30th Annual BASS

SGLT2 inhibitors and cancer prognosis –using SEER-Medicare linked data

Juhua Luo, PhD, Professor
Department of Epidemiology and Biostatistics
School of Public Health, Indiana University
Oct. 24, 2023

Disclosures

- Funding : This research was supported by the NIH-NCI (R03CA256238).
- Potential Conflict of Interest: No
- Disclaimer: This presentation only reflects the views of the presenter

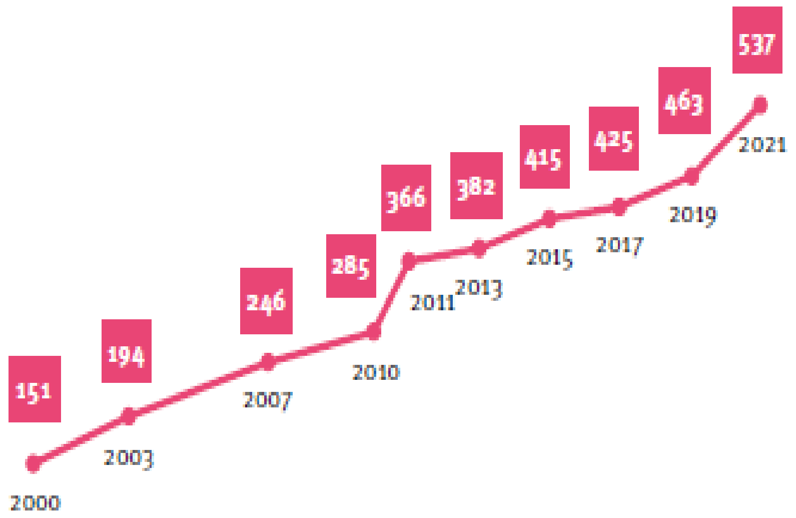
Outline

- Background on diabetes, diabetes treatments and SGLT2 inhibitors
- SGLT2 inhibitors and cancer prognosis
  - SEER-Medicare linked data
  - Challenges of using real-world data to estimate the causal effect of a treatment
  - Several approaches to address some of the challenges

Number of people with diabetes

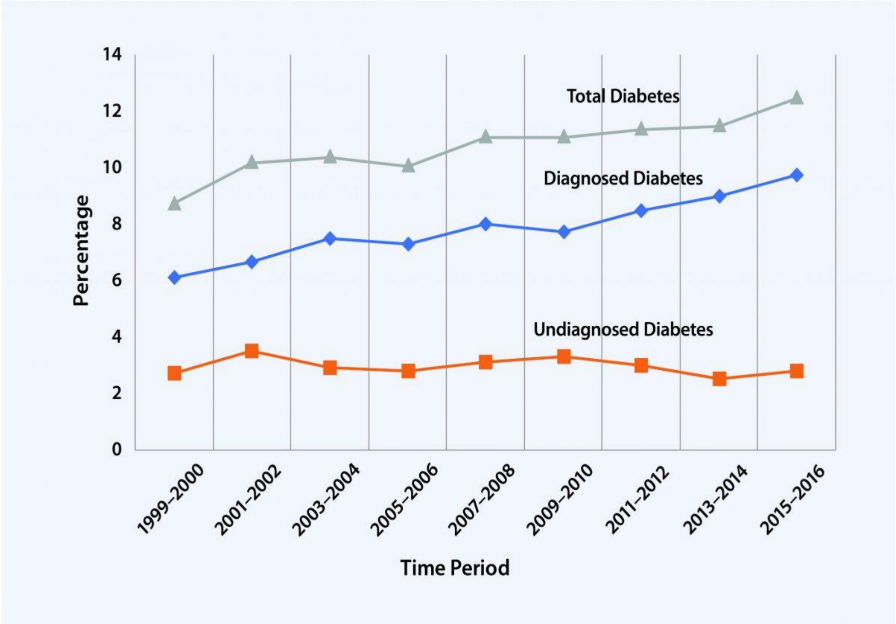
Worldwide

537millions (10.5%)



US

37.3 millions (11.3%)



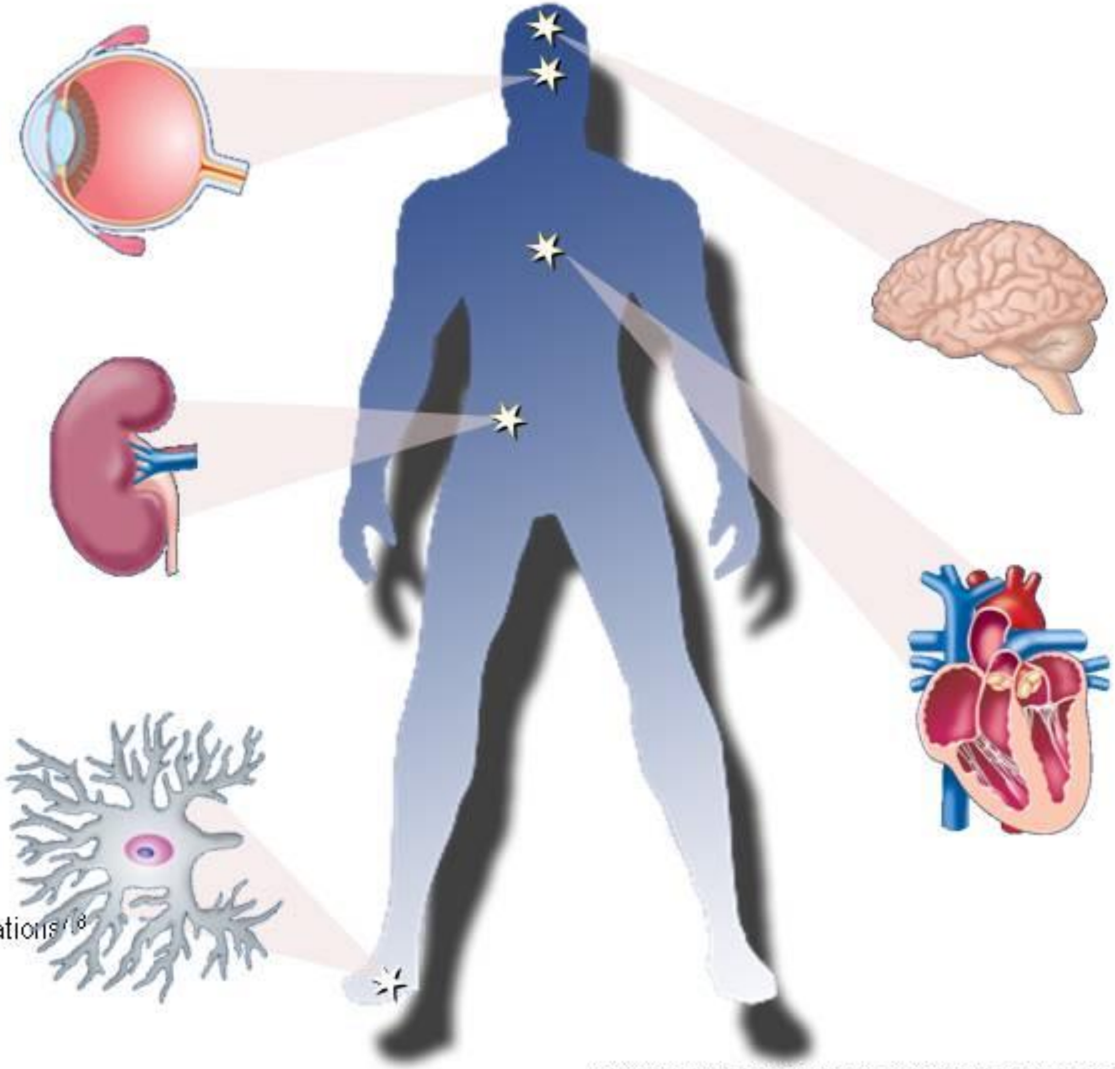
The estimated prevalence of diabetes in adults aged 20-79 years has **more than tripled**, from 151 million in 2000 to 537 million today.

Prediction: 783 million (12.2%) by 2045

International Diabetes Federation (IDF) Diabetes Atlas 10th edition in 2021.

https://www.cdc.gov/diabetes/library/reports/reportcard/national-state-diabetes-trends.html

Diabetes is a lifelong condition associated with serious complications



Diabetic Retinopathy

Leading cause of blindness in adults[1,2]

Diabetic Nephropathy

Leading cause of endstage renal disease[3,4]

Diabetic Neuropathy

Leading cause of non-traumatic lower extremity amputations[7,8]

Stroke

2- to 4-fold increase in CV mortality and stroke[5]

Cardiovascular Disease

8/10 individuals with diabetes die from CV events[6]

SANOFI DIABETES

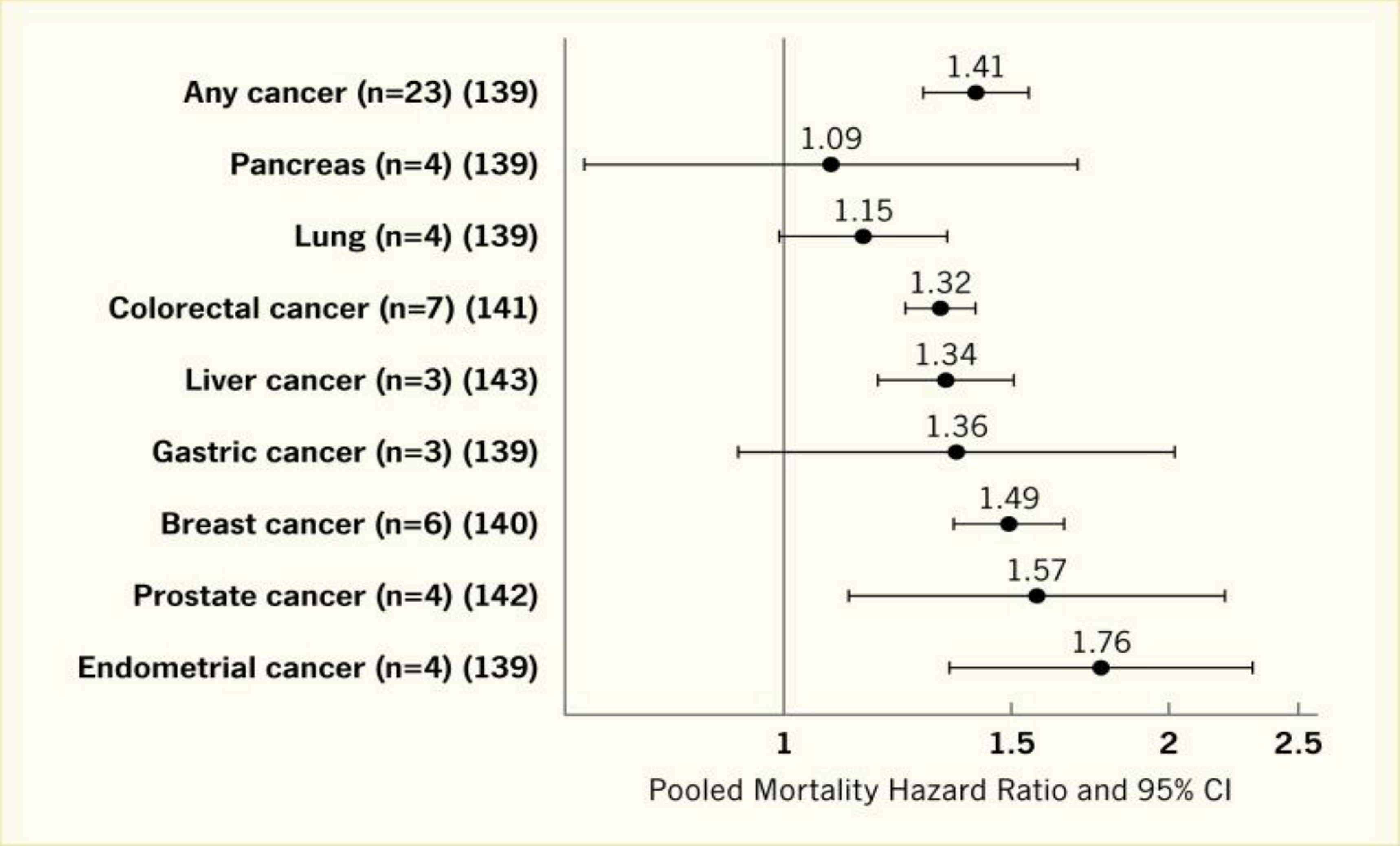
1. UKPDS Group. *Diabetes Res* 1990;13(1):1–11. 2. Fong DS *et al. Diabetes Care* 2003;26(Suppl 1):S99–S102. 3. HDS. *J Hypertens* 1993;11(3):309–317.
4. Molitch ME *et al. Diabetes Care* 2003;26(Suppl 1):S94–S98. 5. Kannel WB *et al. Am Heart J* 1990;120:672–676. 6. Gray RP *et al.* In *Textbook of Diabetes* 2nd Edition, 1997. 7. King's Fund. London: British Diabetic Association, 1996.
8. Mayfield JA *et al. Diabetes Care* 2003;26(Suppl 1):S78–S79

Summary of Meta-analysis: Diabetes as a risk factor for cancer

TUMOR TYPE	NUMBER OF STUDIES	Relative RISK (95% CI)
Pancreas	35 cohort	1.94 (1.66–2.27)
Colon and rectum	14 (6 case-control; 8 cohort)	Colon: 1.38 (1.26–1.51) Rectum: 1.20 (1.09–1.31)
Liver	49	2.2 (1.7–3.0)
Kidney	9 cohort	1.42 (1.06–1.91)
Bladder	29 cohort	1.29 (1.08–1.54)
Breast	39	1.27 (1.16–1.39)
Endometrium	16 (13 case-control; 3 cohort)	2.10 (1.75–2.53)
Prostate	19	0.84 (0.76–0.93)

From Chapter 29. Cancer and Diabetes. Diabetes in America. 3rd edition. Cowie CC, Casagrande SS, Menke A, et al., editors. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases (US); 2018 Aug.

Summary of Meta-Analysis of diabetes and mortality in patients with cancer



From Chapter 29. Cancer and Diabetes. Diabetes in America. 3rd edition. Cowie CC, Casagrande SS, Menke A, et al., editors. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases (US); 2018 Aug.

Do diabetes treatments influence cancer?

Glycemic control

- 1. Lifestyle changes
  - Dietary modification
  - Increased physical activity
- 2. Medications
  - Insulin injection
  - Oral medications
    - Metformin
    - Sulfonylureas
    - Thiazolidinediones (TZDs)
    - DPP-4 inhibitors
    - GLP-1 and dual GLP-1/GIP receptor agonists
    - **SGLT2 inhibitors**

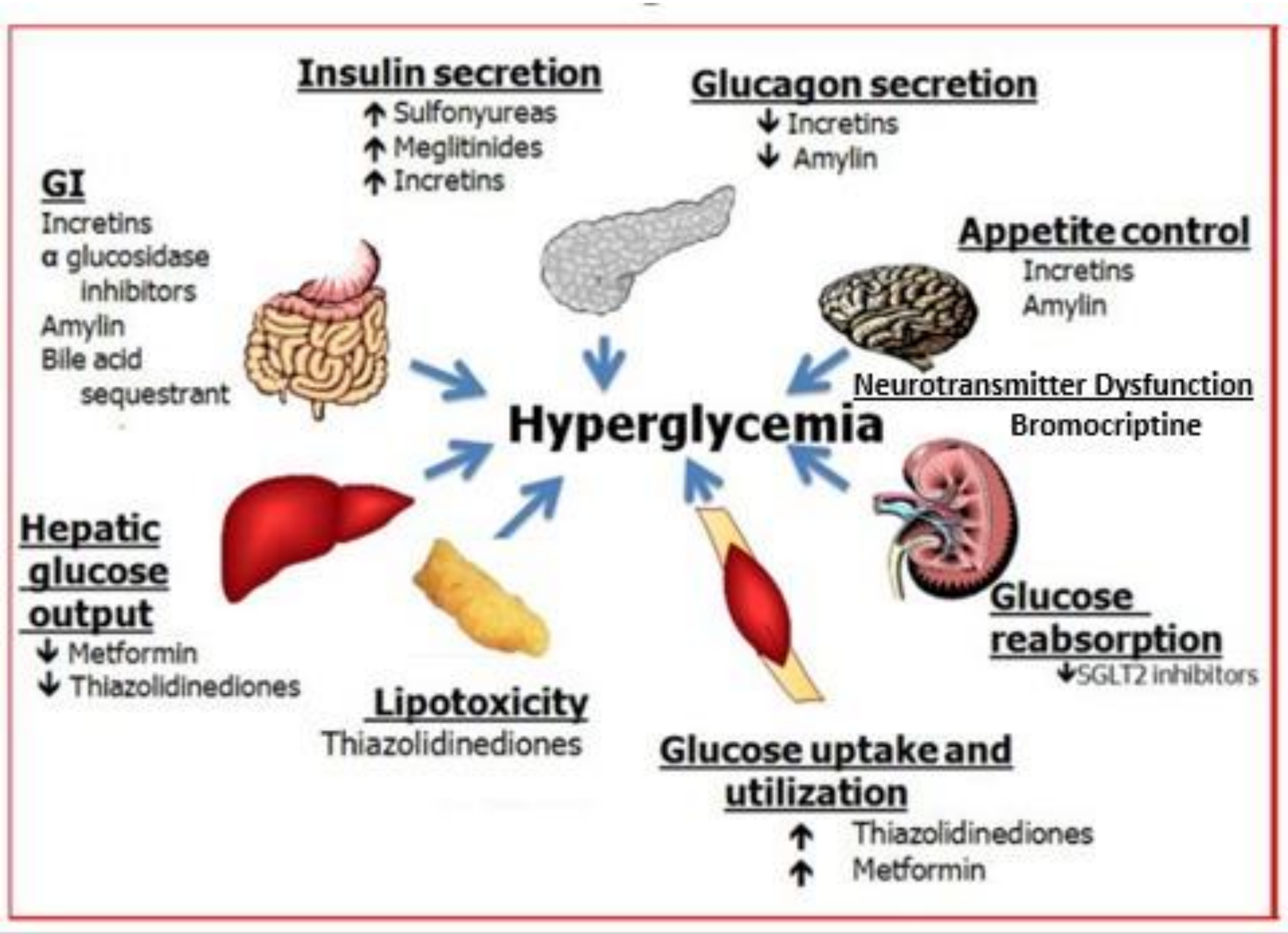
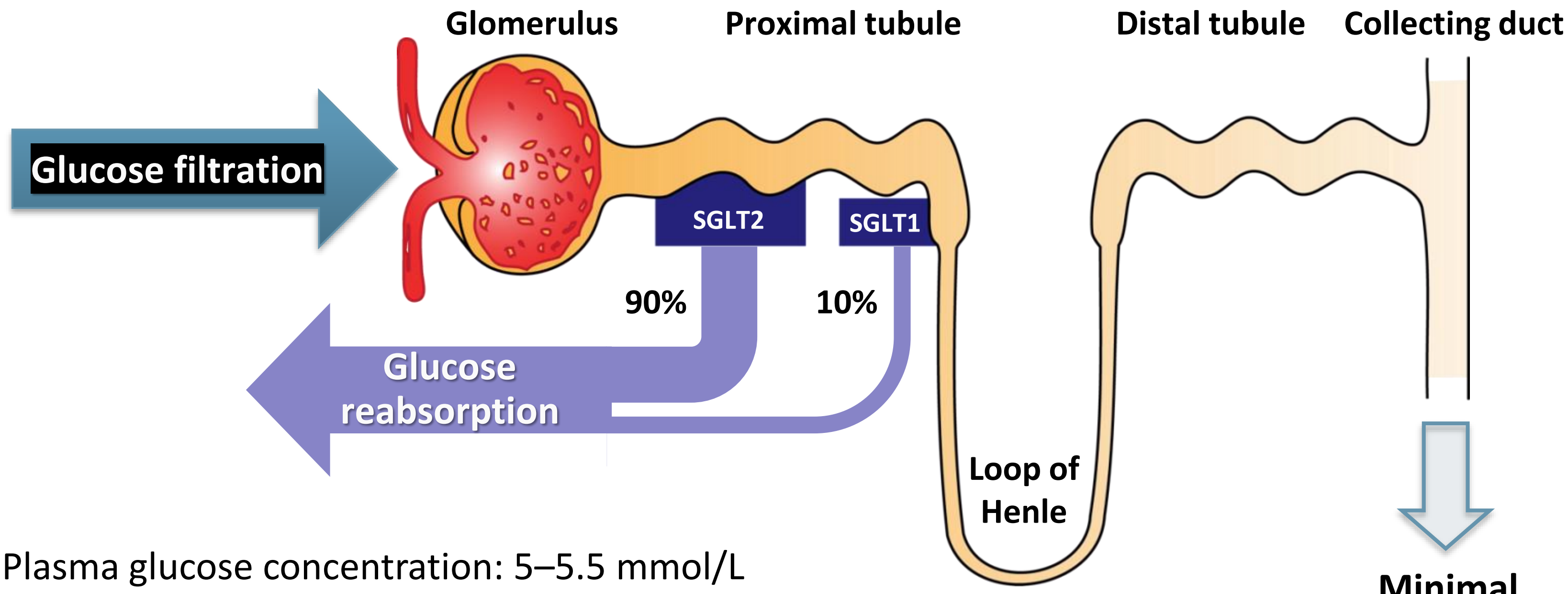


Image source: Oral and injectable (No-insulin) pharmacological agents for the treatment of type 2 diabetes. Endotext, Feingold KR, Anawalt B, Blackman MR, et al., editors. South Dartmouth (MA): MDText.com, Inc.

Sodium-Glucose Cotransporter 2 (SGLT2) inhibitors

- The newest class of glucose-lowering agents
- Currently available drugs in the US
  - Canagliflozin (Invokana, first approved by FDA in 2013)
  - Dapagliflozin (Farxiga, 2014)
  - Empagliflozin (Jardiance, 2014)
  - Ertugliflozin (Steglatro, 2017)

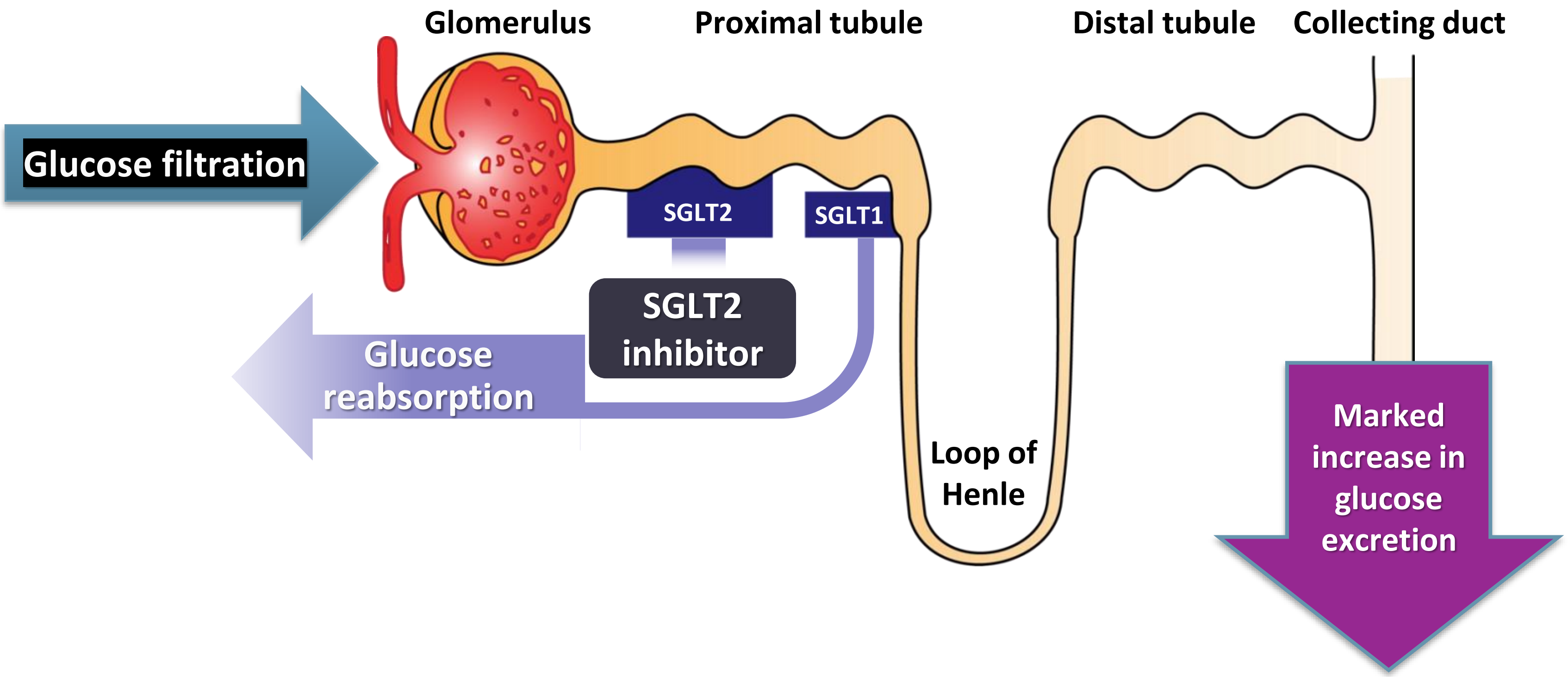
Renal glucose handling in the nephron of the healthy individual

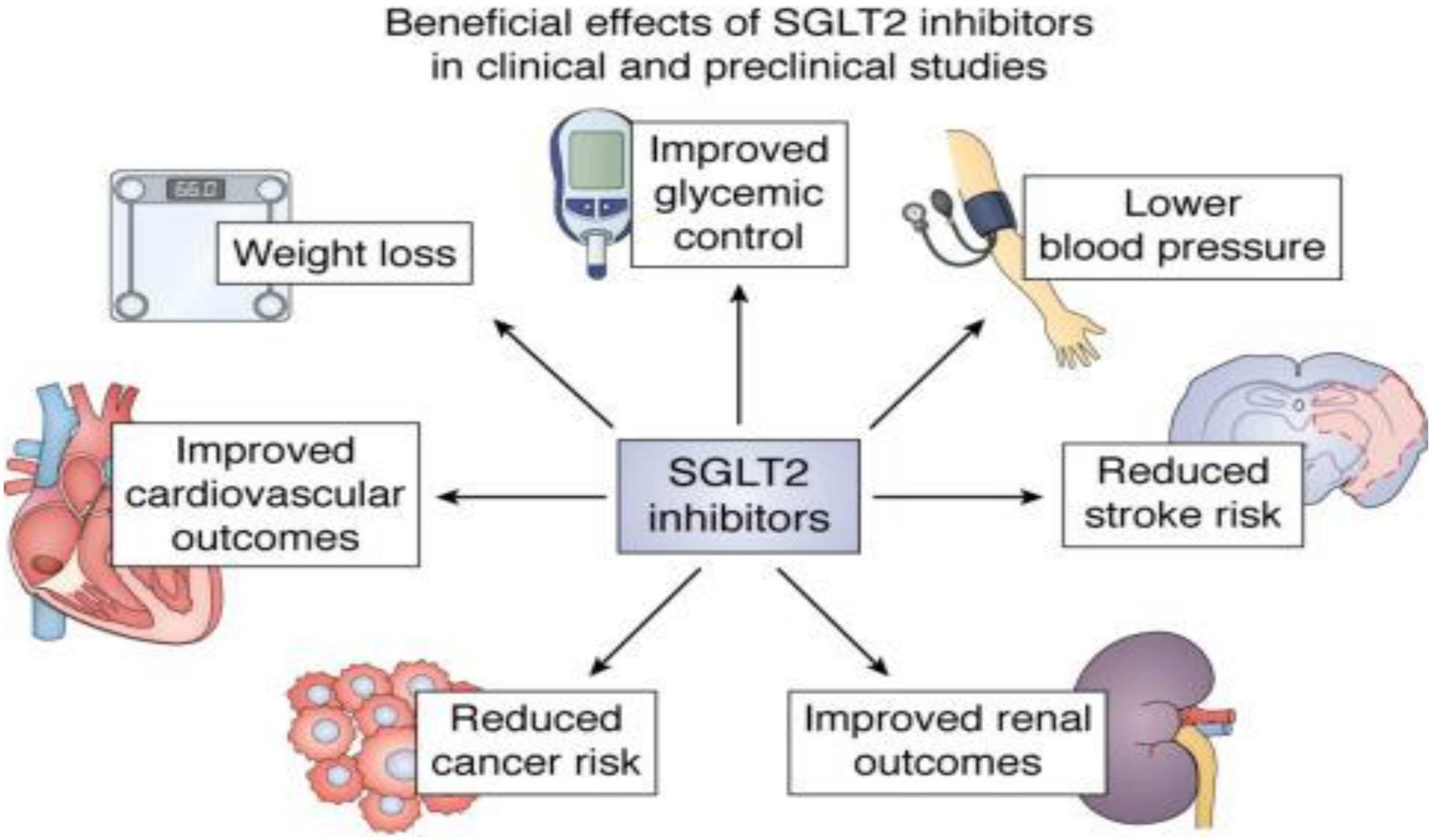


- Plasma glucose concentration: 5–5.5 mmol/L
- Plasma filtered: 180 L/day
- Glucose filtered: 160–180 g/day
- Glucose excreted: Minimal

Image adapted from: Bailey CJ. *Trends Pharmacol Sci.* 2011;32:63–71.

SGLT2 inhibition lowers the elevated renal threshold for glucose in type 2 diabetes[1]





** Perry and Shulman J. Biol Chem, 2020, 295(42): 14379-390*

SGLT2 inhibitors and cancer

- SGLT2 inhibitors may have benefits for several cancers, including lung, liver, kidney, colon, pancreatic cancer and breast cancers. E.g.
  - Studies have demonstrated functional SGLT2 activity in pancreatic and prostate adenocarcinomas.
  - SGLT2 inhibitors reduced tumor growth of early-stage lung adenocarcinoma and prolonged survival in mouse models and xenographs.
  - SGLT2 inhibitors suppressed proliferation in hepatocellular carcinoma (HCC) cells
  - SGLT2 inhibitors reduced human HCC xenograph tumor growth or fewer hepatic tumors in mouse models

- Hypothesis: SGLT2 inhibitors may improve lung, HCC and pancreatic cancer prognosis

Research gaps

- **Population-level epidemiological studies investigating the impact of SGLT2 inhibitors on cancer prognosis are scarce**

Our project aim

- **To test the hypothesis that SGLT2 inhibitors may be promising for certain cancer therapy**

SGLT2 inhibitors → HCC mortality ? ↓

Estimating Causal Effects of a Treatment

- RCTs are ideal for estimating causal effects
  - Strengths: balance both known and unknown confounding factors between groups -> high internal validity
  - Limitations: Limited external validity or generalizability, due to strict inclusion and exclusion criteria; ethical consideration, generally higher cost and shorter-duration (vs observational studies), inefficiency for rare or delayed outcomes.
- Need alternative approach

Observational studies using real-world data

Present unique opportunities

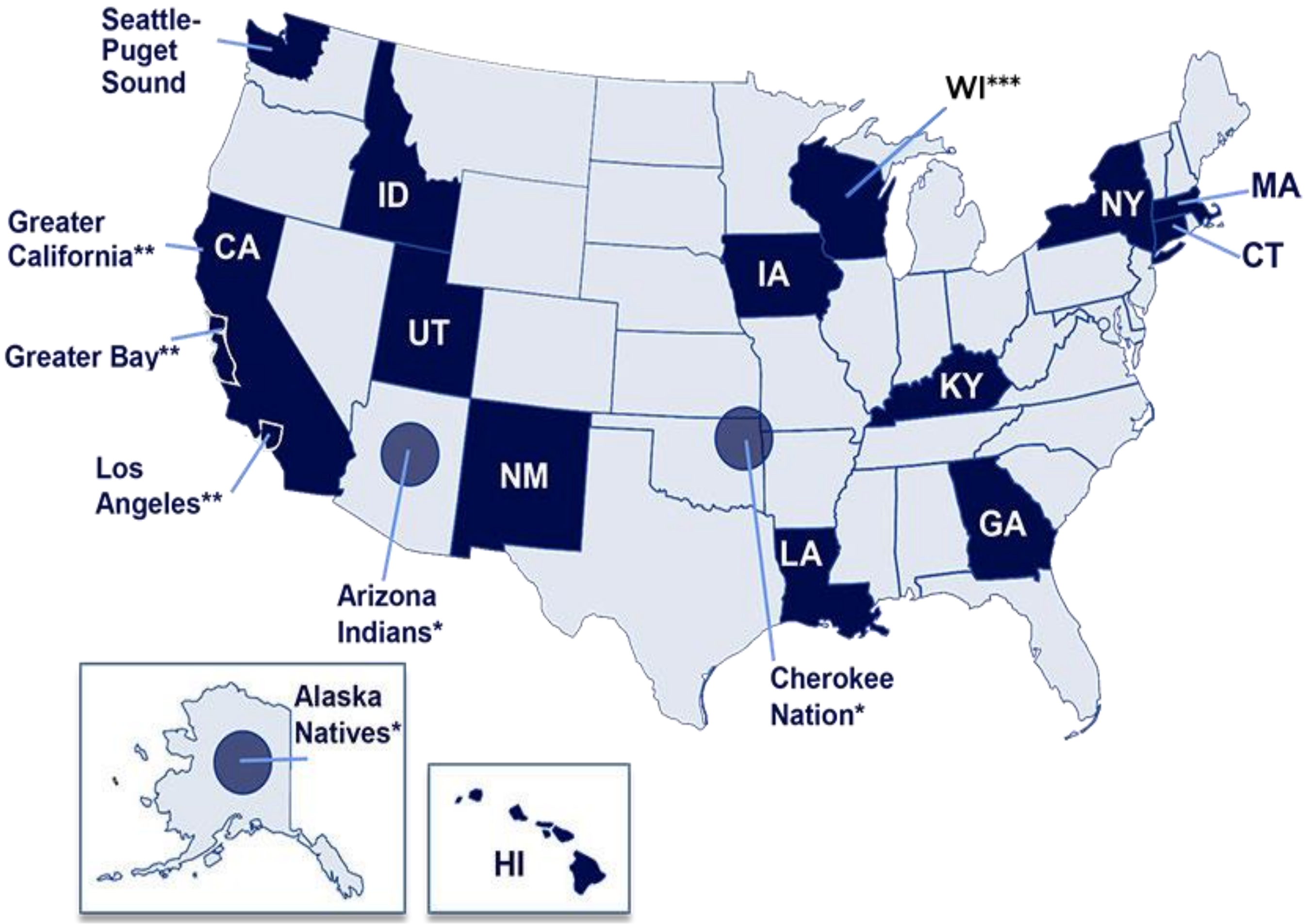
- Can address broader research questions
- Few filtering out of patients through inclusion/exclusion criteria -> increases generalizability
- With retrospective data, there are no biases driven by known hypotheses
- Cost-effectiveness: Less cost and time consuming (vs RCTs)

Database – the linked SEER-Medicare data

- Data from SEER cancer registers combined with data from the Medicare program
  - The linkage was first completed in 1991, updated in 1995, 1999, and biennially in recent years
  - As of 2023, the data includes all Medicare eligible persons who were diagnosed with cancer through 2019, and their Medicare claims through 2019.

SEER Registries (2018-present)

- SEER: cancer registries that routinely collect information on all newly diagnosed cancer cases in SEER areas
- It began collecting data on cancer cases on January 1, 1973
- Current data available– 48% of U.S. population covered



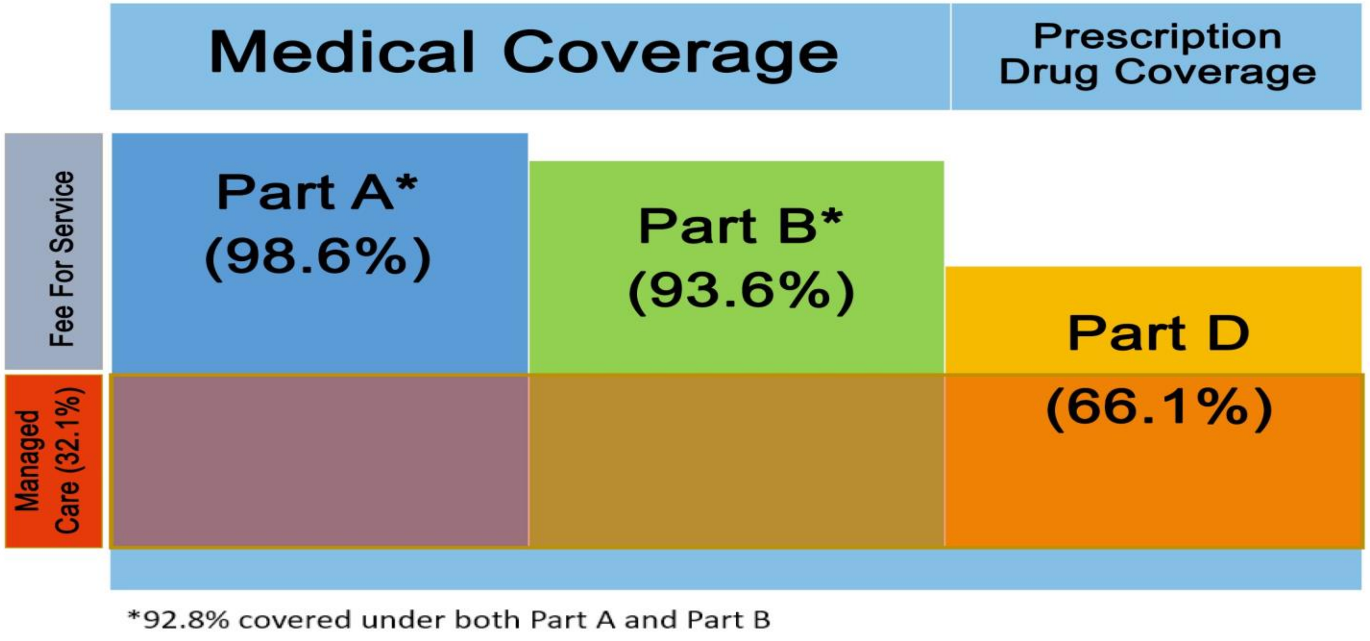
* Subcontract under New Mexico
** Three regions represent the state of California: Greater Bay, Los Angeles, and Greater California
*** Research support registry only; not under contract to submit data

Data Collected by the SEER Registries

- Demographic
  - Residence at diagnosis, sex, age at diagnosis, race/Hispanic origin, marital status
- Information about the cancer
  - Identification of the cancer
  - Tumor related information
- Treatment
- Mortality

The Medicare Program and coverage

- The Medicare program is a federal health insurance plan available to qualifying elders (age 65 and older) and select disabled adults



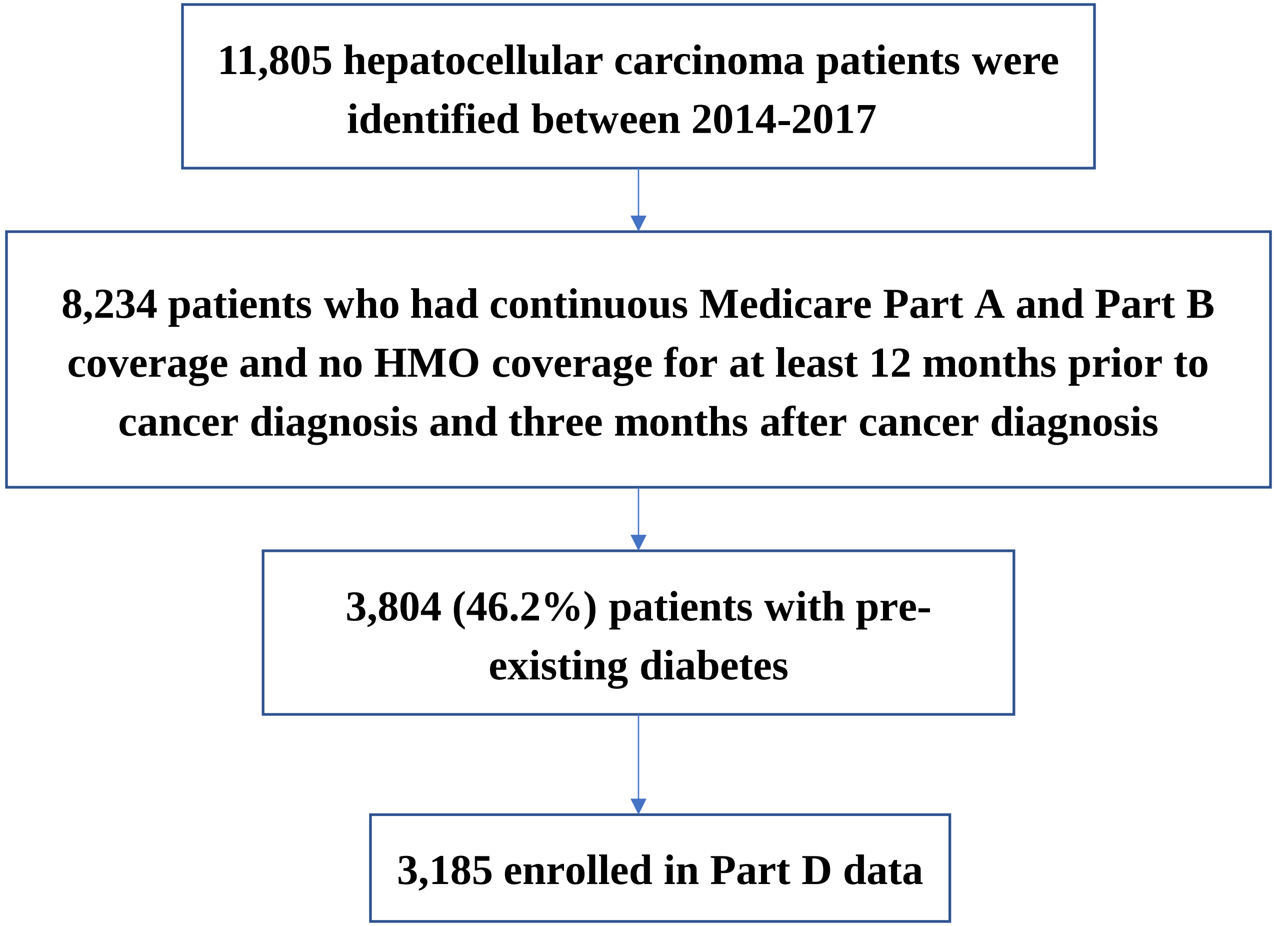
What Information in the Medicare Data?

- The Medicare program collects all claims for fee- for-service care provided from the time a person enrolls in Medicare until death
  - It can include care prior to the diagnosis, peri-diagnosis, and following diagnosis

However,

- Behavioral risk factors rarely coded, such as
  - Obesity, smoking, and family history

Study population flow chart



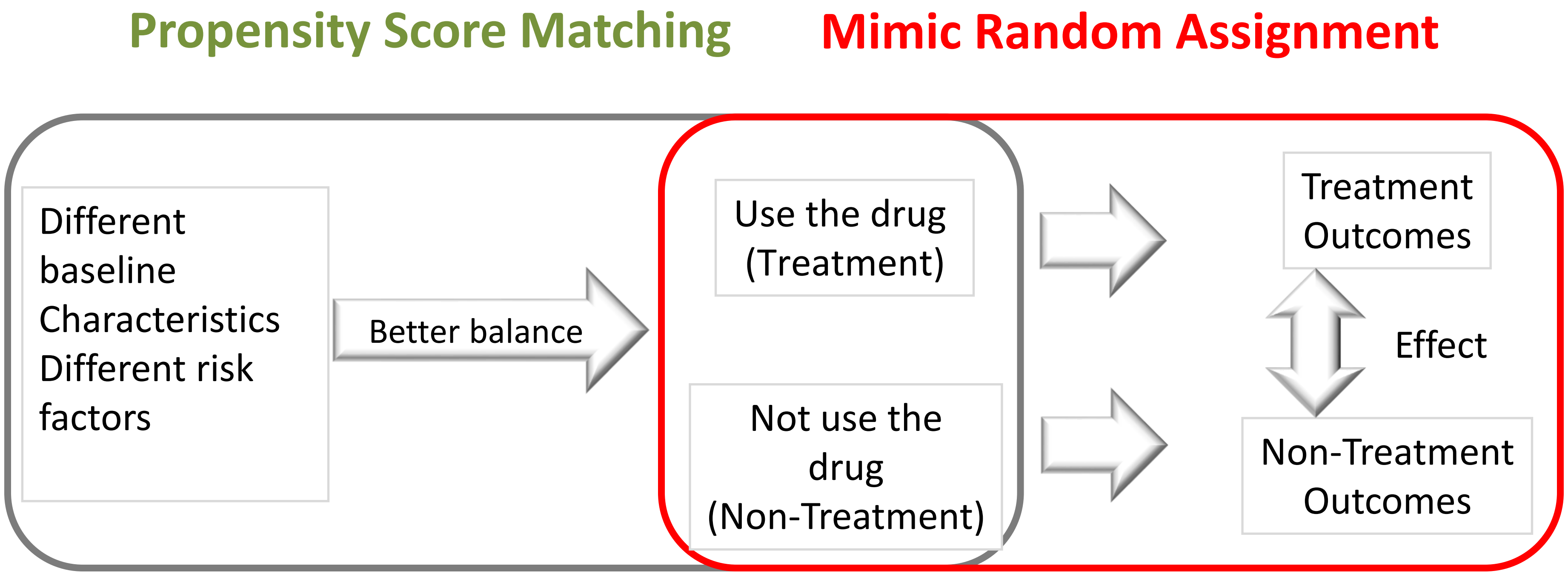
Measurements

- **<u>Exposures</u>**
  - SGLT2 inhibitors use was extracted from the Medicare Part D event file, including canagliflozin, dapagliflozin, empagliflozin, and ertugliflozin
    - SGLT2 inhibitors initiated or not
    - For duration of SGLT2 inhibitor use, only duration before cancer diagnosis was considered
- **<u>Outcomes</u>**: All cause mortality
- **<u>Covariates</u>**
  - Demographic characteristics
    - age, sex, race/ethnicity, and marital status
  - Diabetes related
    - metformin use, dipeptidyl peptidase 4 (DPP4) inhibitors, diabetes duration
  - Cancer characteristics and treatment
    - stage, treatment (surgery, chemotherapy and radiation),
  - Comorbidities
    - Kidney diseases, hypertension, cardiovascular disease
    - Cirrhosis, hepatitis B virus infection, hepatitis C infection, alcohol-related diseases

Statistical analysis

- Multivariate Cox proportional hazards regression models
  - Survival time was measured as months from date of cancer diagnosis to death, or the end of 2019 if the patients survived
- Subgroup analyses according to covariates
- Several sensitivity analyses
  - Use active control
    - DPP-4 inhibitor new use as an active comparator
    - Matched the distance between the date of HCC diagnosis and initiation of the two classes of drugs to minimize possible immortal time bias
  - Propensity score matching

Propensity Score Methods | Estimating Causal Effect



PS estimation

- **Logistic regression**: estimating the conditional probability of assignment to treatment group given observed covariates
  - Most widely used

$$\text{logit}\,(W_i = 1 \mid X_i = x_i) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i$$

where k is the number of covariates; w denotes the binary treatment conditions

- **Machine learning methods**, such as classification trees, boosting, bagging, random forests

  Non-parametric estimation method

  Do not rely on functional form assumption

Table 1. Patients' characteristics by SGLT2 inhibitor use status among 3,185 hepatocellular carcinoma patients

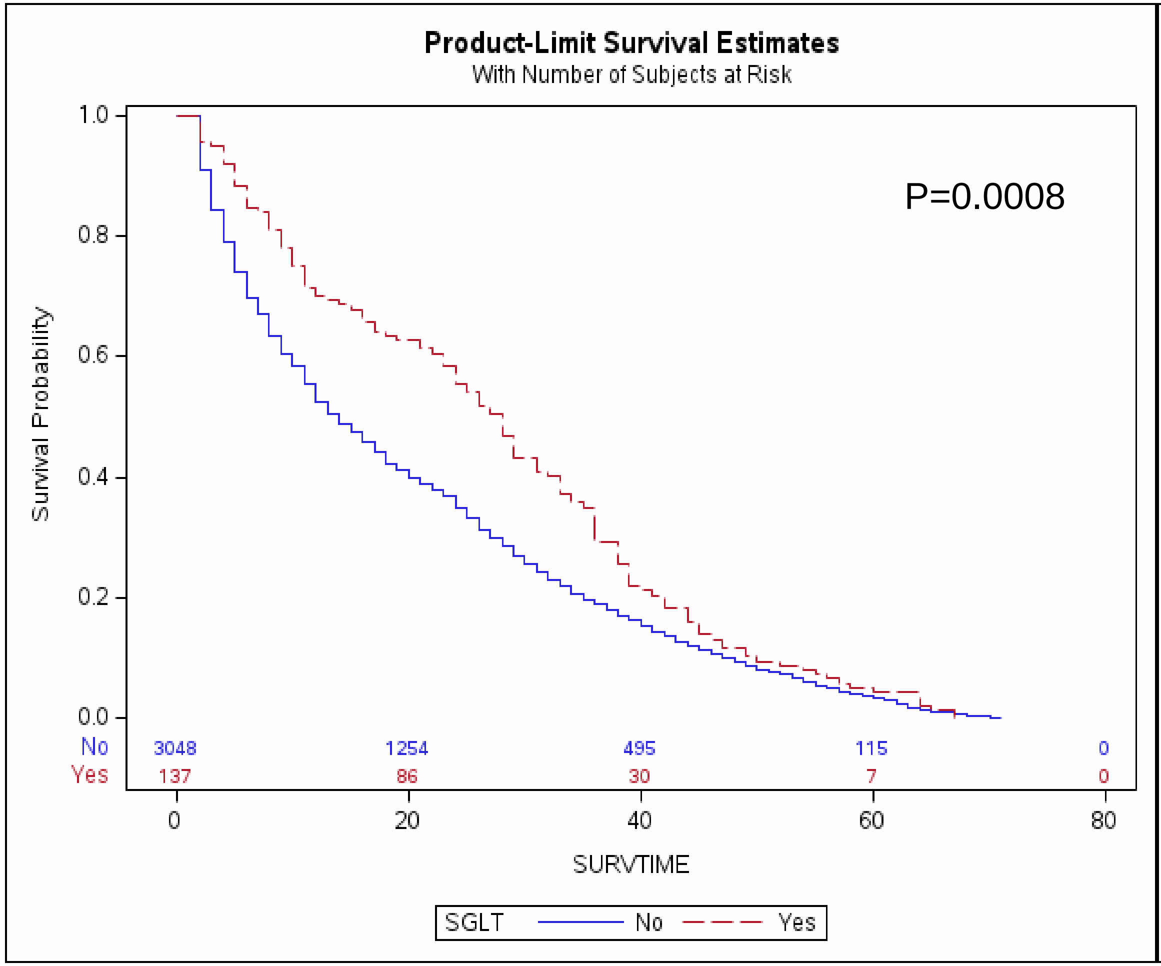
	Use SGLT inhibitor use N (%)			
	Total (3185)	No (3048)	Yes (137)	P value
Age at diagnosis (years), mean (std)	74.8 (6.48)	74.9 (6.52)	72.5 (5.20)	<0.0001
Survival time, mean (std)	20.4 (17.6)	20.1 (17.5)	27.1 (17.4)	<0.0001
Diabetes duration (years), mean (std)	7.7 (4.81)	7.8 (4.80)	7.2 (5.00)	0.19
Sex				0.93
Male	2175 (68.29)	2081 (68.27)	94 (68.61)	
Female	1010 (31.71)	967 (31.73)	43 (31.39)	
Race/ethnicity				0.27
Non-Hispanic White	1624 (50.99)	1555 (51.02)	69 (50.36)	
Non-Hispanic Black	311 (9.76)	-	-	
Asian or Pacific Islander	518 (16.26)	488 (16.01)	59 (11.11)	
Hispanic/Latino	684 (21.48)	656 (21.52)	52 (9.79)	
Others	48 (1.51)	-	-	
Marital Status				0.18
No	1004 (32.72)	1004 (32.94)	38 (27.74)	
Yes	1205 (37.83)	1143 (37.50)	62 (45.26)	
Missing	938 (29.45)	901 (29.56)	37 (27.01)	
Chronic conditions				
Chronic kidney disease (yes, %)	1395 (43.80)	1323 (43.41)	72 (52.95)	0.03
Hypertension (yes, %)	2957 (92.84)	2833 (92.95)	124 (90.51)	0.28
Cardiovascular disease (yes, %)	2143 (67.28)	2053 (67.36)	90 (65.69)	0.69
Cancer stage				0.17
Localized	1773 (55.67)	1685 (55.28)	88 (64.23)	
Regional	824 (25.87)	797 (25.98)	32 (23.36)	
Distant	341 ((10.71)	-	-	
Missing	247 (7.76)	-	-	
Cancer treatment				
Surgery (yes, %)	776 (24.36)	726 (23.82)	50 (36.50)	0.0007
Chemotherapy (yes, %)	1274 (40.00)	1225 (40.19)	49 (35.77)	0.30
Radiation (yes, %)	807 (25.34)	767 (25.16)	40 (29.20)	0.29
Metformin use (yes, %)	1477 (46.37)	1374 (45.08)	103 (75.18)	<0.0001
Hepatitis C virus infection (yes%)	823 (25.84)	789 (25.89)	34 (24.82)	0.78
Hepatitis B virus infection (yes, %)	605 (19.00)	583(19.13)	22 (16.06)	0.37
Alcohol-related diseases (yes, %)	1142 (35.86)	1093 (35.86)	49 (35.77)	0.98
Liver cirrhosis (yes, %)	2183 (68.54)	2089 (68.54)	94 (68.61)	0.99

Results

- Of 3,185 HCC patients with pre-existing diabetes, **137 (4.3%)** patients used SGLT2 inhibitors.

- Over an average of **20.4 months** of follow-up, there were **2,426 (76.2%)** patients who died in total.
  - 78 patients died out of 137 who used SGLT2 inhibitors
    - Crude mortality among users: 56.9%
  - 2,348 died out of 3,048 patients who did not use SGLT2 inhibitors
    - Crude mortality among non-users: 77.0%

Kaplan-Meier curves



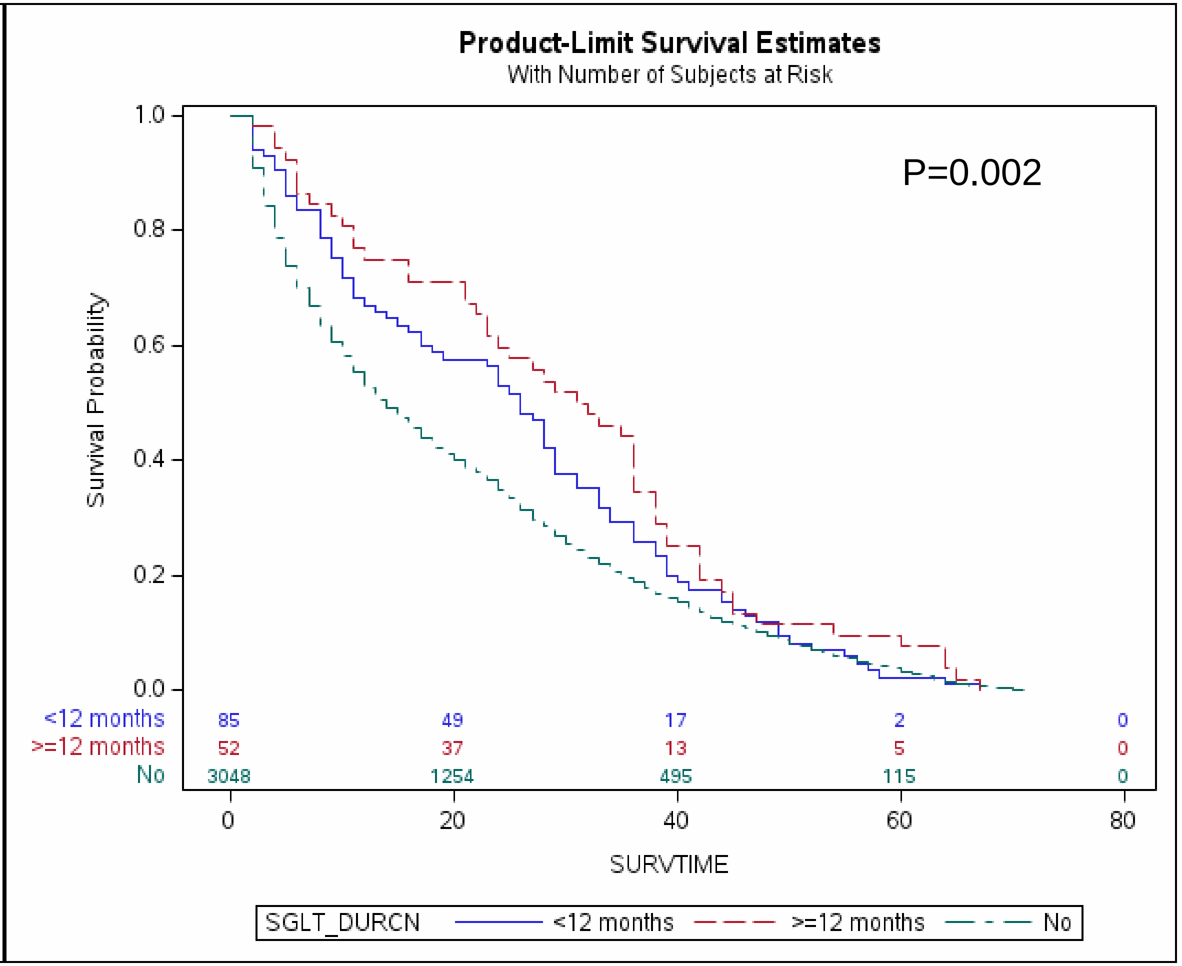
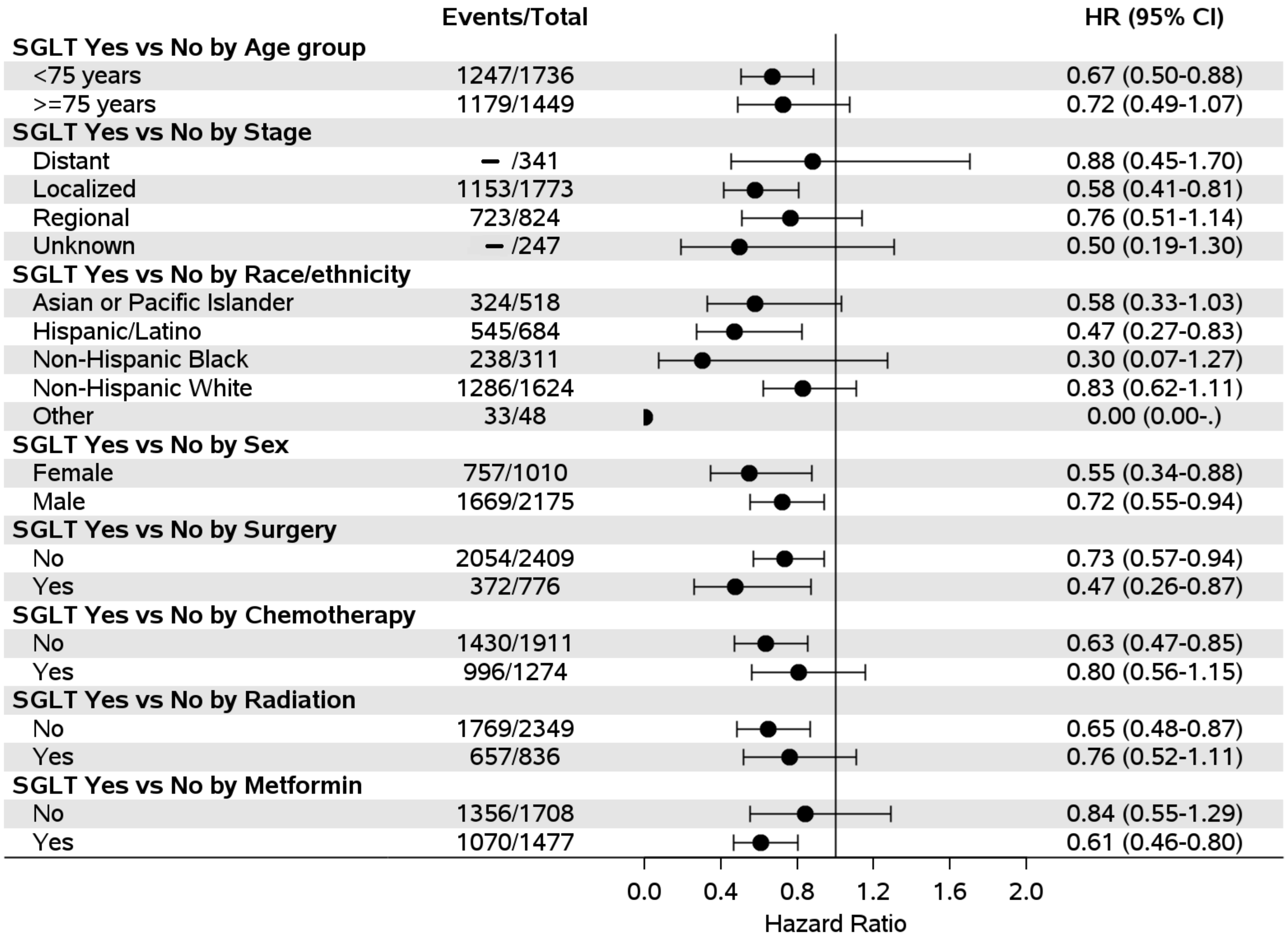


Table 2. Associations between sodium-glucose cotransporter 2 (SGLT2) inhibitor use and mortality among 3,185 hepatocellular carcinoma patients with T2DM

	Number of deaths /Number of patients	HR (95% CI) *
SGLT2 inhibitor use		
No	2348/3048	Reference
Yes	78/137	0.68 (0.54-0.86)
Duration of use		
<12 months	51/85	0.73 (0.55-0.97)
≥ 12 months	27/52	0.60 (0.41-0.88)

*Models were adjusted for age at diagnosis, sex, race/ethnicity, marital status, chronic conditions (chronic kidney disease, hypertension, cardiovascular disease), cancer stage, cancer treatment (cancer-directed surgery, chemotherapy, radiation), diabetes duration (<5 years, 5-<10 years, 10 or more years), metformin use, hepatitis C virus infection, hepatitis B virus infection, alcohol-related diseases, and cirrhosis.



Association between SGLT2 inhibitor use and liver cancer mortality by covariates

Table 3. Associations between sodium-glucose cotransporter 2 (SGLT2) inhibitor use and mortality among hepatocellular carcinoma patients with T2DM*

	Sensitivity analysis 1		Sensitivity analysis 2		Sensitivity analysis 3	
	Deaths /Patients	HR (95% CI)	Deaths /Patients	HR (95% CI)	Deaths /Patients	HR (95% CI)
SGLT2 inhibitor use						
No	86/137	Reference	88/137	Reference	77/122	Reference
Yes	78/137	0.50 (0.37-0.69)	78/137	0.78 (0.57-1.08)	74/122	0.78 (0.55-1.10)
Duration of use						
<12 months	51/85	0.52 (0.37-0.75)	51/85	0.89 (0.62-1.29)	48/67	0.86 (0.58-1.29)
≥ 12 months	27/52	0.47 (0.30-0.74)	27/52	0.63 (0.40-0.998)	26/55	0.66 (0.41-1.07)

*Models were adjusted for age at diagnosis, sex, race/ethnicity, marital status, chronic conditions (chronic kidney disease, hypertension, cardiovascular disease), cancer stage, cancer treatment (cancer-directed surgery, chemotherapy, radiation), diabetes duration (<5 years, 5-<10 years, 10 or more years), metformin use, hepatitis C virus infection, hepatitis B virus infection, alcohol-related diseases, and cirrhosis.

Sensitivity analysis 1 – **employed propensity score** among SGLT2i users and non-users and then matched the two groups by propensity score with a caliper of 0.04 standard deviation of propensity score.

Sensitivity analysis 2 – employed propensity score among **SGLT2i users and DPP4i users** and then matched the two groups by propensity score with a caliper of 0.12 standard deviation of propensity score.

Sensitivity analysis 3 – employed propensity score among SGLT2i users and DPP4i users and then matched the two groups by propensity score with a caliper of 0.12 standard deviation of propensity score and **distance between HCC diagnosis and initiation** of the two compared drugs (± 3 months).

Strengths and limitations

- **Strengths:** a large, population-based, nationally representative database with a rich source for comorbidity data, tumor characteristics and cancer treatments.
- **Limitations:**
  - Only included patients who were 66 years or older
  - Did not account for changes in therapy

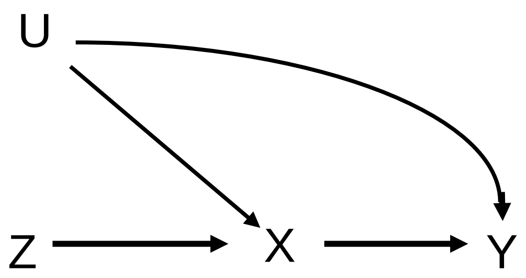
  - Unable to adjust for important lifestyle factors
  - Confounding by indication

→ **Unmeasured confounding issue**

Instrumental variable (IV) analysis

- A method widely used in econometrics and social sciences, to account for unmeasured confounding
- IV analysis relies on the use of instrumental variables and requires that an IV is available.
- Three assumptions hold:
  - I. IV is associated with the exposure
  - II. IV affects the outcomes only through the exposure
  - III. The association between the IV and the outcome is unconfounded

Randomized Controlled Trial



In RCTs, assignment Z, actual treatment X, outcome Y, and unmeasured factors U.

Z is an instrument, because

(i) trial participants are more likely to receive treatment if they were assigned to treatment.

(ii) Z affects the outcome Y only through X (or no direct effect of Z on Y)

(iii) Z does not share common causes with the outcome Y.

* IV mimics the role played by randomization in an RCT

One challenge in IV analysis

- Choose a good IV in a real clinical study
- A weak IV may lead to a treatment estimate with large variance and sensitive to a slight departure from the three IV assumptions.

What kinds of variables make for good instruments?

Instrument Type	Instrument	Treatment -> Outcome
Genetic variants	SNPs	Mendelian Randomization analysis
Distance	Distance to nearest hospital with disease of interest[1]	Acute myocardial infarction (AMI) treatment -> mortality
Physician Preference	Prescribing MD’s preference for conventional or atypical antipsychotics, as indicated by most recent new Rx[2]	Antipsychotic medication type -> mortality
Geography	Regional catheterization rate[3]	Invasive cardiac management -> AMI survival

[1]McClellan 1994; [2]Wang 2005; [3]Stukel 2007

Two-stage estimation

- In the first stage, a regression model is fitted for the exposure (X), using L and Z as regressors (*where L – known confounders, Z – IV*).
- In the second stage, another regression model is fitted for the outcome (Y), using L and $\hat{X}$ as regressors
- Two-stage estimation
  - Computationally simple
  - Gives unbiased estimates in linear models
  - Gives biased estimates in nonlinear models

R codes for G-estimation with a time-to-event outcome

>library(ivtools)

>library(survival)

>data(VitD)

G-estimation requires one model for Z regressed on L, and one Cox model of T regressed on L, Z, and X (exposure)

>fitZ.L <-glm(formula= Z ~ L, family=“binominal”, data=VitD)

>fitT.LZX <- coxph(formula=surv(time, death) ~ L+Z+X, data=VitD)

>fitIV_g <- ivcoxph(estmethod=“g”, X=“X”, fitZ.L=fitZ.L, fitT.LZX=fitT.LZX, data=VitD)

What Can Go Wrong In the Use of IV

- **Failure of the assumptions!**
  - *Z* influences *Y* though other variables than *X*
  - *Z* has only a weak relation to *X*
  - But also, IV is a large sample procedure, even when assumptions are met it is only guaranteed to be unbiased in *large* samples
  - Assumption that the instrument is uncorrelated with error term in the outcome equation is untestable (though aspects can be tested if extra instruments available).

Exogenous, but Weak Instruments

- Weak instruments (i.e., those that are weakly correlated with treatment) can accentuate bias and provide unreliable estimates
- Rule of thumb to check if an instrument is weak:
  - From 1st stage of 2SLS, compute the F-statistic testing the hypothesis that the instrument's coefficient equals 0
  - "Rule of Ten": F-statistic > 10 indicates the instrument isn't weak
  - In general, more instruments increases the relevance of the instrument set (increases the first stage F)
  - But, too many instruments increases small sample bias (compared to few instruments)
  - In general it is best to have as few instruments as possible, and for them to be strongly correlated with X (the endogenous variable)

In Summary

- SGLT2 inhibitor use was associated with improved overall survival of HCC patients with pre-existing diabetes.
- Further studies are needed to confirm our findings and elucidate the possible mechanisms behind the association.

In Summary: challenges of using real-world data

- Confounding
  - Measured
  - Unmeasured
- Data quality: may be incomplete or inaccurate or inconsistencies
- Biases
  - Selection bias: patients with certain characteristics or conditions may be overrepresented
  - Measurement errors: errors in drug exposure measurement, health outcome assessment, or other variables
  - Other biases, such as survivorship bias, immortal time bias, and so on

In Summary: approaches to control for confounding

- RCTs – gold standard for eliminating confounding
- Measured confounding
  - Multivariable-adjusted regression analysis
  - Propensity score methods
- Unmeasured confounding
  - Instrumental variable analysis

Thanks for your attention!

Any questions?