

Statusartikel

Kritisk læsning af artikler om kohortestudier

Lars Christian Lund¹, Anton Pottgård¹, Henrik Toft Sørensen² & Jesper Hallas^{1, 3}

1) Klinisk Farmakologi, Farmaci og Miljømedicin, Syddansk Universitet, 2) Klinisk Epidemiologisk Afdeling, Aarhus Universitetshospital og Aarhus Universitet, 3) Farmakologi, Odense Universitetshospital

Ugeskr Læger 2024;186:V06230401. doi: 10.61409/V06230401

HOVEDBUDSKABER

- De hyppigste fejlkilder i kohortestudier er selektionsbias, informationsbias og konfounding.
- Active comparator, new user-designet reducerer risikoen for de hyppigste fejlkilder i kohortestudier.
- Læsere af et kohortestudie bør være opmærksom på risikoen for tidsrelaterede bias.

I denne artikel vil vi vejlede læseren i, hvordan man kritisk gennemgår et moderne kohortestudie. Vi tager udgangspunkt i et konkret klinisk relevant forskningsspørgsmål og gennemgår et studie, der søger at besvare dette spørgsmål. Epidemiologiske fagtermer forklares ikke løbende, men er beskrevet i Tabel 1.

TABEL 1 Terminologi.

Term	Betydning
Followup	Den mængde tid, som indgår i analysen, og hvor en person potentielt kan opleve det udfald, der undersøges
Konfounding	En forveksling af årsager til, at et givet resultat observeres
Bias	Alle forhold i forskningsprocessen, der medfører fund, der afviger systematisk fra sandheden
Intern validitet	I hvilket omfang resultaterne for en given undersøgelse reflekterer den sande situation i studiepopulationen
Ekstern validitet	I hvilket omfang resultaterne er gældende for andre populationer
Matching	Pardannelse af personer med ens karakteristika. Typisk vil der dannes par mellem eksponerede og ueksponerede personer i et kohortestudie

Randomiserede forsøg har vist lavere niveau af serumurat blandt brugere af SGLT2-hæmmere ift. placebo [1]. Der er en plausibel mekanisme [2], øget diurese og dermed øget udskillelse af urat, men forsøgene har ikke kunnet vise, at reduktionen af urat også nedsætter forekomsten af urinsyreigt [3]. Da patienter med type 2-diabetes er i øget risiko for at få urinsyreigt [4], kunne en beskyttende effekt over for urinsyreigt i nogle tilfælde være udslagsgivende ved præparatvalg.

Ideelt set burde spørgsmålet besvares ved et randomiseret forsøg. Urinsyreigt er ikke en hyppig sygdom, og dette studie ville derfor blive prohibitivt ressourcekrævende. Alternativt kunne spørgsmålet besvares med et studie, hvor man sammenligner brugere af SGLT2-hæmmere med ikkebrugere, uden at randomisere. Men hvordan kan vi vide, om et sådant studie er retvisende?

FORMULERING AF FORSKNINGSSPØRGSMÅL

Til formulering af forskningsspørgsmålet kan population, intervention, comparator, outcome (PICO)-strukturen anvendes: P: personer med type 2-diabetes; I: SGLT2-hæmmere; C: glukagonlignende peptid 1 (GLP-1)-receptoragonister eller dipeptidylpeptidase-4-hæmmere; O: urinsyreigt.

Studieeksempel

Ovennævnte problemstilling blev undersøgt af *Fralick et al* [5]. I et kohortestudie sammenlignede man forekomsten af urinsyreigtdiagnoser blandt brugere af SGLT2-hæmmere og GLP-1-receptoragonister (GLP-1-

RA) med type 2-diabetes. Grupperne blev fulgt fra dagen, hvor de indløste deres første recept for hhv. SGLT2-hæmmere eller GLP-1-RA. Personer med tidligere urinsyreigt blev ekskluderet. Brugere af SGLT2-hæmmere og GLP-1-RA kan dog ikke forventes at være umiddelbart sammenlignelige. F.eks. foretrækkes GLP-1-RA ved overvægt, mens SGLT2-hæmmere foretrækkes ved nedsat nyrefunktion. For at tage højde for disse forskelle blev der anvendt propensity score matching [6]. Raten af urinsyreigt i grupperne blev sammenlignet vha. Cox-regressionsanalyse, hvorved man fandt en 36% lavere rate blandt brugere af SGLT2-hæmmere end blandt brugere af GLP-1-RA.

CENTRALE STUDIEKOMPONENTER

Active comparator new user-designet

I moderne farmakoepidemiologiske kohortestudier anvendes ofte active comparator new user-designet [7]. Brugen af en aktiv komparator (fremfor sammenligning med ikkebrug af lægemidlet) er begrundet i, at man mindsker confounding by indication (Tabel 2). I eksemplet ville ikkebrugere hovedsageligt udgøres af personer uden diabetes og ville således afvige fra brugerne af SGLT2-hæmmere på faktorer såsom en lavere forekomst af overvægt og hjertesygdom. Dette ville medføre en lavere forekomst af urinsyreigt i kontrolgruppen og maskere en beskyttende effekt af SGLT2-hæmmere. Hvis man derimod sammenligner brugere af SGLT2-hæmmer med brugere af et andet antidiabetisk lægemiddel, opnår man en sammenlignelig fordeling af prognostiske faktorer grupperne imellem. En forudsætning er dog, at det andet diabetesmiddel ikke i sig selv påvirker risikoen for urinsyreigt.

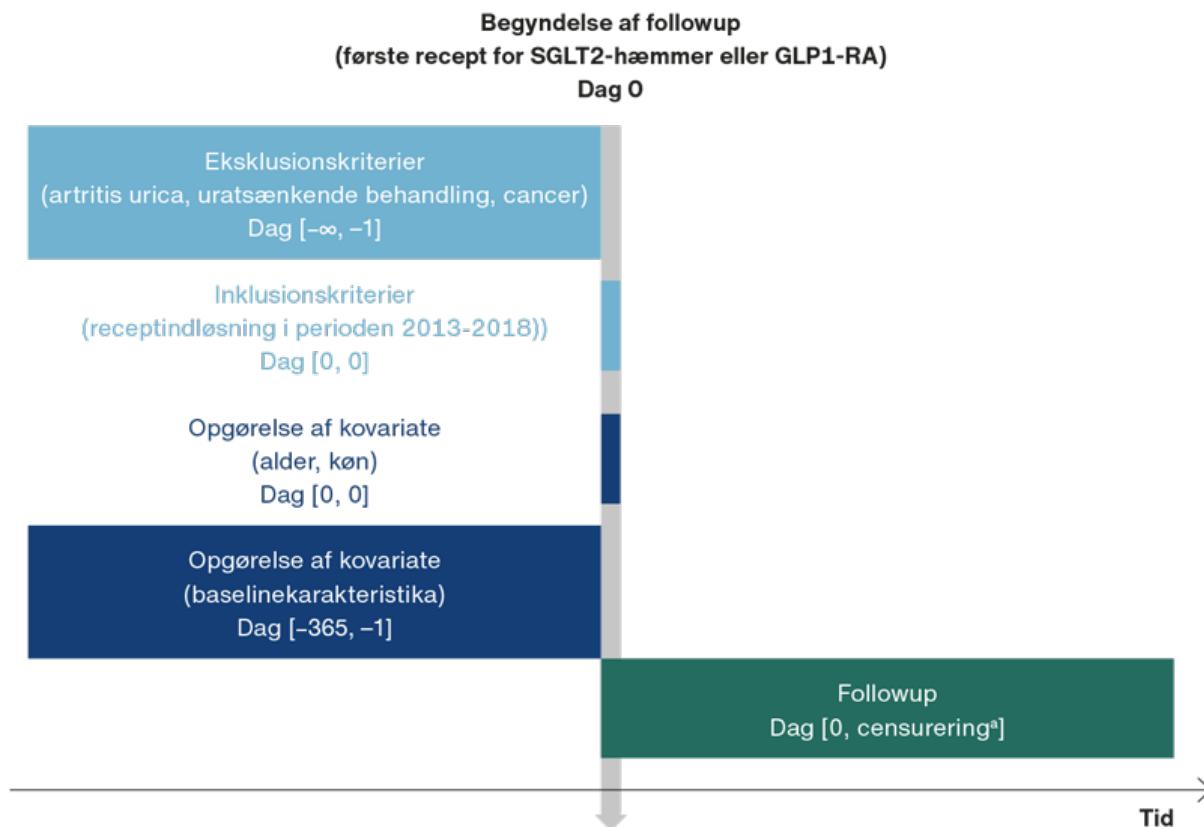
TABEL 2 Hyppige designproblemer i kohortestudier.

Problem	Definition	Eksempel	Konsekvenser for resultaterne	Løsning
Confounding by indication	Indikationen for en eksponering modificerer sandsynligheden for udfaldet, der undersøges, og denne effekt tilskrives eksponeringen	Brugere af efedrinbaserede slankemidler vil hyppigere opleve AMI end ikke-brugere, i kraft af at de hyppigere er overvægtige	En falsk (eller forstærket) association mellem lægemiddel og udfald	Brug af en aktiv komparator eller på anden vis mere sammenlignelig kontrolgruppe
Protopatisk bias	Eksponeringen er en konsekvens af et tidligt symptom på outcome, som endnu ikke er klinisk erkendt, og effekten tilskrives fejlagtigt eksponeringen	Personer, der starter behandling med protonpump hæmmere, får flere coloncancer diagnosticeret end ikkebrugere, fordi coloncancer kan give uspecifikke smerter i epigastriet, før canceren er erkendt, og at disse smerter behandles med protonpump hæmmere	Man vil finde en øget forekomst af gastrointestinal kræft blandt patienter, der påbegynder behandling med protonpump hæmmere	Den allerede eksisterende, men udiagnosticerede kræftsygdom vil typisk diagnosticeres inden for kort tid – kortere end biologisk plausibelt, hvis denne skulle være forårsaget af behandlingen. Løsningen er derfor at udskyde starten af followup svarende til det korteste interval, hvor man kan forvente en biologisk plausibel effekt
Selektionsbias	Ved at begrænse populationen til en given gruppe vil forekomsten af udfaldet være anderledes i denne gruppe end for den resterende (klinisk relevante) population	Der sammenlignes personer, der påbegynder en ny slags menopausal hormonerapi, med eksisterende brugere af en anden menopausal hormonerapi mhp. forekomsten af venøs trombose. Kun personer, der ikke tidligere har haft venøs trombose, kan indgå i studiet	Eksisterende brugere af anden menopausal hormonerapi er selekteret ved, at ikke har haft venøs trombose. De eksisterende brugere er i lavere risiko for at udvikle venøs trombose, da de ellers ville have oplevet dette tidligere. Dette vil give falsk forhøjet effekt estimerer for det nye lægemiddel	I dette eksempel ville sammenligning med nye brugere løse problemet
Misklassifikation	Fejlagtig klassifikation af en given variabel, oftest eksponering eller udfald	Man ønsker at undersøge sammenhængen mellem brug af NSAID og akut myokardieinfarkt. Brug af NSAID defineres som en indløsning af en NSAID-recept i de 30 dage op til et myokardieinfarkt. Nogle personer, der indløser en recept i dette tidsrum, har alligevel ikke brugt det i dagene op til udfaldet, og andre kan have indtaget NSAID, selvom de ikke har indløst en recept. Da data om eksponeringen indsamles inden udfaldet, er misklassifikationen defintorisk nondifferentiel	Konsekvenser afhænger af, om misklassifikationen er differentiell eller nondifferentiel, dvs. om misklassifikationen er ens for de personer, der har haft udfaldet, og dem, der ikke har. Ved nondifferentiel misklassifikation vil effektestimater forskydes tættere på 1, ved differentiell misklassifikation kan det forskydes både mod 1 eller væk fra 1	Misklassifikation er nærmest umulig at undgå i kliniske studier. Ønsker man at begrænse misklassifikation, kræver det manuel dataindsamling, at data beriges fra andre kilder end blot registerdata (f.eks. journal gennemgang), eller at der udføres et valideringsstudie
Justering for en mediator af effekten	Effektophævelse grundet, at man korrigerer for en proxy af det udfald, man ønsker at undersøge	Man ønsker at undersøge effekten af behandling med olanzapin på risikoen for kardiovaskulær sygdom efter 5 år. Der sammenlignes med et metabolisk neutralt antipsykotikum. For at gøre grupperne mere sammenlignelige justeres analysen for niveauet af lavdensitets lipoprotein 6 måneder efter start af followup	Hvis olanzapin øger risikoen for kardiovaskulær sygdom via et øget niveau af lipider i blodet, vil man fjerne en del af lægemiddeleffekten ved justering og få et falsk lavt resultat	Denne problematik undgås ved, at der udelukkende justeres for karakteristika ved initiering af behandling (baselinekonfundere). Justering for faktorer, der ændrer sig over tid, kan være relevant (tidsvarierende kovariate), men de må ikke være en del af årsagskæden
Immortal time bias	Fejlagtig klassifikation af followuptid betinget på fremtidige hændelser	Der ønskes undersøgt, hvorvidt behandling med SGLT2-hæmmere reducerer risikoen for genindlæggelse blandt patienter, der udskrives efter indlæggelse for hjertesvigt. Hvorvidt der påbegyndes behandling med SGLT2-hæmmere for patienterne, vurderes i de første 3 måneder efter udskrivelsen, og al persontid klassificeres ifølge dette. Followup slutter ved eventuel genindlæggelse	Grundet klassificeringen af persontid, der ligger før påbegyndelsen af SGLT2-hæmmere som eksponeret, er det hermed sikret, at den (senere hen) eksponerede person ikke er genindlagt op til receptindløsningen. Dette reducerer raten blandt eksponerede og giver falsk lave effekt-estimerer, dvs. falsk beskyttende effekt	Al persontid skal klassificeres på baggrund af den aktuelle eksponeringsstatus, ikke en fremtidig eksponering. Alternativt kan al followup før slutningen af perioden, hvor eksponeringen vurderes, ekskluderes fra analysen

AMI = akut myokardieinfarkt; NSAID = nonsteroidale antiinflammatoriske stoffer.

New user-elementet indebærer, at man kun inkluderer nye brugere af lægemidlerne og følger dem, fra de påbegynder brugen af lægemidlet. Dette kan illustreres vha. p-piller og risikoen for venøs tromboemboli. Personer, der har en høj risiko for at få denne bivirkning, vil ofte få den tidligt i deres forløb, og efterhånden vil der fortrinsvis være kvinder i kohorten med lav risiko, da de har tålt lægemidlet i længere tid. Sammenlignes risikoen ved et nyt præparat, hvor alle er nye brugere, med risikoen ved et gammelt præparat, hvor der kun er personer tilbage, der tåler behandlingen, vil det nye præparat fejlagtigt fremstå som havende den største risiko. Ovenstående studiedesign er illustreret i **Figur 1**, som demonstrerer den tidsmæssige rækkefølge af elementerne i *Fralick et al* [5], i overensstemmelse med principperne beskrevet af *Schneeweiss et al* [8].

FIGUR 1 Designdiagram for studiet af *Fralick et al* [5].



GLP1-RA = glukagonlignende peptid 1-receptoragonist.

a) Censurering grundet, at personen ikke længere kan observeres i database, død eller første tilfælde af urinsyregigt.

Persontidsbaserede analyser

Fralick et al målte raten af urinsyregigtsdiagnoser blandt brugere af SGLT2-hæmmere hhv. GLP-1-RA. Rater er forekomsten af et udfald over et bestemt tidsrum divideret med mængden af den opfølgningstid, hvor udfaldene forekom [9]. Hvis der er tale om nye (incidente) tilfælde, kaldes raten en incidensrate. Rater kan beregnes, selvom nogle personer kun er fulgt i få måneder, mens andre er blevet fulgt i flere år. Dette står i modsætningen til en kumuleret incidensproportion, oftest kaldet risiko, hvor alle deltagere skal følges lige lang tid (f.eks. risikoen for urinsyregigt inden for seks mdr. efter påbegyndelse af behandling). Ved at foretage den simple beregning »rate lig med antallet af hændelser delt med persontid«, antager man, at raten er konstant over den indsamlede persontid – noget, der i mange situationer ikke er gældende [10]. En hyppig måde at undgå denne begrænsning er ved brug af en Cox-regressionsanalyse [11]. Her beregnes relative forskelle i raten mellem to grupper, en hazard ratio, hvor raten frit kan variere over den indsamlede persontid. Det kræver dog en antagelse om, at effekten af en given eksponering er konstant over tid.

Håndtering af konfounding

I observationelle studier vil de grupper, man ønsker at sammenligne, som regel ikke være sammenlignelige. I *Fralick et al*-studiet har brugere af GLP-1-RA en højere forekomst af nyresygdom og brug af diuretika, begge prognostiske faktorer for at udvikle urinsyregigt. En direkte sammenligning af de observerede rater vil fejlagtigt tilskrive effekter fra disse faktorer til SGLT2-behandlingen, dvs. der er tale om konfounding eller på dansk:

årsagsforveksling. Konfounding kan håndteres i design- eller analysefasen. Matching af brugere af SGLT2-hæmmere til lignende brugere af GLP-1-RA, f.eks. på baggrund af alder, køn, forekomsten af nyresygdom og brug af diuretika, kan gøres i designfasen. Man vil opnå mindre grupper, der dog vil være mere sammenlignelige. En betydelig fordel ved denne metode er, at sammenligneligheden kan vises direkte. En anden udbredt måde er ved anvendelse af regressionsmodeller, som f.eks. Cox-regressionsanalyse. Her justeres (læs: ophæves) effekten af faktorer, der disponerer for udfaldet, der undersøges.

KRITISKE SPØRGSMÅL

Når et observationelt studie viser en sammenhæng, er det centrale spørgsmål, om denne vurderes at afspejle en årsagssammenhæng, dvs. kausalitet. For at vurdere, om en sammenhæng er kausal, må man overveje, om sammenhængen alternativt kunne være forårsaget af bias, herunder konfounding, eller statistiske tilfældigheder. Bias defineres som forhold i forskningsprocessen, der medfører fund, der afviger fra sandheden. Bias kan opstå enten i udvælgelse af studiepopulationen eller bortfald under opfølgningen, dvs. selektionsbias, ved fejlbehæftet måling af eksponering eller effektvariable, dvs. informations- eller misklassifikationsbias, eller ved konfounding (se ovenfor).

Det er vigtigt at være opmærksom på, at kliniske målinger og alle registerdata er behæftet med en vis grad af misklassifikation. I mange studier bestemmes eksponering på baggrund af indløste recepter fra Lægemiddelstatistikregisteret og sygdomsudfald på baggrund af diagnoser fra Landspatientregisteret. Den indløste recept afspejler ikke den eksakte indtagelse af lægemidlet og tager f.eks. ikke højde for lav adhærens. Diagnoser i Landspatientregisteret kan være behæftet med fejl i de kliniske diagnoser, egentlige kodningsfejl og manglende specificitet af diagnosekoderne. Selv nogle af de mest valide diagnoser i Landspatientregisteret udviser en ikkenegligabel grad af misklassifikation [12]. Hvis denne misklassifikation er uafhængig af andre variable (også kaldet tilfældig misklassifikation), vil det medføre, at man underestimerer den reelle effekt [13].

Det er vores erfaring, at de fleste tekniske problemer i kohortestudier kan henføres til en af disse tre kategorier: selektionsbias, informationsbias eller konfounding. Der findes dog andre tekniske fejlkilder (Tabel 2), og man bør altid forholde sig til de ikketekniske problemer i studiets kontekst, tolkning, generaliserbarhed og valg af design.

ANDRE DESIGNVARIANTER

Kohortestudier kan give stærkt misvisende resultater, hvis man i deres implementering bryder de grundlæggende epidemiologiske principper. Et af disse lyder, at data skal analyseres i den rækkefølge, de blev indsamlet, dvs. man må ikke betinge på fremtidige begivenheder. Et eksempel på en sådan bias er immortal-time bias [14]: I et studie ønsker man at undersøge overlevelsesgevinsten ved behandling med pladehæmmere efter et akut myokardieinfarkt. Man følger derfor folk, fra de udskrives efter deres infarkt, inddeler dem efter, om de inden for 90 dage efter udskrivelse indløser en recept på pladehæmmer eller ej, og følger dem fra udskrivelsen og et år frem. Her vil man, uafhængigt af lægemiddeleffekten, finde, at lægemidlet er associeret med en reduceret risiko for død.

Dette skyldes, at personer, der overlever frem til receptindløsningen, bringer »udødelig« eksponeret tid med ind i analysen – tiden fra udskrivelse og frem til indløsningen. Tiden er udødelig, da den klassificeres på basis af fremtidige begivenheder (receptindløsningen). Den korrekte håndtering er at klassificere denne persons tid frem til indløsningen som ikkebrugertid, eller at alle personer i studiet først følges fra dag 91, dvs. efter den periode, der bruges til at identificere brug og ikkebrug af lægemidlet. Problemer med tidsrelateret bias er hyppige i kohortestudier [15], og den kritiske læser bør forholde sig til, om der forekommer followup, der

klassificeres på baggrund af fremtidige begivenheder.

TOLKNING

Generaliserbarhed

En afgørende styrke af de observationelle studier sammenlignet med lodtrækningsforsøg er, at studiepopulationen afspejler patienterne, der behandles i klinisk praksis. Selvom en repræsentativ studiepopulation normalt ikke videnskabeligt er det vigtigste af hensyn til håndtering af bias og konfounding, er det vigtigt, at patienter, der inkluderes i studiet, omfatter segmenter af populationen med høj risiko for det udfald, der ønskes undersøgt. Mange lodtrækningsstudier er små, bruger surrogatmarkører eller omfatter ikke de patientgrupper, der har højest risiko for at få given sygdom. Observationelle undersøgelser baseret på registre er derfor et alternativ til lodtrækningsforsøg, da de omfatter store populationer, hvilket gør det muligt at studere sjældne eksponeringer og sygdomme billigt og hurtigt.

Absolutte og relative effektestimater

Kohortestudiers resultater angives som regel enten som et relativt effektestimater, f.eks. en relativ risiko (RR), en hazard ratio (HR) eller et absolut effektestimater, f.eks. en incidensrateredifference (IRD). Det relative effektestimater fortæller, hvor stærk sammenhængen er mellem eksponeringen og outcome, men ikke noget om, hvorvidt effekten vil være klinisk relevant. *Fralick et al* rapporterer en absolut forskel på -3 tilfælde pr. 1.000 personår i en population af diabetespatienter, hvilket næppe i sig selv ville diktere behandling med en SGLT2-hæmmer. Det er vigtigt, at relative effektestimater så vidt muligt kontekstualiseres af absolutte effektestimater, samt at effekter estimeret i én population ikke nødvendigvis kan overføres til en anden.

AVANCEREDE DESIGNVARIANTER

Mange kohortestudier anvender propensity score-baserede metoder, især når et lægemiddel undersøges. En propensity score er en beregnet sandsynlighed for, at en given person er eksponeret [16] baseret på målte patientkarakteristika og bruges til at modvirke konfounding. Propensity scores bruges hyppigst til at matche individer en til en og på den måde opnå to grupper med ensartede fordelinger af prognostiske faktorer. Alternative måder at bruge propensity scores er at stratificere eller skabe to sammenlignelige pseudopopulationer, ved at personer tildeles en given vægt på baggrund af deres propensity score [17].

RAPPORTERINGSLINJER

Kohortestudier kan rapporteres ud fra »The strengthening the reporting of observational evidence statement« [18]. Denne tjekliste på 22 punkter sikrer en grundig gennemgang af metode, resultater og diskussion. Mere detaljerede tjeklister for studier, der gør brug af i forvejen indsamlede data, såsom registerdata, er også blevet udarbejdet i form af »The reporting of studies conducted using observational routinely-collected health data statement« [19].

Herudover bliver det også almindeligt at registrere protokoller for observationelle studier, inden dataanalysen påbegyndes, som det kendes fra lodtrækningsforsøg. Til dette formål kan EU-PAS-registeret [20] eller Real World Evidence Registry [21] anvendes.

Antaget 4. januar 2024

Publiceret på ugeskriftet.dk 26. februar 2024

Interessekonflikter Der er anført potentielle interessekonflikter. Forfatterernes ICMJE-formularer er tilgængelige sammen med artiklen på ugeskriftet.dk

Referencer findes i artiklen publiceret på ugeskriftet.dk

Artikelreference Ugeskr Læger 2024;186:V06230401

doi [10.61409/V06230401](https://doi.org/10.61409/V06230401)

Open Access under Creative Commons License [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

SUMMARY

Critical evaluation of cohort study results

Being able to critically evaluate modern cohort studies is important when being presented with claims based on observational evidence. In this review article, key aspects of the cohort design are presented using an example of a cohort study investigating the association between the use of SGLT2 inhibitors and gout. We describe the active comparator, new user design, modern methods used to address confounding, how to identify the most common sources of bias, and how to interpret study results appropriately.

REFERENCER

1. Zinman B, Wanner C, Lachin JM. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373(22):2117-28. doi: [10.1056/NEJMc1600827](https://doi.org/10.1056/NEJMc1600827).
2. Van Bommel EJM, Muskiet MHA, Tonneijck L et al. SGLT2 inhibition in the diabetic kidney - from mechanisms to clinical outcome. *Clin J Am Soc Nephrol*. 2017;12(4):700-710. doi: [10.2215/CJN.06080616](https://doi.org/10.2215/CJN.06080616).
3. Zhao Y, Xu L, Tian D et al. Effects of sodium-glucose cotransporter 2 (SGLT2) inhibitors on serum uric acid level: a meta-analysis of randomized controlled trials. *Diabetes Obes Metab*. 2018;20(2):458-462. doi: [10.1111/dom.13101](https://doi.org/10.1111/dom.13101).
4. Yang Y, Xian W, Wu D et al. The role of obesity, type 2 diabetes, and metabolic factors in gout: a Mendelian randomization study. *Front Endocrinol (Lausanne)*. 2022;13:917056. doi: [10.3389/fendo.2022.917056](https://doi.org/10.3389/fendo.2022.917056).
5. Fralick M, Chen SK, Paterno E, Kim SC. Assessing the risk for gout with sodium-glucose cotransporter-2 inhibitors in patients with type 2 diabetes: a population-based cohort study. *Ann Intern Med*. 2020;172(3):186-194. doi: [10.7326/M19-2610](https://doi.org/10.7326/M19-2610).
6. Webster Clark M, Stürmer T, Wang T et al. Using propensity scores to estimate effects of treatment initiation decisions: state of the science. *Stat Med*. 2021;40(7):1718-1735. doi: [10.1002/sim.8866](https://doi.org/10.1002/sim.8866).
7. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep*. 2015;2(4):221-228. doi: [10.1007/s40471-015-0053-5](https://doi.org/10.1007/s40471-015-0053-5).
8. Schneeweiss S, Rassen JA, Brown JS et al. Graphical depiction of longitudinal study designs in health care databases. *Ann Intern Med*. 2019;170(6):398-406. doi: [10.7326/M18-3079](https://doi.org/10.7326/M18-3079).
9. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. *Modern epidemiology*. Fourth edition. Lippincott Williams And Wilkins, 2021.
10. Stensrud MJ, Hernán MA. Why test for proportional hazards? *JAMA*. 2020;323(14):1401-1402. doi: [10.1001/jama.2020.1267](https://doi.org/10.1001/jama.2020.1267).
11. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society: series B (methodological)*. 1972;34:187-220. doi: [10.1111/j.2517-6161.1972.tb00899.x](https://doi.org/10.1111/j.2517-6161.1972.tb00899.x).
12. Schmidt M, Schmidt SAJ, Sandegaard JL et al. The Danish national patient registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449-90. doi: [10.2147/CLEP.S91125](https://doi.org/10.2147/CLEP.S91125).
13. Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. *Int J Epidemiol*. 2005;34(3):680-7. doi: [10.1093/ije/dyi060](https://doi.org/10.1093/ije/dyi060).
14. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167(4):492-9. doi: [10.1093/aje/kwm324](https://doi.org/10.1093/aje/kwm324).

15. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf.* 2020;29(9):1101-1110. doi: 10.1002/pds.5083.
16. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. *Biometrika.* 1983;70:41-55. doi: 10.2307/2335942.
17. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ.* 2019;367:l5657. doi: 10.1136/bmj.l5657.
18. Von Elm E, Altman DG, Egger M et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med.* 2007;4(10):e296. doi: 10.1016/S0140-6736(07)61602-X.
19. Benchimol EI, Smeeth L, Guttman A et al. The REporting of studies conducted using observational routinely-collected health Data (RECORD) statement. *PLoS Med.* 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885.
20. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. The European Union electronic Register of Post-Authorisation Studies (EU PAS Register), 2023. <https://www.encepp.eu/encepp/studiesDatabase.jsp> (26. okt 2023).
21. Real World Evidence Registry, RWE. RWE's study registration site, 2023. <https://osf.io/registries/rwe/discover> (26. okt 2023).