents/smop/chmp-post-authorisation-summary-positive-opinion-forxiga-ws-1941_en.pdf. Accessed 28 August 2025.

43. EMA. Assessment report, variation. Forxiga. EMA/954956/2022. 15 December 2022. Available at: https://www.ema.europa.eu/en/documents/variation-report/forxiga-h-c-ws-2299-epar-assessment-report-variation_en.pdf. Accessed 28 August 2025.

44. EMA. Forxiga: procedural steps taken and scientific information after the authorisation. 2024. Available at: https://www.ema.europa.eu/en/documents/procedural-steps-after/forxiga-epar-procedural-steps-taken-scientific-information-after-authorisation_en.pdf. Accessed 28 August 2025.

45. EMA. Wegovy (semaglutide) [authorised: this medicine is authorised for use in the European Union]. European Medicines Agency; 2025. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/wegovy#authorisation-details. Accessed 28 August 2025.

46. Englum BR, Prasad NK, Lake RE, Mayorga-Carlin M, Turner DJ, Siddiqui T, et al. Impact of the COVID-19 pandemic on diagnosis of new cancers: A national multicenter study of the Veterans Affairs Healthcare System. Cancer 2022;128(5):1048-56.

47. Farxiga PI. Farxiga (dapagliflozin) tablets, for oral use, prescribing information. US Food and Drug Administration; January 2014. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2014/202293s000lbl.pdf. Accessed 28 August 2025.

48. FDA. Best practices for conducting and reporting pharmacoepidemiologic safety studies using electronic healthcare data. Guidance for industry and FDA staff. US Department of Health and Human Services, Food and Drug Administration; May 2013. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic. Accessed 28 August 2025.

49. FDA. Qternmet XR, NDA 210874 [archived documents; product discontinued]. Food and Drug Administration; 2020. Available at: https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=210874. Accessed 28 August 2025.

50. Federation of Medical Specialists. Pharmacotherapy in very high-risk patients with type 2 diabetes mellitus (DM2) Federation of Medical Specialists; 15 November 2021. Available at: https://richtlijnendatabase.nl/richtlijn/medicamenteuze_behandeling_zeerhoogrisico_patienten_dm2/startpagina_-_farmacotherapie_bij_zeerhoogrisicopati_nten_met_diabetes_mellitus_type_2_dm2.html. Accessed 28 August 2025.

51. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013;49(6):1374-403.

52. Forxiga SmPC. Forxiga summary of product characteristics. European Medicines Agency; 29 October 2024. Available at: https://www.ema.europa.eu/en/documents/product-information/forxiga-epar-product-information_en.pdf. Accessed 28 August 2025.

53. Fung V, Brand RJ, Newhouse JP, Hsu J. Using Medicare data for comparative effectiveness research: opportunities and challenges. Am J Manag Care 2011;17(7):488-96.

54. Gallagher AM, Dedman D, Padmanabhan S, Leufkens HGM, de Vries F. The accuracy of date of death recording in the Clinical Practice Research Datalink GOLD database in England compared with the Office for National Statistics death registrations. Pharmacoepidemiol Drug Saf 2019;28(5):563-9.

55. Garcia-Esquinas E, Guino E, Castano-Vinyals G, Perez-Gomez B, Llorca J, Altzibar JM, et al. Association of diabetes and diabetes treatment with incidence of breast cancer. Acta Diabetol 2016;53(1):99-107.

56. Garry EM, Buse JB, Lund JL, Pate V, Sturmer T. Comparative safety of pioglitazone versus clinically meaningful treatment alternatives concerning the risk of bladder cancer in older US adults with type 2 diabetes. Diabetes Obes Metab 2018;20(1):129-40.

57. Glasheen WP, Renda A, Dong Y. Diabetes Complications Severity Index (DCSI)—update and ICD-10 translation. J Diabetes Complications 2017;31(6):1007-13.

58. Green KM, Stuart EA. Examining moderation analyses in propensity score methods: application to depression and substance use. J Consult Clin Psychol 2014;82(5):773-83.

59. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 2016;31(4):337-50.

60. Herings RMC, Pedersen L. Pharmacy-based medical record linkage systems. In: Strom BL, Kimmel SE, Hennessy S, editors. Pharmacoepidemiology. 5th ed. John Wiley & Sons, Ltd; 2012. p. 270-86.

61. Hernán M, Robins J. Causal inference: what if. Boca Raton: Chapman & Hall/CRC; 2020. Available at: https://static1.squarespace.com/static/675db8b0dd37046447128f5f/t/6835bdb7dc9e482163d8712d/1748352541538/hernanrobins_WhatIf_27may25.pdf. Accessed 28 August 2025.

62. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol 2015;44(3):827-36.

63. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60.

64. HMA-EMA Catalogue of RWD Studies. Comparison of the risk of cancer between patients with type 2 diabetes exposed to dapagliflozin and those exposed to other antidiabetic treatments. 13 December 2024. Available at:

https://catalogues.ema.europa.eu/node/3411/administrative-details. Accessed 28 August 2025.

65. Hong JL, Jonsson Funk M, Buse JB, Henderson LM, Lund JL, Pate V, et al. Comparative effect of initiating metformin versus sulfonylureas on breast cancer risk in older women. Epidemiology 2017;28(3):446-54.

66. Hossain FM, Danos DM, Fu Q, Wang X, Scribner RA, Chu ST, et al. Association of obesity and diabetes with the incidence of breast cancer in Louisiana. Am J Prev Med 2022;63(1 Suppl 1):S83–s92.

67. Hu Y, Zhang X, Ma Y, Yuan C, Wang M, Wu K, et al. Incident type 2 diabetes duration and cancer risk: a prospective study in two US cohorts. J Natl Cancer Inst 2021;113(4):381-9.

68. ICMJE. Preparing a manuscript for submission to a medical journal. 2. Reporting guidelines. International Committee of Medical Journal Editors; 2024. Available at: http://icmje.org/recommendations/browse/manuscript-preparation/preparing-for-submission.html#two. Accessed 28 August 2025.

69. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. International Society for Pharmacoepidemiology; June 2015. Available at: https://www.pharmacoepi.org/resources/policies/guidelines-08027/. Accessed 28 August 2025.

70. Jamal-Allial A, Sponholtz T, Vojjala SK, Paullin M, Papazian A, Eshete B, et al. Validation of mortality data sources compared to the National Death Index in the Healthcare Integrated Research Database. Pragmat Obs Res 2025;16:19-25.

71. Jubber I, Ong S, Bukavina L, Black PC, Comperat E, Kamat AM, et al. Epidemiology of bladder cancer in 2023: a systematic review of risk factors. Eur Urol 2023;84(2):176-90.

72. Kaiser Family Foundation (KFF). Medicare Advantage in 2022: enrollment update and key trends. 2022. Available at: https://www.kff.org/medicare/issue-brief/medicare-advantage-in-2022-enrollment-update-and-key-trends/. Accessed 28 August 2025.

73. Kaiser Family Foundation (KFF). Proposed coverage of anti-obesity drugs in Medicare and Medicaid would expand access to millions of people with obesity. 2024. Available at: https://www.kff.org/policy-watch/proposed-coverage-of-anti-obesity-drugs-in-medicare-and-medicaid-would-expand-access-to-millions-of-people-with-obesity/. Accessed 28 August 2025.

74. Korhonen P, Heintjes EM, Williams R, Hoti F, Christopher S, Majak M, et al. Pioglitazone use and risk of bladder cancer in patients with type 2 diabetes: retrospective cohort study using datasets from four European countries. BMJ 2016;354:i3903.

75. Kuiper JG, Bakker M, Penning-van Beest FJA, Herings RMC. Existing data sources for clinical epidemiology: the PHARMO Database Network. Clin Epidemiol 2020;12:415-22.

76. Lash TL, Fox MP, Fink AK. Unmeasured and unknown confounders. In: Applying quantitative bias analysis to epidemiologic data. Springer Science and Business Media, LLC; 2009. p. 59-78.

77. Lee DJ, Shelton JB, Brendel P, Doraiswami R, Makarov D, Meeks W, et al. Impact of the COVID-19 pandemic on urological care delivery in the United States. J Urol 2021;206(6):1469-79.

78. Li YR, Liu CH, Sun WC, Fan PY, Liu FH, Chen TH, et al. The risk of bladder cancer in type 2 diabetes mellitus with combination therapy of SGLT-2 inhibitors and pioglitazone. J Pers Med 2021;11(9).

79. Ling S, Brown K, Miksza JK, Howells L, Morrison A, Issa E, et al. Association of type 2 diabetes with cancer: a meta-analysis with bias analysis for unmeasured confounding in 151 cohorts comprising 32 million people. Diabetes Care 2020;43(9):2313-22.

80. Lu Y, Hajjar A, Cryns VL, Trentham-Dietz A, Gangnon RE, Heckman-Stoddard BM, et al. Breast cancer risk for women with diabetes and the impact of metformin: a meta-analysis. Cancer Med 2022.

81. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep 2015;2(4):221-8.

82. Mafi JN, Craff M, Vangala S, Pu T, Skinner D, Tabatabai-Yazdi C, et al. Trends in US ambulatory care patterns during the COVID-19 pandemic, 2019-2021. JAMA 2022;327(3):237-47.

83. Maskarinec G, Jacobs S, Park SY, Haiman CA, Setiawan VW, Wilkens LR, et al. Type II diabetes, obesity, and breast cancer risk: the Multiethnic Cohort. Cancer Epidemiol Biomarkers Prev 2017;26(6):854-61.

84. Munsell MF, Sprague BL, Berry DA, Chisholm G, Trentham-Dietz A. Body mass index and breast cancer risk according to postmenopausal estrogen-progestin use and hormone receptor status. Epidemiol Rev 2014;36(1):114-36.

85. Murff HJ, Roumie CL, Greevy RA, Hackstadt AJ, McGowan LED, Hung AM, et al. Metformin use and incidence cancer risk: evidence for a selective protective effect against liver cancer. Cancer Causes Control 2018;29(9):823-32.

86. National Cancer Institute. SEER* Explorer. Breast: cancer, SEER 5-year age-adjusted incidence rates, 2018-2022. Surveillance, Epidemiology, and End Results Program; 2025. Available at: https://seer.cancer.gov/statistics-

network/explorer/application.html?site=55&data_type=1&graph_type=3&compareBy=sex&chk_sex_3=3&rate_type=2&race=1&advopt_precision=1&advopt_show_ci=on&hdn_view=1#resultsRegion1. Accessed 28 August 2025.

87. National Cancer Institute. SEER*Explorer. Urinary bladder (invasive & in situ). SEER 5-year age-adjusted incidence rates, 2018-2022. Surveillance Epidemiology and End Results Program; 2025. Available at: https://seer.cancer.gov/statistics-network/explorer/application.html?site=71&data_type=1&graph_type=10&compareBy=sex&chk_sex_1=1&chk_sex_3=3&chk_sex_2=2&series=age_range&chk_age_range_157=157&chk_age_range_122=122&race=1&stage=101&advopt_precision=1&advopt_show_ci=on&hdn_view=1&advopt_show_apc=on&advopt_display=2#resultsRegion1. Accessed 29 August 2025.

88. National Cancer Institute. SEER*Stat software, version 9.0.40.0. Surveillance Epidemiology and End Results Program; 2025. Available at: www.seer.cancer.gov/seerstat. Accessed 29 August 2025.

89. *Nature* editorial. It's time to talk about ditching statistical significance. Nature 2019;567(7748):283.

90. NCR Data. Incidence by year, Crude Rate (CR), Bladder (age 65 years and older). Netherlands Comprehensive Cancer Organisation (IKNL); 2025. Available at: https://nkr-cijfers.iknl.nl/viewer/incidentie-per-jaar?language=en_GB&viewerId=1423a5e3-145f-43f8-896f-55664e035f5c. Accessed 28 August 2025.

91. NCR Data. Incidence by year, Crude Rate (CR), Stomach (excl. cardia) + Colon + Colon carcinoma + Rectum carcinoma + Lung / bronchus + Skin melanoma + Corpus uteri + Ovary and fallopian tube + Indolent non-Hodgkin lymphoma (females, age 45 years and older). Netherlands Comprehensive Cancer Organisation (IKNL); 2025. Available at: https://iknl.nl/international/ncr. Accessed 28 August 2025.

92. NCR Data. Incidence by year, Crude Rate (CR), Stomach (excl. cardia) + Colon + Colon carcinoma + Rectum carcinoma + Lung / bronchus + Skin melanoma + Prostate + Indolent non-Hodgkin lymphoma (males, age 65 years and older). Netherlands Comprehensive Cancer Organisation (IKNL); 2025. Available at: https://iknl.nl/international/ncr. Accessed 28 August 2025.

93. NHCI. Number of DDDs 2019-2023 for ATC subgroup A10BK: sodium glucose cotransporter-2 inhibitors (SGLT-2 inhibitors). National Health Care Institute; 26 June 2024. Available at: https://www.gipdatabank.nl/databank?infotype=g&label=00-totaal&tabel=B_01-basis&geg=ddd&item=A10BK. Accessed 28 August 2025.

94. NHG working group. NHG guidelines: type 2 diabetes mellitus, version 5.7. Dutch College of General Practitioners; December 2024. Available at: https://richtlijnen.nhg.org/standaarden/diabetes-mellitus-type-2. Accessed 28 August 2025.

95. Novo Nordisk. Wegovy™ demonstrated significant and sustained weight loss in two-year study in adults with obesity. 05 November 2021. Available at: https://www.novonordisk-us.com/media/news-archive/news-details.html?id=85378. Accessed 29 August 2025.

96. Overbeek JA, Swart KMA, Houben E, Penning-van Beest FJA, Herings RMC. Completeness and Representativeness of the PHARMO General Practitioner (GP) Data: A Comparison with National Statistics. Clin Epidemiol 2023;15:1-11.

97. Overbeek JA, Swart KMA, van der Pal EYM, Blom MT, Beulens JWJ, Nijpels G, et al. The DIAbetes MANagement and Treatment (DIAMANT) Cohort. Clin Epidemiol 2022;14:1453-62.

98. PALGA. National Pathology Registry in the Netherlands. Available at: https://www.palga.nl/. Accessed 28 August 2025.

99. Park YM, Bookwalter DB, O'Brien KM, Jackson CL, Weinberg CR, Sandler DP. A prospective study of type 2 diabetes, metformin use, and risk of breast cancer. Ann Oncol 2021;32(3):351-9.

100. Peila R, Rohan TE. Diabetes, glycated hemoglobin, and risk of cancer in the UK Biobank Study. Cancer Epidemiol Biomarkers Prev 2020;29(6):1107-19.

101. Perkins SM, Tu W, Underhill MG, Zhou XH, Murray MD. The use of propensity scores in pharmacoepidemiologic research. Pharmacoepidemiol Drug Saf 2000;9(2):93-101.

102. Piccoli GF, Mesquita LA, Stein C, Aziz M, Zoldan M, Degobi NAH, et al. Do GLP-1 receptor agonists increase the risk of breast cancer? A systematic review and meta-analysis. J Clin Endocrinol Metab 2021;106(3):912-21.

103. Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. CA Cancer J Clin 2017;67(5):378-97.

104. Qtern PI. Qtern (dapagliflozin and saxagliptin) tablets, for oral use, prescribing information. US Food and Drug Administration; February 2017. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/209091s000lbl.pdf. Accessed 28 August 2025.

105. Qtern SmPC. Qtern 5 mg/10 mg film-coated tablets, summary of product characteristics. European Medicines Agency; 25 November 2024. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/qtern. Accessed 28 August 2025.

106. Qternmet XR PI. Qternmet XR (dapagliflozin, saxagliptin, and metformin hydrochloride) extended-release tablets, for oral use, prescribing information. US Food and Drug Administration; May 2019. Available at:

https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/210874s000lbl.pdf. Accessed 28 August 2025.

107. Qtrilmet SmPC. Qtrilmet modified-release tablets, summary of product characteristics [product no longer authorised]. European Medicines Agency; 09 October 2020. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/qtrilmet. Accessed 28 August 2025.

108. Recalde M, Pistillo A, Davila-Batista V, Leitzmann M, Romieu I, Viallon V, et al. Longitudinal body mass index and cancer risk: a cohort study of 2.6 million Catalan adults. Nat Commun 2023;14(1):3816.

109. Research Data Assistance Center. Death information in the research identifiable Medicare data. 2022. Available at: https://resdac.org/articles/death-information-research-identifiable-medicare-data. Accessed 28 August 2025.

110. Rothman KJ, Greenland S. Planning study size based on precision rather than power. Epidemiology 2018;29(5):599-603.

111. Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008.

112. Rothman KJ, Lash TL. Precision and study size. In: Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ, editors. Modern epidemiology, 4th edition. Wolters Kluwer; 2021.

113. Sahai H, Khurshid A. Statistics in epidemiology: methods, techniques, and applications. Boca Raton, Florida: CRC Press LLC; 1996.

114. Sanderson M, Lipworth L, Shrubsole MJ, Andersen SW, Shu XO, Zheng W, et al. Diabetes, obesity, and subsequent risk of postmenopausal breast cancer among white and black women in the Southern Community Cohort Study. Cancer Causes Control 2019;30(5):425-33.

115. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, et al. Graphical depiction of longitudinal study designs in health care databases. Ann Intern Med 2019;170(6):398-406.

116. Spiazzi BF, Naibo RA, Wayerbacher LF, Piccoli GF, Farenzena LP, Londero TM, et al. Sodium-glucose cotransporter-2 inhibitors and cancer outcomes: A systematic review and meta-analysis of randomized controlled trials. Diabetes Res Clin Pract 2023;198:110621.

117. STIZON. Stichting Informatievoorziening voor Zorg en Onderzoek (Foundation for Information Provision for Care and Research). Utrecht, the Netherlands. Available at: https://stizon.nl/. Accessed 28 August 2025.

118. Suissa K, Schneeweiss S, Lin KJ, Brill G, Kim SC, Patorno E. Validation of obesity-related diagnosis codes in claims data. Diabetes Obes Metab 2021;23(12):2623-31.

119. Sun X, Zhang Q, Kadier K, Hu P, Liu X, Liu J, et al. Association between diabetes status and breast cancer in US adults: findings from the US National Health and Nutrition Examination Survey. Front Endocrinol (Lausanne) 2023;14:1059303.

120. Suzuki Y, Kaneko H, Okada A, Ko T, Jimba T, Fujiu K, et al. Association of SGLT2 inhibitors with incident cancer. Diabetes Metab 2024;50(6):101585.

121. Tang H, Dai Q, Shi W, Zhai S, Song Y, Han J. SGLT2 inhibitors and risk of cancer in type 2 diabetes: a systematic review and meta-analysis of randomised controlled trials. Diabetologia 2017;60(10):1862-72.

122. Tang H, Shi W, Fu S, Wang T, Zhai S, Song Y, et al. Pioglitazone and bladder cancer risk: a systematic review and meta-analysis. Cancer Med 2018;7(4):1070-80.

123. The Commonwealth Fund. Medicare Advantage vs traditional Medicare: how do beneficiaries' characteristics and experiences differ? 14 October 2021. Available at: https://www.commonwealthfund.org/publications/issue-briefs/2021/oct/medicare-advantage-vs-traditional-medicare-beneficiaries-differ. Accessed 28 August 2025.

124. Tipirneni R, Solway E, Malani P, Luster J, Kullgren JT, Kirch M, et al. Health insurance affordability concerns and health care avoidance among US adults approaching retirement. JAMA Netw Open 2020;3(2):e1920647.

125. Toh S, Garcia Rodriguez LA, Hernan MA. Analyzing partially missing confounder information in comparative effectiveness and safety research of therapeutics. Pharmacoepidemiol Drug Saf 2012;21 Suppl 2(0 2):13-20.

126. Tseng CH. Pioglitazone and breast cancer risk in female patients with type 2 diabetes mellitus: a retrospective cohort analysis. BMC Cancer 2022;22(1):559.

127. Ueda P, Svanstrom H, Hviid A, Eliasson B, Svensson AM, Franzen S, et al. Sodium-glucose cotransporter 2 inhibitors and risk of bladder and renal cancer: Scandinavian Cohort Study. Diabetes Care 2022;45(5):e93–e6.

128. US Census Bureau. American Community Survey. S0101: Age and sex. 2021: ACS 1-year estimates subject tables. 2021. Available at: https://data.census.gov/table?q=PEPAGE&t=Age+and+Sex&tid=ACSST1Y2021.S0101. Accessed 28 August 2025.

129. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013;18(2):137-50.

130. van den Brandt PA. Diabetes and the risk of bladder cancer subtypes in men and women: results from the Netherlands Cohort Study. Eur J Epidemiol 2024;39(4):379-91.

131. van Herk-Sukel MP, van de Poll-Franse LV, Lemmens VE, Vreugdenhil G, Pruijt JF, Coebergh JW, et al. New opportunities for drug outcomes research in cancer patients: the linkage of the Eindhoven Cancer Registry and the PHARMO Record Linkage System. Eur J Cancer 2010;46(2):395-404.

132. Wang F, Hendryx M, Liu N, Bidulescu A, Mitra AK, Luo J. SGLT2 inhibitor use and risk of breast cancer among adult women with type 2 diabetes. Drug Saf 2024;47(2):125-33.

133. Wang T, Chai B, Chen WY, Holmes MD, Erdrich J, Hu FB, et al. Metformin and other anti-diabetic medication use and breast cancer incidence in the Nurses' Health Studies. Int J Cancer 2024;155(2):211-25.

134. Wasserstein RL, Lazar NA. The ASA statement on *P* values: context, process, and purpose. Am Stat 2016;70(2):129-33.

135. Wirtz HS, Calip GS, Buist DSM, Gralow JR, Barlow WE, Gray S, et al. Evidence for detection bias by medication use in a cohort study of breast cancer survivors. Am J Epidemiol 2017;185(8):661-72.

136. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. N Engl J Med 2018: doi: 10.1056/NEJMoa1812389.

137. Wu JW, Azoulay L, Majdan A, Boivin JF, Pollak M, Suissa S. Long-term use of long-acting insulin analogs and breast cancer incidence in women with type 2 diabetes. J Clin Oncol 2017;35(32):3647-53.

138. Xie J, Liu C. Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. Stat Med 2005;24(20):3089-110.

139. Xigduo SmPC. Xigduo summary of product characteristics. European Medicines Agency; 25 November 2024. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/xigduo. Accessed 28 August 2025.

140. Xigduo XR PI. Xigduo XR (dapagliflozin and metformin HCl extended-release) tablets, for oral use, prescribing information. US Food and Drug Administration; October 2014. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2014/205649s000lbl.pdf. Accessed 28 August 2025.

141. Xiong F, Dai Q, Zhang S, Bent S, Tahir P, Van Blarigan EL, et al. Diabetes and incidence of breast cancer and its molecular subtypes: A systematic review and meta-analysis. Diabetes Metab Res Rev 2024;40(1):e3709.

142. Xiong F, Wang J, Nierenberg JL, Van Blarigan EL, Kenfield SA, Chan JM, et al. Diabetes mellitus and risk of breast cancer: a large-scale, prospective, population-based study. Br J Cancer 2023;129(4):648-55.

143. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernan MA. A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med 2020;39(8):1199-236.

144. Zhang F, de Haan-Du J, Sidorenkov G, Landman GWD, Jalving M, Zhang Q, et al. Type 2 diabetes mellitus and clinicopathological tumor characteristics in women diagnosed with breast cancer: a systematic review and meta-analysis. Cancers (Basel) 2021;13(19).

145. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159(7):702-6.

Appendix A List of Stand-Alone Documents

Table 27 List of Stand-Alone Documents

Number	Document reference number	Date	Title
1	MB102-118 ST Version 3.1	06 June 2023	Final study protocol
2	Version 2.0	28 June 2018	Validation plan, HIRD
3	Version 2.0	27 July 2018	Validation plan, CPRD and Medicare
4	Version 3.0	22 April 2020	Statistical analysis plan
5	Version 2.1	22 September 2020	Data analytic plan
6	Version 1.0	16 May 2024	Sensitivity analysis plan
7	Version 1.0	20 November 2024	Pooled analysis plan

Appendix B List of Relevant Antidiabetic Drugs

Table 28 Antidiabetic Drugs Eligible for Inclusion in the Comparator Group

Blood glucose–lowering drugs (excluding insulin) by ATC subgroup	Active substance
A10BA, Biguanides[a]	Metformin
A10BB, Sulfonylureas[a]	Glibenclamide/glyburide
	Tolbutamide
	Glibornuride
	Gliclazide
	Glimepiride
	Carbutamide
	Chlorpropamide
	Tolazamide
	Glipizide
	Gliquidone
	Acetohexamide
	Glisoxepide
A10BC, Sulfonamides (heterocyclic)	Glymidine
A10BD, Combinations	Metformin/sulfonylureas
	Metformin/rosiglitazone
	Rosiglitazone/glimepiride
	Pioglitazone/metformin
	Pioglitazone/glimepiride
	Sitagliptin/metformin
	Vildagliptin/metformin
	Pioglitazone/alogliptin
	Metformin/saxagliptin
	Metformin/linagliptin
	Pioglitazone/sitagliptin
	Metformin/alogliptin
	Metformin/repaglinide
	Metformin/acarbose
	Metformin/gemigliptin
A10BF, Alpha-glucosidase inhibitors	Acarbose
	Voglibose
	Miglitol

Blood glucose–lowering drugs (excluding insulin) by ATC subgroup	Active substance
A10BG, Thiazolidinediones	Pioglitazone
	Rosiglitazone
A10BH, DPP-4 (dipeptidyl peptidase-4) inhibitors	Sitagliptin
	Vildagliptin
	Saxagliptin
	Linagliptin
	Alogliptin
	Sitagliptin/simvastatin
A10BH, DPP-4 inhibitor combinations	Alogliptin/metformin
	Linagliptin/metformin
	Saxagliptin/metformin
A10BJ Glucagon-like peptide-1 (GLP-1) receptor agonists	Exenatide
	Liraglutide
	Lixisenatide
	Albiglutide
	Dulaglutide
	Semaglutide
A10BX, Other blood glucose–lowering drugs, excl. insulins	Repaglinide[b]
	Nateglinide[b]
	Mitiglinide[b]
	Tirzepatide[c]
	Pramlintide[d]
	Bromocriptine[e]

[a] Drugs in this class will qualify as comparator antidiabetic drugs only if prescribed in combination with other antidiabetic drugs.

[b] Repaglinide, nateglinide, and mitiglinide are in the meglitinide class.

[c] Tirzepatide is a dual GIP/GLP-1 receptor co-agonist.

[d] Pramlintide is an amylin mimetic.

[e] Low-strength (eg, 0.8 mg) bromocriptine, a dopamine-2 agonist, is indicated for the treatment of T2DM and is classified in ATC G02CB01.

Source: World Health Organization Collaborating Centre for Drug Statistics Methodology. ATC/DDD index 2024. Available at: http://www.whocc.no/atc_ddd_index/. Accessed 04 January 2024.

Appendix C List of Codes for Dapagliflozin

Table 29 Codes for Dapagliflozin Products in CPRD

Product code	Gemscript code	Product name
63031	23008021	Dapagliflozin 5 mg/metformin 850 mg tablets
60012	23006021	Dapagliflozin 5 mg/metformin 1 g tablets
54265	46673020	Dapagliflozin 5 mg tablets
54182	46671020	Dapagliflozin 10 mg tablets
54480	46674020	Forxiga 5 mg tablets (AstraZeneca UK Ltd)
54203	46672020	Forxiga 10 mg tablets (AstraZeneca UK Ltd)
63516	40763021	Forxiga 10 mg tablets (Waymade Healthcare Plc)
84232	86575021	Forxiga 10 mg tablets (Pharmaram Ltd)
60643	23007021	Xigduo 5 mg/1,000 mg tablets (AstraZeneca UK Ltd)
65059	23009021	Xigduo 5 mg/850 mg tablets (AstraZeneca UK Metformin hydrochloride/Dapagliflozin Ltd)
69540	73838021	Saxagliptin 5 mg/Dapagliflozin 10 mg Saxagliptin hydrochloride/Dapagliflozin tablets propanediol monohydrate
69654	73839021	Qtern 5 mg/10 mg tablets (AstraZeneca UK Saxagliptin hydrochloride/Dapagliflozin Ltd) propanediol monohydrate
87570	88733021	Forxiga 10 mg tablets (CST Pharma Ltd)

Table 30 Codes for Dapagliflozin Products in PHARMO

ATC code	Description
A10BX09/A10BK01	Dapagliflozin
A10BD15	Dapagliflozin + metformin
A10BD21	Saxagliptin + dapagliflozin
A10BD25	Metformin + saxagliptin + dapagliflozin

Table 31 Codes for Dapagliflozin Products in the HIRD and Medicare

NDC	GPI name	Product name
00310620530	Dapagliflozin propanediol tab 5 mg (base equivalent)	Farxiga
00003142711	Dapagliflozin propanediol tab 5 mg (base equivalent)	Farxiga
00310620595	Dapagliflozin propanediol tab 5 mg (base equivalent)	Farxiga
50090348200 [a]	Dapagliflozin propanediol tab 5 mg (base equivalent)	Farxiga
00003142811	Dapagliflozin propanediol tab 10 mg (base equivalent)	Farxiga
00310621030	Dapagliflozin propanediol tab 10 mg (base equivalent)	Farxiga
00310621095	Dapagliflozin propanediol tab 10 mg (base equivalent)	Farxiga
50090348100 [a]	Dapagliflozin propanediol tab 10 mg (base equivalent)	Farxiga
00310622560	Dapagliflozin-metformin HCl tab ER 24HR 2.5-1000 mg	Xigduo XR
00310625030	Dapagliflozin-metformin HCl tab ER 24HR 5-500 mg	Xigduo XR
00310626060	Dapagliflozin-metformin HCl tab ER 24HR 5-1000 mg	Xigduo XR
00310626030	Dapagliflozin-metformin HCl tab SR 24HR 5-1000 mg	Xigduo XR
00310627030	Dapagliflozin-metformin HCl tab ER 24HR 10-500 mg	Xigduo XR
00310628030	Dapagliflozin-metformin HCl tab ER 24HR 10-1000 mg	Xigduo XR
00310677030	Dapagliflozin-saxagliptin HCl tab 5-5 mg	Qtern
00310678030	Dapagliflozin-saxagliptin HCl tab 10-5 mg	Qtern
55154693208	Dapagliflozin propanediol	Farxiga
55154693308	Dapagliflozin propanediol	Farxiga
00310621039	Not applicable	Farxiga
00310627095	Dapagliflozin and metformin hydrochloride	Xigduo
00310626095	Dapagliflozin and metformin hydrochloride	Xigduo
00310628095	Dapagliflozin and metformin hydrochloride	Xigduo

[a] Manufactured by A-S Medication Solutions.

Note: AstraZeneca is the manufacturer of all these products unless otherwise indicated.

24HR = 24 hours; ER = extended release; GPI = Generic Product Identifier; NDC = National Drug Code; SR = sustained release; XR = extended release.

Appendix D Electronic Algorithms for Identifying Cancer Outcomes

Table 32 Electronic Algorithms for Identifying Provisional Cases of the Cancer Outcomes, by Data Source

Data source	Algorithm
	Female breast cancer (female patients only)
CPRD	• At least one cancer Read code for invasive breast cancer[a] in CPRD GOLD data. OR • A hospital record in HES with a diagnosis code (ICD-10 C50.*) for invasive breast cancer.
PHARMO	An ICD-O-3 diagnosis code for invasive breast cancer (ICD-O-3 topography code C50.*) as registered in the Pathology Registry and Netherlands Cancer Registry.[b]
The HIRD and Medicare	• At least two invasive breast cancer diagnosis codes (ICD-9-CM 174.* or ICD-10-CM C50.*)[c] reported as part of a medically attended inpatient, emergency department, outpatient, or physician visit (CPT codes 992xx to 994xx). • The two invasive breast cancer diagnosis codes must be reported on different dates and within 60 days of each other; the outcome index date is defined as the date of the first cancer diagnosis.
	Bladder cancer
CPRD	• At least one cancer Read code in CPRD GOLD data for in situ or invasive bladder cancer[d]. OR • A hospital record in HES with a diagnosis code (ICD-10 D09.0 or C67.*) for in situ or invasive bladder cancer.
PHARMO	An ICD-O-3 diagnosis code for bladder cancer (ICD-O-3 topography code C67.*) as registered in the Pathology Registry and Netherlands Cancer Registry.[b]
The HIRD and Medicare	• At least two in situ or invasive bladder cancer diagnosis codes (ICD-9-CM 233.7 or 188.* or ICD-10-CM D09.0 or C67.*)[c] reported as part of a medically attended inpatient, emergency department, outpatient, or physician visit (CPT codes 992xx to 994xx). • The two in situ or invasive bladder cancer diagnosis codes must be reported on different dates and within 60 days of each other; the outcome index date is defined as the date of the first cancer diagnosis.

[a] Read codes for invasive breast cancer: B34.00, B34.11, B340.00, B340000, B340100, B340z00, B341.00 thru B347.00, B34y.00, B34y000, B34yz00, B34z.00, Byu6.00, BB94.00, BB94.11, BB9J.11, BB9K.00, BB9K000, BB5K.00, BB91000, BB91.00, BB9G.00, BB91100, BB93.00, BB9F.00, BB9H.00, BB9J.00, BBM9.00.

[b] In PHARMO, cases were identified through the Netherlands Cancer Registry for the final 120-month data cut.

[c] ICD-10-CM codes used for diagnoses on or after 01 October 2015.

[d] Read codes for in situ or invasive bladder cancer: B49y.00, B49y000, B49z.00, B837.00, BB4.00, BB43.00, BB43.11, BB47.00, BB4A.00, BB4B.00, BB4C.00, BB4D.00.

Appendix E Study Covariates: Medical Conditions, Medications, Lifestyle, and Health Care Resource Utilisation, as Available in Each Data Source

Table 33 Covariate Medical Conditions, by Outcome

Condition	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Composite cancers
Medical conditions for exclusion						
Type 1 diabetes mellitus	During all available time up to and including index date	0 = Absent 1 = Present	Cohort exclusion indicator variable	√	√	√
Any invasive cancer diagnosis (except NMSC)	During all available time up to and including index date	0 = Absent 1 = Present	Cohort exclusion indicator variable	√	√	√
In situ bladder cancer	During all available time up to and including index date	0 = Absent 1 = Present	Cohort exclusion indicator variable	—	√	—
Medical conditions/comorbidities						
Benign mammary dysplasia (females only)	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	—	—
Hereditary nonpolyposis colon cancer syndrome (Lynch syndrome) or genetic susceptibility to malignant neoplasm	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	√	√
Urinary infections (chronic or recurrent)	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	√	—
Urinary cystitis (chronic or recurrent)	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	√	—
Kidney stones	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	√	—

Condition	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Composite cancers
Bladder stones	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	√	—
Benign prostatic hyperplasia (males only)	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Polycystic ovarian syndrome (females only)	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Colon polyps, including familial adenomatous polyposis	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Crohn's disease	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Ulcerative colitis	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Immunosuppressive diseases (AIDS/HIV)	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Peptic ulcer disease	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Helicobacter pylori infection	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Autoimmune diseases (rheumatoid arthritis, systemic lupus erythematosus, Sjögren disease, celiac sprue [gluten-sensitive enteropathy])	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	—	—	√
Heart failure	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Chronic kidney disease	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√

Condition	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Composite cancers
Type 2 diabetes mellitus	During all available time up to and including index date	0 = Absent 1 = Present	To describe proportion of patients with a recorded previous diagnosis of T2DM	—	—	√
Indicators of diabetes severity						
Diabetic nephropathy or renal insufficiency	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Retinopathy	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Peripheral neuropathy	During all available time up to and including index date	0 = Absent 1 = Present	Potential, confounder	√	√	√
Peripheral vascular disease	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Coronary heart disease	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Cerebrovascular disease	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Amputation	During all available time up to and including index date	0 = Absent 1 = Present	Potential confounder	√	√	√
Time since first recorded diagnosis of T2DM (years), if available (CPRD and a subset from PHARMO only)	During all available time up to and including index date	Number of years, mean, median, IQR, SD, min, max	Potential confounder	√	√	√

Condition	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Composite cancers
HbA1c, % (mmol/mol) (CPRD and a subset from PHARMO)[a]	The most recent recorded laboratory value in the 180 days before the index date	• < 7.0 (< 53) • 7.0-10.0 (53-86) • > 10.0 (> 86) • Unknown	Potential confounder	√	√	√
HbA1c (the HIRD and Medicare)	Number of HbA1c measurements in the 180 days before the index date	Number of measurements [categorise based on frequency distribution]	Potential confounder	√	√	√

[a] Conversion from HbA1c (%) to HbA1c (mmol/mol): 7.0% = 53 mmol/mol; 10.0% = 86 mmol/mol.

IQR = interquartile range; max = maximum; min = minimum; NMSC = non-melanoma skin cancer.

Table 34 Covariate Medications/Procedures, by Outcome

Medication category	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Other cancers
SGLT2 inhibitor other than dapagliflozin	During all available time up to and including index date	0 = Absent 1 = Present	Cohort exclusion	√	√	√
Cystoscopy	During the 180 days up to and including the index date	0 = Absent 1 = Present	Cohort exclusion	—	√	—
Bladder biopsy	During the 180 days up to and including the index date	0 = Absent 1 = Present	Cohort exclusion	—	√	—
Urine cytology	During the 180 days up to and including the index date	0 = Absent 1 = Present	Cohort exclusion	—	√	—
Breast biopsy (females only)	During the 180 days up to and including the index date	0 = Absent 1 = Present	Cohort exclusion	√	—	—
Insulin	Current use at index date	0 = Absent 1 = Present	Indicator variable for stratification; propensity score (included in propensity score models for cohorts not stratified by insulin use)	√	√	√
Pioglitazone	Current use at index date	0 = Absent 1 = Present	Indicator variable for stratification; propensity score (included in propensity score models for cohorts not stratified by pioglitazone use)	—	√	—
Antidiabetic drug use at baseline	During the 90 days before the index date, on the index date, and during the 90 days after the index date		Propensity score	√	√	√
Monotherapy		0 = No, 1 = Yes				

Medication category	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Other cancers
Index combination therapy		0 = No, 1 = Yes				
Add-on therapy		0 = No, 1 = Yes				
Switched-to therapy		0 = No, 1 = Yes				
Add-on plus switched-to therapy		0 = No, 1 = Yes				
Non-evaluable		0 = No, 1 = Yes				
Prior use of antidiabetic drugs[a]	• Within the 12 months before index date • More than 12 months through 24 months before index date • More than 24 months before index date	Number of drug classes[b:] 0, 1-2, 3-4, 5-8, not applicable[c]	Propensity score	√	√	√
Combined oestrogen-progesterone hormone-replacement therapy (females only)	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	√	—	√
Selective oestrogen receptor modulators (raloxifene, tamoxifen) (females only)	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	√	—	√
Cyclophosphamide	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	—	√	—
Unopposed oestrogen therapy (females only)	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	—	—	√
Immunosuppressant (including systemic steroids, excluding inhaled corticosteroids)	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	—	—	√
Inhaled corticosteroids	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	—	—	√

Medication category	Evaluation period	Coding	Purpose in analyses	Breast cancer	Bladder cancer	Other cancers
Opioids	During 180 days before and including the index date	0 = Absent 1 = Present	Propensity score	√	√	√

[a] Separate variables were created for each time window.

[b] Antidiabetic drug classes that were considered: insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

[c] "Not applicable" was assigned for the evaluation periods "More than 12 months through 24 months before index date" or "More than 24 months before index date" if the patient did not have enough lookback time in their medical records.

Table 35 Demographic, Lifestyle, and Health Care Resource Utilisation Variables for Use With All Cancer Outcomes

Variables	Evaluation period	Coding
Demographic and lifestyle variables		
Body mass index[a] (using recorded BMI variable or data on height and weight; CPRD and GP subset in PHARMO only)	During all available history (including index date); use the most recent value recorded	1 = < 20 (underweight) 2 = 20 to < 25 (normal) 3 = 25 to < 30 (overweight) 4 = 30 to < 40 (obese) 5 = ≥ 40 (severely obese) 6 = Unknown/missing
Overweight	During all available history (including index date)	0 = Absent 1 = Present 2 = Unknown
Smoking history (CPRD and GP subset in PHARMO only)	During all available history (including index date); use the most recent value recorded[b]	1 = Current 2 = Former 3 = Never 4 = Unknown/missing
Alcohol use (CPRD and GP subset in PHARMO only)	During all available history (including index date); use the most recent value recorded	1 = Non-drinker 2 = Low to moderate intake (1 to 6 units/week) 3 = Heavy or very heavy intake (≥ 7 units/week) 4 = Drinker, unknown quantity 5 = Unknown/missing
History of alcohol abuse	During all available history (including index date)	0 = Absent 1 = Present
Duration of lookback time	Number of days from cohort entry date up to and including the index date	Continuous variable with a minimum of 180 days; may be categorised after examining frequency distribution

Variables	Evaluation period	Coding		
Socioeconomic status (SES)				
Index of multiple socioeconomic deprivation (CPRD only)	During all available history (including index date); use the most recent value recorded	Quintiles: Q1 (least deprived) through Q5 (most deprived)		
SES score based on postal code of residence (in PHARMO only)		1 = Low 2 = Middle 3 = High 4 = Unknown		
Low-income subsidy status (Medicare only)	At index date	0 = No 1 = Yes		
Race/ethnicity (CPRD, the HIRD, and Medicare)	At index date	<u>CPRD</u> 1 = Asian or Asian British 2 = Black or Black British 3 = White 4 = Mixed 5 = Other ethnic group 6 = Unknown / missing	<u>HIRD</u> 1 = American Indian or Alaska Native, non-Hispanic or Latino 2 = Asian, non-Hispanic or Latino 3 = Black or African American, non-Hispanic or Latino 4 = Hispanic or Latino, any race 5 = Native Hawaiian or Other Pacific Islander, non-Hispanic or Latino 6 = White, non-Hispanic or Latino 7 = Other Race, non-Hispanic or Latino 8 = Unknown or undisclosed	<u>Medicare</u> 1 = Asian 2 = Black 3 = Hispanic 4 = White 5 = Other 6 = Unknown
Age (in years)	At index date	Descriptive analyses include age categories for 40-54 years, 55-64 years, 65-74 years, 75-84 years, and ≥ 85 years. Propensity score models: age modelled as a continuous variable.		

Variables	Evaluation period	Coding
Calendar year (time period) of the index date	At index date	Categories generated based on the following: • For index dates occurring before 2020: separate category for each calendar year from start of study period to 2019. • For index dates occurring starting in 2020: one category for January 2020 through February 2020; starting with March 2020 (ie, COVID-19 era), single category for each 6-month period until the end of study period; the most recent category may include a period of < 6 months or > 6 months. Propensity score models: if needed, some COVID-19 era time period categories were collapsed for propensity score modelling due to small numbers.
Primary care practice/geographic region or outpatient pharmacy region	At index date	CPRD: England, Wales, Scotland, Northern Ireland PHARMO: County: Drenthe, Flevoland, Friesland, Gelderland, Groningen, Limburg, Noord-Brabant, Noord-Holland, Overijssel, Utrecht, Zeeland, Zuid-Holland Medicare: Midwest, Northeast, South, West HIRD: Midwest, Northeast, South, West
Health care resource utilisation variables		
Outpatient visits (to GP or outpatient hospital clinic)	In the 180 days before but not including the index date	Categories for all health care resource utilisation variables will be data driven depending on the distribution of this variable at each data source
Hospitalisations	In the 180 days before but not including the index date	
Emergency department visits (not available in PHARMO)	In the 180 days before but not including the index date	
Speciality care visits (not available in PHARMO)	In the 180 days before but not including the index date	

[a] Possible categories for BMI are listed. The distribution of BMI values will be evaluated to determine whether it is necessary to collapse categories because of sparse data.

[b] In the CPRD, BMI was calculated based on the most recent BMI value or the height and weight values recorded closest to the index date within up to 5 years before the index date.

Q*n* = quintile.

Appendix F Characteristics of the Data Sources

Table 36 Selected Characteristics of the Data Sources and Variables of Interest

Characteristics	CPRD	PHARMO	HIRD	Medicare
Database population	As of January 2021, approximately 19.5 million total patients with research quality data, including 3,000,000 active patients	More than 4 million residents (~20%) of a well-defined population in Netherlands for an average of 12 years. Currently, the PHARMO Data Network covers over 7 million active persons	As of July 2024, 91 million unique individuals with medical coverage (24 million of whom are actively enrolled) and more than 72 million lives with medical and pharmacy claims information in the Northeastern, Mid-Atlantic, Southeastern, Midwestern, Central, and Western regions of the US	As of 2021, 55,900,000 enrolees aged 65 years or older from all US states (which accounts for ~100% of the US population aged 65 years or older); 36,400,000 with original Medicare (fee-for-service insurance)
Database type	Primary health care, GP electronic medical record database plus linkage to hospital data through the Hospital Episode Statistics (HES) database for about 44% of patients eligible for linkage as of April 2023	The PHARMO Data Network of health care data sources linked on a patient level through validated algorithms. For this study, outpatient pharmacy data, hospital data, clinical laboratory data, general practitioner data, Pathology Registry, and the Netherlands Cancer Registry were used.	Healthcare Integrated Research Database Insurance claims database	Federally funded insurance claims database; Part A (hospital insurance), Part B (physician services and outpatient care), Part D (outpatient prescription drug coverage)
Data source lag times	CPRD GOLD and HES-linked data: 6 months Office for National Statistics (mortality data): 6 months Cancer register; due to the 2-year lag time, there are currently no plans to use this linkage	Data sources linked annually (average lag time is 1 year)	Medical claims: 3 months Paid pharmacy claims: 30 days	Parts A, B, and D data: up to 24 months
Exposures				
Prescribed/ dispensed drugs	Defined by GP prescriptions, no inpatient data	Defined by medication dispensings	Defined by medication dispensing claims	Defined by medication dispensing claims, no inpatient data

Characteristics	CPRD	PHARMO	HIRD	Medicare
Study variables				
Lookback period	The earliest date is the latter of the practice up-to-standard date and the patient registration date	The earliest date is the first enrolment date in the database	The earliest date is the first enrolment date in a health plan for the continuous period of enrolment that overlaps the index date	The earliest date is the date that a patient has enrolment in Medicare Parts A, B, and D for the continuous period of enrolment that overlaps the index date
Medical conditions, medications, and lifestyle variables	See details on identification and coding in Section 9.4.3, Section 9.8, and Appendix E.	See details on identification and coding in Section 9.4.3, Section 9.8, and 0.	See details on identification and coding in Section 9.4.3, Section 9.8, and 0.	See details on identification and coding in Section 9.4.3, Section 9.8, and 0.
Other variables	Socioeconomic status: Index of Multiple Deprivation Race/ethnicity available at the time of the final analysis Alcohol use available Body mass index (using recorded BMI variable or data on height and weight) Smoking history	Socioeconomic status: based on postal code, derived from the Netherlands Institute for Social Research Race/ethnicity not available Alcohol use available for a subset of patients from the GP database BMI (using recorded BMI variable or data on height and weight) available for a subset of patients from the GP database Smoking history available for a subset of patients from the GP database	Socioeconomic status not available at the time of this analysis Race/ethnicity available at the time of the final analysis Alcohol use available on a subset of members of the HIRD BMI available on a subset of members of the HIRD Smoking history available on a subset of members of the HIRD	Socioeconomic status: Medicare low-income subsidy status Race/ethnicity available Alcohol use not available BMI not available Smoking history not available
Laboratory data	Data available	Available on a subset of patients from the GP Database/Clinical Laboratory Database	Available for a subset of patients (approximately 30%) but not used in this analysis	Not available
Specific medical conditions	Time since first diagnosis of T2DM HbA1c, % (the most recent recorded laboratory value in the 180 days before the index date)	Subset of PHARMO (GP data): Time since first recorded diagnosis of T2DM HbA1c, % (the most recent recorded laboratory value in the 180 days before the index date)	HbA1c test (number of HbA1c measurements in the 180 days before the index date)	HbA1c test (number of HbA1c measurements in the 180 days before the index date)

Characteristics	CPRD	PHARMO	HIRD	Medicare
Study endpoints				
Primary and secondary cancer outcomes	Read diagnosis codes in the CPRD GOLD ICD-10 diagnosis codes from HES (hospital outpatient and inpatient data) General practice electronic medical records to identify diagnoses of targeted malignancies HES data to identify hospital inpatient discharge diagnoses of targeted malignancies not listed in the CPRD GOLD	ICD-O-3 morphology codes in the Pathology Registry and the Netherlands Cancer Registry	ICD-9-CM diagnosis codes ICD-10-CM diagnosis codes starting in October 2015 Inpatient admissions, emergency department visits, outpatient hospital clinic visits, and physician visits associated with the diagnoses of interest	ICD-9-CM diagnosis codes ICD-10-CM diagnosis codes starting in October 2015 Inpatient admissions, emergency department visits, outpatient hospital clinic visits, and physician visits associated with the diagnoses of interest

Characteristics	CPRD	PHARMO	HIRD	Medicare
Identification of health care resource utilisation, secondary outcomes	Read codes in the CPRD GOLD and consultation entries for medical encounters and procedures HES disease procedural coding linked to screening and diagnostic tests of interest, coding for inpatient admissions	An internal unique admission identifier will be used to identify hospital admissions and the Dutch Classification of Procedures for procedures linked to screening and diagnostic tests of interest; for a subset of patients, GP visit encounters based on invoiced visits according to VEKTIS codes and encounters with other health care providers (eg, specialists, physical therapists, dieticians) based on communications with the GP coded according to a Dutch coding system (Dutch College of General Practitioners-45) Data on emergency care or specialists visits not available Data from the linked Pathology Registry were used to identify procedural data related to biopsies for primary cancer outcomes	ICD-9-CM procedure codes (before October 2015; ICD-10-CM thereafter), CPT codes, and HCPCS codes	ICD-9-CM encounter codes (before October 2015; ICD-10-CM thereafter) and CPT codes
Reporting requirements	Any cell with 1-4 patients or any cell that allows derivation of a 1-4 value must be masked	Any strata with fewer than 5 patients must be masked	Any cell with 1-4 patients, or any cell that allows derivation of a 1-4 value must be masked	Suppress any cell with 1-10 patients or any cell that allows derivation of a 1-10 value; reporting of minima, maxima, medians, modes, and percentiles is not allowed

CPT = Current Procedural Terminology; GOLD = General Practitioner Online Database of CPRD; HCPCS = Healthcare Common Procedure Coding System; HES = Hospital Episode Statistics.

Appendix G Algorithms for Identifying Medical Conditions

Table 37 Covariate Medical Condition Algorithms for the HIRD and Medicare

Variable	Algorithm
Type 1 diabetes mellitus (cohort exclusion variable)	At least two claims on separate dates for any medically attended visit with a diagnosis code for T1DM on or before the potential index date
Type 2 diabetes mellitus (descriptive variable)	At least one claim for a medically attended visit with any diagnosis code for type 2 diabetes mellitus on or before the potential index date
Indicators of diabetes severity	
Diabetic nephropathy or renal insufficiency	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Peripheral neuropathy	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Peripheral vascular disease	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Retinopathy	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis
Coronary heart disease	At least two outpatient claims on separate dates with any diagnosis code OR one or more diagnoses in the emergency department/inpatient setting OR at least one medical claim with the procedure codes (CPT, ICD-9-CM)
Cerebrovascular disease	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Amputation	At least one medical claim with one of the diagnosis, procedure, or CPT codes
HbA1c test	One occurrence of one of the CPT procedure codes in the final code list
Other medical conditions	
Chronic kidney disease	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Heart failure	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Kidney and genitourinary stones	At least one medical claim with any diagnosis code in any setting
Bladder stones	At least one medical claim with any diagnosis code in any setting
Autoimmune diseases	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Urinary infections (chronic or recurring)	At least two claims with one of the diagnosis codes in any setting on separate dates separated by at least 30 days
Urinary cystitis (chronic or recurring)	At least two claims with one of the diagnosis codes in any setting on separate dates separated by at least 30 days
Colon polyps	At least one claim in any setting with any diagnosis code

Variable	Algorithm
Crohn's disease	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Ulcerative colitis	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Hereditary nonpolyposis colon cancer syndrome (Lynch syndrome)	At least two outpatient claims on separate dates with any diagnosis code OR at least one inpatient claim with the diagnosis code
Immunosuppressive diseases, such as HIV/AIDS	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Peptic ulcer disease	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
All malignancies other than non-melanoma skin cancer	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Helicobacter pylori infection	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Benign mammary dysplasia (females only)	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Polycystic ovarian syndrome (females only)	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code
Benign prostatic hypertrophy (males only)	At least two outpatient claims on separate dates with the diagnosis code OR at least one inpatient claim with the diagnosis code

Note: Outpatient claims in Medicare data were defined as any claim from a hospital outpatient facility or any carrier (physician) claim with a date that did not overlap with an inpatient hospitalisation. Inpatient claims in Medicare and the HIRD were defined as hospital inpatient or skilled nursing facility claims.

CPT = Current Procedural Terminology.

Appendix H Diabetes Severity Complications Index: Categories and Scores

Table 38 Diabetes Severity Complications Index: Ophthalmic (Retinopathy)

Score	ICD-9-CM diagnosis code		ICD-10-CM diagnosis code	
1	250.5x	Diabetic ophthalmologic disease	Main codes E08 E09 E11 E13	 Diabetes mellitus due to underlying conditions Drug or chemical-induced diabetes mellitus Type 2 diabetes mellitus Other specified diabetes mellitus
1	249.5x	Secondary diabetes with ophthalmic manifestations		
1	362.0x, excluding 362.02	Diabetic retinopathy, excluding proliferative diabetic retinopathy		
			Relevant subcodes E**.3x, excluding E**.34x & E**.35x	With ophthalmic complications
1	362.1x	Other background retinopathy and retinal vascular changes	H35.0x	Background retinopathy and retinal vascular changes
1	362.53	Cystoid macular degeneration	H35.35x	Cystoid macular degeneration
1	362.81-362.83	Retinal haemorrhage, retinal exudates and deposits, retinal oedema	H35.6x H35.8x H35.9	Retinal haemorrhage Other specified retinal disorders Unspecified retinal disorder
2	361.x	Retinal detachment, retinoschisis, retinal cysts	H33.x	Retinal detachments and breaks
2	362.02	Proliferative retinopathy	E**.34x E**.35x	Severe non-proliferative diabetic retinopathy Proliferative diabetic retinopathy
2	369.x	Blindness and low vision	H54.x	Blindness and low vision
2	379.23	Vitreous haemorrhage	H43.1x	Vitreous haemorrhage

Table 39 Diabetes Severity Complications Index: Nephropathy

Score	ICD-9-CM diagnosis code		ICD-10-CM diagnosis code	
1	250.4x	Diabetes with renal manifestations	Main codes E08 E09 E11 E13 Relevant subcodes E**.21 E**.22 E**.29	 Diabetes mellitus due to underlying condition Drug- or chemical-induced diabetes mellitus Type 2 diabetes mellitus Other specified diabetes mellitus With diabetic nephropathy With diabetic chronic kidney disease With other diabetic kidney complication
	249.4x	Secondary diabetes with renal manifestations		
1	580.x	Acute glomerulonephritis	N00.x	Acute nephritic syndrome
1	581.x[a]	Nephrotic syndrome	N04.x	Nephrotic syndrome
1	582.x	Chronic glomerulonephritis	N03.x	Chronic nephritic syndrome
1	583.x	Nephritis and nephropathy not specified as acute or chronic	N05.x	Unspecified nephritic syndrome
1	585.1	CKD, Stage 1	N18.1	CKD, Stage 1
	585.2	CKD, Stage 2 (mild)	N18.2	CKD, Stage 2 (mild)
	585.3	CKD, Stage 3 (moderate)	N18.3	CKD, Stage 3 (moderate)
	585.9	CKD, unspecified	N18.9	CKD, unspecified
2	585.4	CKD Stage 4 (severe)	N18.4	CKD, Stage 4 (severe)
	585.5	CKD Stage 5	N18.5	CKD, Stage 5
	585.6	End-stage renal disease	N18.6	End-stage renal disease
2	586, 593.9	Renal failure, unspecified Unspecified disorder of kidney and ureter	N19	Unspecified kidney failure

[a] The description for 581.81 on http://www.icd9data.com/2015/Volume1/580-629/580-589/581/default.htm is "Nephrotic syndrome in diseases classified elsewhere." This code would be implied by 581.x.

CKD = chronic kidney disease.

Table 40 Diabetes Severity Complications Index: Neuropathy

Score	ICD-9-CM diagnosis code		ICD-10-CM diagnosis code	
1	250.6x 357.2	Diabetes with neurological manifestations Polyneuropathy in diabetes	Main codes E08 E09 E11 E13 Relevant subcodes E**.4x	 Diabetes mellitus due to underlying condition Drug- or chemical-induced diabetes mellitus Type 2 diabetes mellitus Other specified diabetes mellitus With neurological complications
	249.6x	Secondary diabetes with neurological manifestations		
1	337.0x	Idiopathic peripheral autonomic neuropathy	G90.09	Other [than carotid sinus syncope] idiopathic peripheral autonomic neuropathy
1	337.1	Peripheral autonomic neuropathy in disorders classified elsewhere	G90.8 G90.9 G99.0	Other disorders of autonomic nervous system Disorder of the autonomic nervous system, unspecified Autonomic neuropathy in diseases classified elsewhere
1	354.x	Mononeuritis of upper limb and mononeuritis multiplex	G56.x	Mononeuropathies of upper limb
1	355.x	Mononeuritis of lower limb and unspecified site	G57.x	Mononeuropathies of lower limb
1	356.9	Unspecified hereditary and idiopathic peripheral neuropathy	G60.9	Hereditary and idiopathic neuropathy, unspecified
1	358.1	Myasthenic syndromes in diseases classified elsewhere	G73.3	Myasthenic syndromes in other diseases classified elsewhere
1	—	—	G90.01	Carotid sinus syncope
1	—	—	H49.x	Paralytic strabismus
1	458.0	Orthostatic hypotension	I95.1	Orthostatic hypotension
1	536.3	Gastroparesis	K31.84	Gastroparesis
1	564.5	Functional diarrhoea	K59.1	Functional diarrhoea
1	596.54	Neurogenic bladder NOS	N31.9	Neuromuscular dysfunction of bladder, unspecified
1	713.5	Arthropathy associated with neurological disorders	M14.6x	Charcot joint
1	951.0 951.1 951.3	Injury to oculomotor nerve Injury to trochlear nerve Injury to abducens nerve	S04.x	Injury to cranial nerve

NOS = not otherwise specified.

Appendix I Results of the Selection of Covariates for the Propensity Score Models

Table 41 Selection of Covariates for Propensity Score Models, Female Breast Cancer Cohorts, Overall and by Insulin Use at the Index Date, by Data Source

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Base model												
Age at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Time period (calendar year) of the index date	1	1	1	1	1	1	1	1	1	1	1	1
Duration of lookback period (years)	1	1	1	1	1	1	1	1	1	1	1	1
Race/ethnicity (CPRD, the HIRD, and Medicare only)	1	1	1	NA	NA	NA	1	1	1	1	1	1
Geographic region (PHARMO, the HIRD, and Medicare only)	NA	NA	NA	1	1	1	1	1	1	1	1	1
Primary care practice region (CPRD only)	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insulin use at the index date	NA	NA	1	NA	NA	1	NA	NA	1	NA	NA	1
Add-on index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Switched-to index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] in the 12 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] 12-24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Number of AD classes[a] > 24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Availability of GP data	NA	NA	NA	1	1	1	NA	NA	NA	NA	NA	NA
Other baseline variables selected into the PS model												
Demographic and lifestyle characteristics												
Socioeconomic status (CPRD, PHARMO, and Medicare only)	1	0	0	0	0	0	NA	NA	NA	1	1	1
Current smoker (CPRD and PHARMO only)	1	1	1	1	1	1	NA	NA	NA	NA	NA	NA
History of alcohol use (CPRD and PHARMO only)	1	0	0	0	0	1	NA	NA	NA	NA	NA	NA
Body mass index (CPRD and PHARMO only)	0	1	1	1	0	0	NA	NA	NA	NA	NA	NA
Medical conditions												
Heart failure	1	0	0	1	1	0	1	1	1	0	0	0
Chronic kidney disease	1	1	1	1	0	0	1	1	0	1	0	0
Benign mammary dysplasia	0	0	0	0	0	0	0	1	0	0	0	0
Diabetic nephropathy or renal insufficiency	0	0	0	1	1	0	1	0	0	1	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Retinopathy	0	0	0	0	0	0	1	1	1	0	0	0
Peripheral neuropathy	1	0	0	1	0	0	0	0	0	0	0	0
Peripheral vascular disease	1	0	0	1	0	0	1	0	0	0	0	0
Coronary heart disease	1	0	0	1	1	1	0	1	1	0	0	0
Cerebrovascular disease	1	0	1	1	0	0	0	1	0	0	0	0
Amputation	1	0	0	0	0	0	1	0	0	0	0	0
Time since first recorded diagnosis of T2DM (CPRD and PHARMO only)	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA
HbA1c level (CPRD and PHARMO only)	1	1	1	0	0	0	NA	NA	NA	NA	NA	NA
No. of HbA1c tests (the HIRD and Medicare only)	NA	NA	NA	NA	NA	NA	1	1	1	1	0	1
Medications												
Hormone-replacement therapy	1	0	0	0	0	0	1	0	0	0	0	0
Selective ERM	0	0	0	0	0	0	1	1	1	0	0	0
Opioids	0	0	0	1	0	0	1	1	1	0	0	0
Index monotherapy	0	0	0	0	1	1	0	0	0	0	0	0
Index combination therapy	0	0	0	0	1	1	0	0	0	0	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**
Non-evaluable index therapy	0	0	0	0	0	0	0	0	0	0	0	0
Health care resource utilisation												
No. of emergency department visits	1	1	1	0	0	0	1	1	1	0	0	0
No. of outpatient visits	0	0	0	1	0	0	1	1	1	0	0	0
No. of hospitalisations	1	1	0	1	1	0	1	1	0	1	0	1
No. of specialty care visits	1	0	0	1	1	1	0	0	0	0	0	0

[a] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

Note: In the table, 1 means that the variable was included in propensity score models; 0 means that the variable was not included.

ERM = oestrogen receptor modulator; NA = not applicable; No. = number; PS = propensity score.

Table 42 Selection of Covariates for the Propensity Score Models, Sex-Combined Bladder Cancer Cohorts, Overall and by Insulin Use at the Index Date, by Data Source

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Base model												
Age at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Sex (male)	1	1	1	1	1	1	1	1	1	1	1	1
Time period (calendar year) of the index date	1	1	1	1	1	1	1	1	1	1	1	1
Duration of lookback period (years)	1	1	1	1	1	1	1	1	1	1	1	1
Race/ethnicity (CPRD, the HIRD, and Medicare only)	1	1	1	NA	NA	NA	1	1	1	1	1	1
Geographic region (PHARMO, the HIRD, and Medicare only)	NA	NA	NA	1	1	1	1	1	1	1	1	1
Primary care practice region (CPRD only)	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insulin use at the index date	NA	NA	1	NA	NA	1	NA	NA	1	NA	NA	1
Pioglitazone use at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Add-on index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Switched-to index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] in the 12 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**
Number of AD classes[a] 12-24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] > 24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Availability of GP data	NA	NA	NA	1	1	1	NA	NA	NA	NA	NA	NA
Other variables selected into the PS model												
Demographic and lifestyle characteristics												
Socioeconomic status (CPRD, PHARMO, and Medicare only)	1	0	0	1	1	1	NA	NA	NA	0	0	0
Current smoker (CPRD and PHARMO only)	1	1	1	1	1	1	NA	NA	NA	NA	NA	NA
History of alcohol use (CPRD and PHARMO only)	1	0	0	1	1	1	NA	NA	NA	NA	NA	NA
Body mass index (CPRD and PHARMO only)	1	1	1	1	0	1	NA	NA	NA	NA	NA	NA
Medical conditions												
Heart failure	1	0	0	1	1	1	1	1	1	0	0	0
Chronic kidney disease	1	1	1	1	0	0	1	1	1	1	1	1
Diabetic nephropathy or renal insufficiency	0	0	0	1	0	0	1	1	1	1	1	1

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Retinopathy	0	0	0	1	0	0	0	1	1	0	0	0
Peripheral neuropathy	1	0	0	1	0	1	1	1	0	0	0	0
Peripheral vascular disease	1	1	1	1	1	1	0	1	1	0	1	0
Coronary heart disease	0	0	0	1	0	1	1	1	1	0	0	0
Cerebrovascular disease	1	0	0	1	0	0	0	1	1	0	0	0
Amputation	1	0	0	1	0	0	1	1	0	0	0	0
Time since first recorded diagnosis of T2DM (CPRD and PHARMO only)	0	0	0	1	0	1	NA	NA	NA	NA	NA	NA
HbA1c level (CPRD and PHARMO only)	1	1	0	1	1	1	NA	NA	NA	NA	NA	NA
No. of HbA1c tests (the HIRD and Medicare only)	NA	NA	NA	NA	NA	NA	1	1	1	0	1	1
Chronic or recurrent urinary infections	0	0	0	1	0	0	1	0	0	0	0	0
Chronic or recurrent urinary cystitis	0	0	0	1	0	1	1	1	1	0	0	0
Kidney stones	1	0	0	1	0	0	1	1	1	0	0	0
Bladder stones	0	0	0	1	0	0	1	0	0	0	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Medications												
Opioids	0	0	0	1	0	0	1	1	1	0	0	0
Index monotherapy	0	0	0	0	1	1	0	0	0	0	0	0
Index combination therapy	0	0	0	1	1	1	0	0	0	0	0	0
Non-evaluable index therapy	0	0	0	1	0	0	0	0	0	0	0	0
Health care resource utilisation												
No. of emergency department visits	1	0	0	0	0	0	0	0	0	0	0	0
No. of outpatient visits	1	0	0	1	1	1	1	1	1	0	1	1
No. of hospitalisations	1	0	1	1	1	0	0	1	0	0	0	0
No. of specialty care visits	1	0	0	1	0	0	0	0	0	0	0	0

[a] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

Note: In the table, 1 means that the variable was included in propensity score models; 0 means that the variable was not included.

NA = not applicable; No. = number; PS = propensity score.

Table 43 Selection of Covariates for the Propensity Score Models, Female Bladder Cancer Cohorts, Overall and by Insulin Use at the Index Date, by Data Source

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Base model												
Age at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Time period (calendar year) of the index date	1	1	1	1	1	1	1	1	1	1	1	1
Duration of lookback period (years)	1	1	1	1	1	1	1	1	1	1	1	1
Race/ethnicity (CPRD, the HIRD, and Medicare only)	1	1	1	NA	NA	NA	1	1	1	1	1	1
Geographic region (PHARMO, the HIRD, and Medicare only)	NA	NA	NA	1	1	1	1	1	1	1	1	1
Primary care practice region (CPRD only)	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insulin use at the index date	NA	NA	1	NA	NA	1	NA	NA	1	NA	NA	1
Pioglitazone use at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Add-on index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Switched-to index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] in the 12 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] 12-24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**
Number of AD classes[a] > 24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Availability of GP data	NA	NA	NA	1	1	1	NA	NA	NA	NA	NA	NA
Other variables selected into the PS model												
Demographic and lifestyle characteristics												
Socioeconomic status (CPRD, PHARMO, and Medicare only)	1	0	0	1	1	1	NA	NA	NA	0	0	0
Current smoker (CPRD and PHARMO only)	1	1	1	1	1	1	NA	NA	NA	NA	NA	NA
History of alcohol use (CPRD and PHARMO only)	1	1	1	1	1	1	NA	NA	NA	NA	NA	NA
Body mass index (CPRD and PHARMO only)	1	0	1	1	1	1	NA	NA	NA	NA	NA	NA
Medical conditions												
Heart failure	1	0	0	1	1	1	1	1	1	0	0	0
Chronic kidney disease	1	1	1	1	1	1	1	0	0	0	0	1
Diabetic nephropathy or renal insufficiency	1	0	1	1	1	1	1	1	1	0	0	1
Retinopathy	1	0	0	1	0	1	1	1	1	0	1	0
Peripheral neuropathy	1	0	0	1	0	1	1	1	1	0	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**
Peripheral vascular disease	1	1	1	1	1	1	1	1	1	0	1	1
Coronary heart disease	0	0	0	1	0	1	1	1	1	0	0	0
Cerebrovascular disease	1	0	0	1	1	1	1	1	0	0	0	0
Amputation	1	0	0	1	0	0	1	0	1	0	0	0
Time since first recorded diagnosis of T2DM (CPRD and PHARMO only)	0	0	0	1	1	1	NA	NA	NA	NA	NA	NA
HbA1c level (CPRD and PHARMO only)	1	0	0	1	0	0	NA	NA	NA	NA	NA	NA
No. of HbA1c tests (the HIRD and Medicare only)	NA	NA	NA	NA	NA	NA	1	1	1	0	0	0
Lynch syndrome	0	0	0	0	0	0	0	0	1	0	0	0
Chronic or recurrent urinary infections	1	0	0	1	0	1	1	1	1	0	0	0
Chronic or recurrent urinary cystitis	1	0	0	1	0	1	1	1	1	0	0	0
Kidney stones	1	0	0	1	1	1	1	1	1	0	0	0
Bladder stones	0	0	0	1	0	0	1	0	1	0	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Medications												
Cyclophosphamide	0	0	0	0	0	0	0	0	1	0	0	0
Opioids	0	0	0	1	1	1	1	0	1	0	0	0
Index monotherapy	0	0	0	0	1	1	0	0	0	0	0	0
Index combination therapy	0	0	0	1	1	1	0	0	0	0	0	0
Non-evaluable index therapy	0	0	0	1	1	1	0	0	0	0	0	0
Health care resource utilisation												
No. of emergency department visits	1	0	0	0	0	0	1	1	1	0	0	0
No. of outpatient visits	1	0	0	1	1	1	1	1	1	0	1	1
No. of hospitalisations	1	0	1	1	1	1	1	1	1	0	0	0
No. of specialty care visits	1	0	1	1	0	0	0	0	0	0	0	0

[a] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

Note: In the table, 1 means that the variable was included in propensity score models; 0 means that the variable was not included.

NA = not applicable; No. = number; PS = propensity score.

Table 44 Selection of Covariates for the Propensity Score Models, Male Bladder Cancer Cohorts, Overall and by Insulin Use at the Index Date, by Data Source

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Base model												
Age at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Time period (calendar year) of the index date	1	1	1	1	1	1	1	1	1	1	1	1
Duration of lookback period (years)	1	1	1	1	1	1	1	1	1	1	1	1
Race/ethnicity (CPRD, the HIRD, and Medicare only)	1	1	1	NA	NA	NA	1	1	1	1	1	1
Geographic region (PHARMO, the HIRD, and Medicare only)	NA	NA	NA	1	1	1	1	1	1	1	1	1
Primary care practice region (CPRD only)	1	1	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insulin use at the index date	NA	NA	1	NA	NA	1	NA	NA	1	NA	NA	1
Pioglitazone use at the index date	1	1	1	1	1	1	1	1	1	1	1	1
Add-on index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Switched-to index therapy	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] in the 12 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Number of AD classes[a] 12-24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**	**Insulin**	**No insulin**	**Overall**
Number of AD classes[a] > 24 months before the index date	1	1	1	1	1	1	1	1	1	1	1	1
Availability of GP data	NA	NA	NA	1	1	1	NA	NA	NA	NA	NA	NA
Other variables selected into the PS model												
Demographic and lifestyle characteristics												
Socioeconomic status (CPRD, PHARMO, and Medicare only)	1	0	0	1	0	1	NA	NA	NA	0	0	0
Current smoker (CPRD and PHARMO only)	1	0	1	1	1	1	NA	NA	NA	NA	NA	NA
History of alcohol use (CPRD and PHARMO only)	1	0	0	1	0	1	NA	NA	NA	NA	NA	NA
Body mass index (CPRD and PHARMO only)	1	1	1	1	0	0	NA	NA	NA	NA	NA	NA
Medical conditions												
Heart failure	1	0	1	1	0	0	1	1	1	0	0	0
Chronic kidney disease	1	1	0	1	0	0	1	1	1	0	1	1
Diabetic nephropathy or renal insufficiency	0	0	0	0	0	0	1	1	0	1	1	1
Retinopathy	0	0	0	1	0	0	1	1	1	0	0	0
Peripheral neuropathy	1	0	0	1	0	1	1	1	1	0	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Peripheral vascular disease	0	1	1	1	1	1	1	1	1	0	0	0
Coronary heart disease	1	0	0	1	0	0	1	1	1	0	1	0
Cerebrovascular disease	1	0	0	1	0	0	1	1	1	0	0	0
Amputation	1	0	0	1	0	0	1	1	1	0	0	0
Time since first recorded diagnosis of T2DM (CPRD and PHARMO only)	0	0	0	1	1	1	NA	NA	NA	NA	NA	NA
HbA1c level (CPRD and PHARMO only)	1	1	0	1	1	1	NA	NA	NA	NA	NA	NA
No. of HbA1c tests (the HIRD and Medicare only)	NA	NA	NA	NA	NA	NA	1	1	1	0	1	1
Lynch syndrome	0	0	0	0	0	0	1	0	0	0	0	0
Chronic or recurrent urinary infections	1	0	0	1	0	0	1	0	0	0	0	0
Chronic or recurrent urinary cystitis	1	0	0	1	0	0	1	0	0	0	0	0
Kidney stones	0	0	0	1	0	0	1	1	1	0	0	0
Bladder stones	0	0	0	1	0	0	1	1	1	0	0	0

Covariates	CPRD			PHARMO			HIRD			Medicare		
	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall	Insulin	No insulin	Overall
Medications												
Cyclophosphamide	0	0	0	0	0	0	1	0	0	0	0	0
Opioids	1	0	0	1	0	0	1	0	0	0	0	0
Index monotherapy	0	0	0	0	1	1	0	0	0	0	0	0
Index combination therapy	0	0	0	1	1	1	0	0	0	0	0	0
Non-evaluable index therapy	0	0	0	1	0	0	0	0	0	0	0	0
Health care resource utilisation												
No. of emergency department visits	1	1	0	0	0	0	1	1	1	0	0	0
No. of outpatient visits	0	1	1	1	1	1	1	1	1	0	1	1
No. of hospitalisations	1	0	1	1	1	0	1	1	1	0	0	0
No. of specialty care visits	1	0	0	1	0	1	0	0	0	0	0	0

[a] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

Note: In the table, 1 means that the variable was included in propensity score models; 0 means that the variable was not included.

NA = not applicable; No. = number; PS = propensity score.

Table 45 Selection of Covariates for the Propensity Score Models, Sex-Combined Bladder Cancer Cohorts, by Pioglitazone Use at the Index Date (CPRD, the HIRD, and Medicare)

Covariates	CPRD		HIRD		Medicare	
	Pioglitazone	No pioglitazone	Pioglitazone	No pioglitazone	Pioglitazone	No pioglitazone
Base model						
Age at the index date	1	1	1	1	1	1
Sex (male)	1	1	1	1	1	1
Time period (calendar year) of the index date	1	1	1	1	1	1
Duration of lookback period (years)	1	1	1	1	1	1
Race/ethnicity	1	1	1	1	1	1
Geographic region (the HIRD and Medicare only)	NA	NA	1	1	1	1
Primary care practice region (CPRD only)	1	1	NA	NA	NA	NA
Insulin use at the index date	1	1	1	1	1	1
Add-on index therapy	1	1	1	1	1	1
Switched-to index therapy	1	1	1	1	1	1
Number of AD classes[a] in the 12 months before the index date	1	1	1	1	1	1
Number of AD classes[a] 12-24 months before the index date	1	1	1	1	1	1
Number of AD classes[a] > 24 months before the index date	1	1	1	1	1	1

Covariates	CPRD		HIRD		Medicare	
	Pioglitazone	**No pioglitazone**	**Pioglitazone**	**No pioglitazone**	**Pioglitazone**	**No pioglitazone**
Other variables selected into the PS model						
Demographic and lifestyle characteristics						
Socioeconomic status (CPRD and Medicare only)	1	0	NA	NA	0	0
Current smoker (CPRD only)	1	1	NA	NA	NA	NA
History of alcohol use (CPRD only)	1	0	NA	NA	NA	NA
Body mass index (CPRD only)	1	1	NA	NA	NA	NA
Medical conditions						
Heart failure	1	0	1	1	0	0
Chronic kidney disease	1	0	1	1	1	1
Diabetic nephropathy or renal insufficiency	1	0	1	1	0	1
Retinopathy	1	0	1	1	0	0
Peripheral neuropathy	0	0	1	1	0	0
Peripheral vascular disease	1	1	1	1	1	0
Coronary heart disease	1	0	0	1	0	0
Cerebrovascular disease	1	0	1	1	0	0
Amputation	1	0	1	1	0	0
HbA1c level (CPRD only)	1	1	NA	NA	NA	NA
No. of HbA1c tests (the HIRD and Medicare only)	NA	NA	1	1	1	0
Chronic or recurrent urinary infections	1	0	1	1	0	0
Chronic or recurrent urinary cystitis	0	0	1	1	0	0
Kidney stones	0	0	1	1	0	0
Bladder stones	1	0	1	1	0	0

Covariates	CPRD		HIRD		Medicare	
	Pioglitazone	No pioglitazone	Pioglitazone	No pioglitazone	Pioglitazone	No pioglitazone
Medications						
Opioids	1	0	1	0	0	0
Health care resource utilisation						
No. of emergency department visits	1	0	1	1	0	0
No. of outpatient visits	1	0	1	1	1	0
No. of hospitalisations	1	1	1	0	0	0
No. of specialty care visits	1	0	0	0	0	0

[a] Antidiabetic drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha-glucosidase inhibitors, and meglitinides.

Note: In the table, 1 means that the variable was included in propensity score models; 0 means that the variable was not included.

Note: Covariate balance plots stratified by pioglitazone use at the index date were not created for PHARMO. In PHARMO, there was an insufficient number of bladder cancer events (ie, < 3) amongst dapagliflozin users in the pioglitazone use group to conduct the comparative analysis stratified by pioglitazone use at the index date in the sex-combined bladder cancer cohort; therefore, propensity score modelling and trimming were not conducted stratified by pioglitazone use at the index date.

NA = not applicable; No. = number; PS = propensity score.

Appendix J Analysis Tables and Figures