

Statusartikel

Ugeskr Læger 2024;186:V06230402. doi: 10.61409/V06230402

Kritisk læsning af artikler om case-kontrol-studier

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, 3) Farmakologi, Odense Universitetshospital

Ugeskr Læger 2024;186:V06230402. doi: 10.61409/V06230402

HOVEDBUDSKABER

- Et case-kontrol-studie er en effektiv sampling af information fra den underliggende kohorte.
- Et korrekt udført case-kontrol-studie giver de samme relative effektestimater som det tilsvarende kohortestudie.
- Hyppige fejlkilder er konfounding, misklassifikation og selektionsbias grundet fejlagtig udvælgelse af kontroller.

Brystkræft er en sjælden bivirkning i forbindelse med længerevarende behandling med menopausal hormonterapi (HT). En konsultation i almen praksis kunne omhandle en kvinde, der henvender sig med perimenopausale gener. Patienten og egen læge bliver enige om, at der er tale om symptomer af en sådan sværhedsgrad, at der er indikation for hormonterapi – men patienten er bekymret for den øgede risiko for brystkræft og spørger: »Hvilken slags hormonterapi indebærer den mindste risiko for at udvikle brystkræft?«

Ovenstående problemstilling kan formuleres vha. PICO-metoden: population: kvinder med perimenopausale gener; intervention: østrogen-, progesteron- eller tibolonholdig hormonterapi; comparator: ikkeanvendelse af hormonterapi; outcome: brystkræft.

Ovenstående spørgsmål er søgt besvaret af *Vinogradova et al* [1]. Der er tale om et case-kontrol-studie baseret på data fra almen praksis [2]. I studiet indgik alle nye tilfælde af brystkræft i en periode på 20 år samt op til fem kontrolpersoner pr. case. Diagnosedatoen for brystkræftdiagnosen for den enkelte case blev anvendt som indeksdato for de kontrolpersoner, som tilhørte den enkelte case. Hvorvidt personerne havde brugt hormonterapi, blev vurderet indtil et år før indeksdatoen. I alt 98.611 kvinder med brystkræft og 457.498 kontrolpersoner indgik i studiet. Herefter blev hyppigheden af hormonterapi sammenlignet blandt cases og kontrolpersoner, og justerede oddsratioer blev beregnet. Herved fandt man en øget relativ risiko for brystkræft ved østrogenmonoterapi (oddsratio 1,15) samt kombinationsbehandling med østrogen og progesteron (oddsratio 1,79).

CENTRALE STUDIEKOMPONENTER

På trods af at case-kontrol-designet traditionelt placeres lavere i evidenshierarkiet [3], er det en misforståelse, at der opnås mindre valide resultater ved udførelse af et case-kontrol-studie end ved et kohortestudie [4]. Et veludført case-kontrol-studie er en effektiv stikprøve af persontiden (opfølgningstiden) i et underliggende kohortestudie og giver derfor de samme resultater på trods af det lavere antal inkluderede personer. Derfor vil

de centrale studiekomponenter i et case-kontrol-studie også overlapper stærkt med komponenterne af et kohortestudie, bortset fra dem der er relateret til udvælgelsen af kontrolpersoner.

Populationen

Det er afgørende, at cases og kontrolpersoner kommer fra den samme underliggende population. En god tommelfingerregel er, at hvis en kontrolperson havde udviklet sygdommen, skulle hun have indgået som case i stedet. Er dette ikke tilfældet, risikerer man, at studiet giver misvisende resultater. I studiet af *Vinogradova et al* blev case-kontrol-studiet foretaget blandt alle kvinder i alderen 50-79 år, som var tilmeldt en almen praksis, som leverer data til en af de to anvendte datakilder.

Stikprøven

Den afgørende forskel mellem et case-kontrol- og et kohortestudie er, at ikke al tilgængelig persontid (opfølgningstiden) i den udvalgte population anvendes i analysen, men kun en stikprøve af persontiden i hele populationen. Det giver de mest valide resultater, hvis personer, der ender med at være cases, også kan udvælges som kontroller, givet at deres indeksdato som kontrolperson ligger forud for tidspunktet, hvor de bliver cases.

Ved den simpleste form af såkaldt density sampling tildeles alle personer fra studiepopulationen en tilfældig indeksdato inden for studieperioden. De personer, der allerede er blevet til cases på det tildelte tidspunkt eller er censureret af andre årsager, udgår. Herefter udvælges en stikprøve af personerne, f.eks. 400 kontrolpersoner, hvis der er 100 cases. Denne 4:1-ratio refereres hyppigt i litteraturen [5], men der kan med fordel vælges et større antal kontrolpersoner [6, 7], f.eks. for at øge den statistiske præcision i subgruppeanalyser. Ved risk set sampling, en variant af density sampling, tildeles alle individer i populationen en indeksdato på baggrund af diagnosedatoen for en konkret case, ofte matchet på fødeår og køn, hvorefter matchede kandidater udvælges tilfældigt.

På denne måde er de udvalgte karakteristika og kalendertid ligeligt fordelt mellem cases og kontrolpersoner. Formålet med denne matching er at øge den statistiske præcision, ikke at ophæve konfounding. Tværtimod risikerer man at indføre selektionsbias, hvis analysen ikke tager højde for de udvalgte matchingvariable [8]. Hvis matching anvendes, anbefales det, at der matches på få stærke konfoundere såsom alder og køn [9]. I hovedparten af nyere case-kontrol-studier anvendes risk set sampling såsom i *Vinogradova et al*. Ud over ovennævnte måder findes der også case-kohorte- og kumulativ case-kontrol-sampling [4]. Disse anvendes sjældent.

Analysen

Strukturen i et case-kontrol-studie er vist i **Tabel 1**, og terminologien herfor er forklaret i **Tabel 2**. Den primære sammenligning er forekomsten af eksponering blandt cases og kontrolpersoner. Hyppigst anvendes en oddsratio til opsummering af denne sammenligning. Oddsrationen beregnes som odds for eksponering blandt cases divideret med odds for eksponering blandt kontrolpersoner og kan tolkes som andre relative effektestimater. Hvis kontrolpersoner er udtrukket ved en form for density sampling, er denne oddsratio et estimat af incidensratioen fra det underliggende kohortestudie.

TABEL 1 Resultaterne fra et case-kontrol-studie opsummeret.

	Cases	Kontrolpersoner
<i>Eksposering</i>		
Eksponeret	a	b
Ueksponeret	c	d
<i>Odds</i>		
Odds _{cases}	= a/c	
Odds _{kontrolpersoner}	= b/d	
Oddsratio	= odds _{cases} /odds _{kontrolpersoner}	

TABEL 2 Hyppigt anvendt terminologi i forbindelse med et case-kontrol-studie.

Term	Forklaring
Indeksdato	For cases oftest den dato, hvor den casedefinerende diagnose stilles For kontrolpersoner enten en tilfældig dato (uparret design) eller samme dato som deres tilsvarende cases diagnosedato
Case	En person, som tilhører den underliggende kohorte, og som får diagnosticeret outcome i løbet af studieperioden
Kontrolperson	En person, som tilhører den underliggende kohorte, og som indtil den tildelte indeksdato ikke har fået diagnosticeret outcome inden for studieperioden
Matching	Pardannelse mellem cases og kontrolpersoner, baseret på udvalgte patientkarakteristika
Density sampling	Udvælgelse af kontrolpersoner, hvor sandsynligheden for at blive valgt er proportional med den akkumulerede persontid
Risk set sampling	Variant af density sampling, hvor kontrolpersonernes indeksdato er baseret på indeksdatoer blandt cases, så der opnås en lige fordeling af indeksdatoer blandt cases og kontrolpersoner

Hvis der er anvendt risk set sampling, opretholdes parringen mellem cases og kontrolpersoner (én case og de tilhørende kontrolpersoner danner et risk set), og data analyseres stratificeret på det enkelte risk set [8].

Absolutte effektestimater fra et case-kontrol-studie

Hvis den underliggende kohorte, som cases og kontrolpersoner er samlet fra, er veldefineret, kan man beregne en incidensratedifference på basis af den estimerede oddsratio fra case-kontrol-studiet. Til dette skal man bruge incidensraten for udfaldet blandt ueksponerede personer. Raten kan være beskrevet i relateret litteratur, hvilket kræver, at studiepopulationerne er sammenlignelige eller kan beregnes som antallet af ueksponerede cases divideret med kohortens totale ueksponerede followuptid, hvis disse oplysninger er tilgængelige. Ved brug af den beregnede oddsratio kan man estimere, hvad incidensraten ville være i den eksponerede del af kohorten, og

dermed også beregne incidensratedifferencen som et absolut risikomål mellem eksponerede og ueksponerede [10].

KRITISKE SPØRGSMAÅL

Grundet case-kontrol-studiets forankring i et underliggende kohortestudie er alle relevante kritikpunkter for et kohortestudie også gældende for et case-kontrol-studie, dvs. spørgsmål om selektionsbias, informationsbias og konfounding. Vi henviser til vores tidligere beskrivelse af disse i statusartiklen om kohortestudier [11]. Hovedparten af tekniske kritikpunkter i case-kontrol-studier vil være at finde blandt disse tre fejltyper.

Som i andre studiedesign er intern validitet en forudsætning for ekstern validitet og dermed generaliserbarhed til andre populationer. Man skal imidlertid være opmærksom på nogle særlige forhold angående udvælgelsen af cases og kontrolpersoner, når man vurderer et case-kontrol-studie: Cases bør være incidente, så man undgår en opblanding med prævalente cases og dermed risiko for at introducere selektionsbias. Cases bør ligeledes repræsentere cases i den underliggende population. Udvalget af kontrolpersoner kan være en særlig udfordring i case-kontrol-studier, hvis man ikke har adgang til populationsbaserede registre som det centrale personregister. Til tider anvendes hospitalskontrolpersoner, som kan give kontrolselektionsbias (se nedenfor). I de klassiske case-kontrol-studier indhentes oplysninger om eksponering via spørgeskema og interview, hvorved manglende deltagelse og hukommelsesbias kan være et problem [12]. Dette modsat mange danske case-kontrol-studier, hvor eksponeringsoplysninger kan indhentes fra registre og dermed indsamles uafhængigt af studiet og accept fra deltagere.

Kontrolselektionsbias

Et eksempel på denne bias er følgende: I en undersøgelse af sammenhængen mellem rygning og lungecancer rekrutteres cases med lungecancer og en kontrolgruppe med andre lungesygdomme. Herefter sammenlignes, hvor udbredt rygning har været blandt cases og kontrolpersoner. Da rygning også er en prognostisk faktor for andre lungesygdomme, risikerer man, at de udvalgte kontrolpersoner lige så ofte har været rygere som cases, og man ville fejlagtigt konkludere, at rygning ikke disponerer for lungecancer. Denne fejlkilde er en variant af selektionsbias [13]. Fejlen opstår ved, at rygning påvirker sandsynligheden for at blive udvalgt som kontrolperson. Læseren skal overveje denne bias, hvis kontrolgruppen er defineret på en anden måde end blot ved fravær af den casedefinerende tilstand.

En anden hyppig fejl er at matche cases og kontrolpersoner på variable, der er stærkt korreleret med eksponeringen, men ikke udfaldet. Herved vil case og de tilhørende kontrolpersoner ofte have samme eksponeringsstatus, hvilket gør, at de ikke bidrager med information til analysen og dermed svækker præcisionen af effektestimater. Begge typer fejl under kontrolselektionen beskrives til tider som »overmatching«.

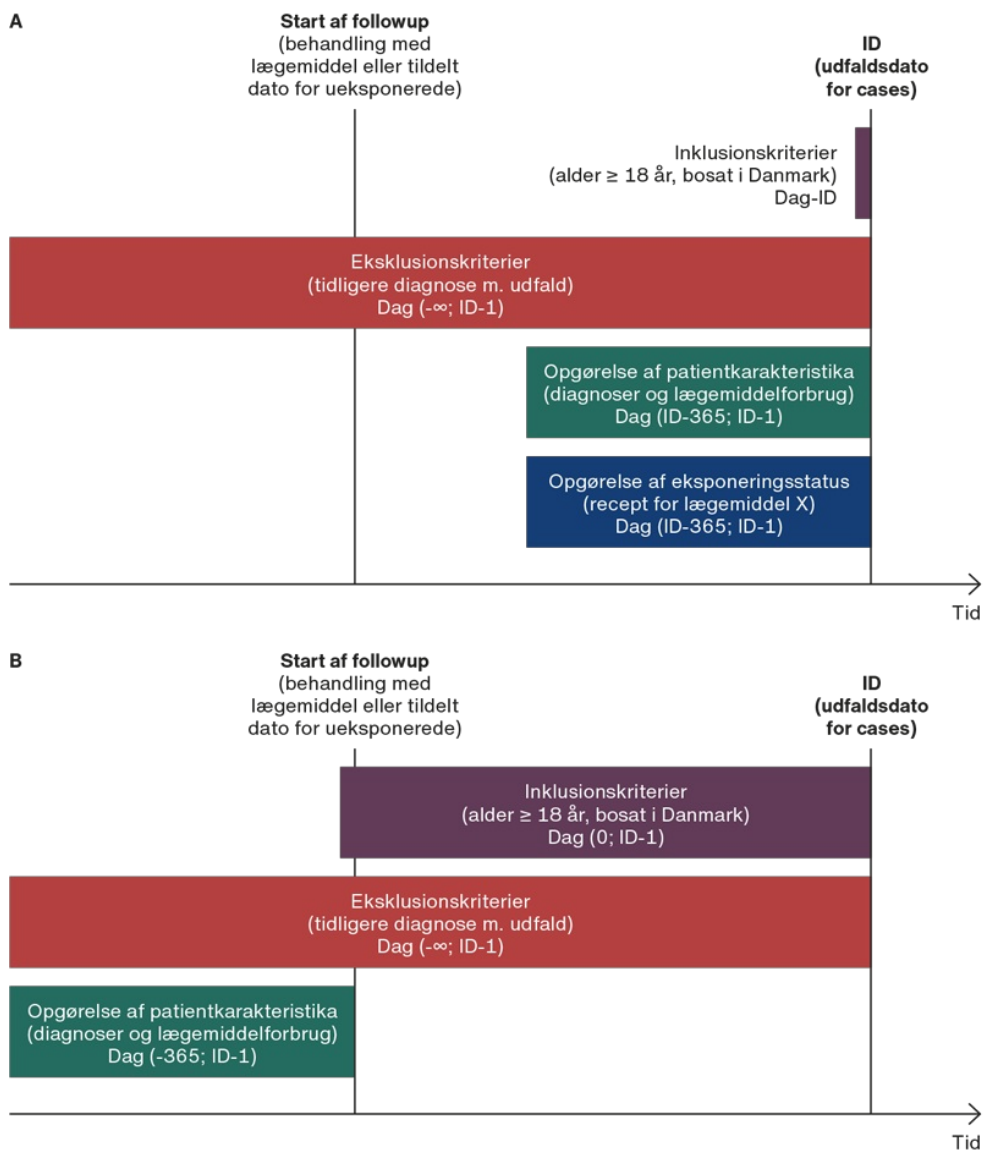
Tolkning af tabel med patientkarakteristika

I både randomiserede forsøg, kohortestudier og case-kontrol-studier præsenteres typisk en tabel med patientkarakteristika. I randomiserede forsøg og kohortestudier ønsker man, at prognostiske faktorer for udfaldet er ligeligt fordelt på inklusionstidspunktet mellem de grupper, man sammenligner, for at udelukke konfounding. I case-kontrol-studier præsenteres personerne dog typisk ikke som eksponerede og ueksponerede, men som cases og kontrolpersoner på udfaldstidspunktet (Tabel 1 i *Vinogradova et al*). Det vil derfor være helt naturligt, at prognostiske faktorer for udfaldet forekommer hyppigere blandt cases end kontrolpersoner. Med andre ord: Man bør forvente, at cases og kontrolpersoner har forskellige karakteristika [14].

Tidspunkt for opgørelse af eksponering

I et kohortestudie opgøres eksponeringen ved begyndelsen af followup. I et case-kontrol-studie derimod har man to muligheder enten ved begyndelsen af followup i det (fiktive) underliggende kohortestudie eller på indeksdatoen. Vælges førstnævnte, forebygger man tidsrelateret bias [15]. Det er ikke forkert at opgøre eksponering på indeksdatoen, men man skal være opmærksom på, hvilken slags effekt der undersøges: I et kohortestudie ville denne analysemodalitet svare til, at personer kan skifte eksponeringsstatus i løbet af followup og bidrage persontid til begge kohorter. Dette svarer til en *as-treated*-analyse af et randomiseret forsøg [16] (Figur 1).

FIGUR 1 Studiedesigndiagram for to relaterede case-kontrol-studier, der samler fra den samme underliggende kohorte. **A.** »As treated«-analyse med opgørelse af eksponering og kovariate på indeksdatoen. **B.** »Intention-to-treat«-analyse med opgørelse af eksponering og kovariate ved start af followup i det underliggende kohortestudie, start af followup for ueksponerede må anses for at være tildelt på baggrund af tidspunktet for start af followup blandt eksponerede.



ID = indeksdato.

VARIANTER AF CASE-KONTROL-STUDIER

Testnegativ case-kontrol

Det testnegative case-kontrol-design er anvendt til overvågning af effektivitet af de årlige influenzavaccinationer [17] og vaccinerne rettet mod SARS-CoV-2 [18]. I sådanne studier vil udfaldet ofte være, hvorvidt en given person er blevet testet positiv for den agens, man vaccinerer imod. Problemet herved er, at kun personer, der er blevet undersøgt for den givne agens, kan opleve udfaldet. Ved brug af et »nested« case-kontrol-studie kan man begrænse studiepopulationen til de personer, der er blevet undersøgt og dermed *reelt* har været i risiko for udfaldet. Cases er de personer, der testede positivt, mens kontrolpersoner udtrækkes blandt de personer, der testede negativt. Dette studiedesign har en høj risiko for selektionsbias [19]: Hvis man udfører et testnegativt studie over risikofaktorer for SARS-CoV-2-infektion, kan man risikere, at kontrolpersonerne fortrinsvis består af patienter med andre respiratoriske virus. Hvis de undersøgte faktorer også påvirker risikoen for at få disse virus, ville man få et falsk negativt resultat.

FORDELAGTIG ANVENDELSE AF CASE-KONTROL-STUDIER

Når et case-kontrol-studie skal evalueres kritisk, er et centralt spørgsmål, om valget af studiedesign er passende ift. formålet:

Case-kontrol-designet kan forenkle databehandlingen, hvis man ønsker at undersøge multiple eksponeringer eller anvende forskellige definitioner af den eksponering, man undersøger, såsom antal års brug, samlede mængde indtaget eller timing af brug. I et case-kontrol-studie kan alle disse forskellige eksponeringer opgøres på den tildelte indeksdato, hvorimod det tilsvarende kohortestudie skulle samples forfra for hver eksponering. En anden situation, hvor case-kontrol-studier er at foretrække, er, hvor udfaldet er sjældent, og det tilsvarende kohortestudie ville blive uforholdsmæssigt stort. Endelig er case-kontrol-tilgangen fordelagtig, hvis manuel dataindsamling er nødvendig, f.eks. for at indhente oplysninger om eksponeringer eller konfoundere, der ikke er tilgængelige i registre. Her vil stikprøven af personer reducere arbejdsbyrden mhp. den manuelle dataindsamling. Et eksempel på anvendelse af manuel dataindsamling på en stikprøve af studiepopulationen er genome-wide association studies, hvor genetiske oplysninger indsamles blandt cases og udvalgte kontrolpersoner fra baggrundspopulationen frem for at genotype alle.

RAPPORTERINGSRETNINGSLINJER

Ligesom andre observationelle studier er det gavnligt, at case-kontrol-studier rapporteres jf. »The strengthening the reporting of observational studies in epidemiology statement« [20], som findes i en udgave specifikt til case-kontrol-studier, eller »The reporting of studies conducted using observational routinely-collected health data« [21].

Korrespondance Lars Christian Lund. E-mail: lclund@health.sdu.dk

Antaget 22. november 2023

Publiceret på ugeskriftet.dk 12. februar 2024

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

Referencer findes i artiklen publiceret på ugeskriftet.dk

Artikelreference Ugeskr Læger 2024;186:V06230402.

doi 10.61409/V06230402

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 articles on case-control studies

The case-control design is one of the key designs used in observational research. In this review, we discuss common pitfalls of case-control studies and describe how case-control studies can be critically evaluated. We further assert that a well-conducted case-control study provides the same results, precision, and level of evidence as a corresponding cohort study.

REFERENCER

1. Vinogradova Y, Coupland C, Hippisley-Cox J. Use of hormone replacement therapy and risk of breast cancer: nested case-control studies using the QResearch and CPRD databases. *BMJ*. 2020;371:m3873.
2. Herrett E, Gallagher AM, Bhaskaran K et al. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-36.
3. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med*. 2016;21(4):125-7.
4. Lash TL, Rothman KJ. Case-control studies. I: Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ, red. *Modern epidemiology*, 4th ed. Wolters Kluwer, 2021.
5. Rothman KJ. Types of epidemiologic studies. *Epidemiology: an introduction*, 2nd ed. Oxford University Press, 2012.
6. Hennessy S, Bilker WB, Berlin JA, Strom BL. Factors influencing the optimal control-to-case ratio in matched case-control studies. *Am J Epidemiol*. 1999;149(2):195-7.
7. Dickerman BA, García-Albéniz X, Logan RW et al. Emulating a target trial in case-control designs: an application to statins and colorectal cancer. *Int J Epidemiol*. 2020;49(5):1637-1646.
8. Pearce N. Analysis of matched case-control studies. *BMJ*. 2016;352:i969.
9. Mansournia MA, Jewell NP, Greenland S. Case-control matching: effects, misconceptions, and recommendations. *Eur J Epidemiol*. 2018;33(1):5-14.
10. Hallas J, Christensen RD, Stürmer T, Pottegård A. Measures of 'exposure needed for one additional patient to be harmed' in population-based case-control studies. *Pharmacoepidemiol Drug Saf*. 2014;23(8):868-74.
11. Lund LC, Pottegård A, Sørensen HT, Hallas J. Kritisk læsning af artikler om kohortestudier. *Ugeskr Læger* 2024;186:V0623040.
12. Sterne J, Hernán M, McAleenan A et al. Chapter 25: risk of bias in a non-randomized study. I: Higgins J, Thomas J, Chandler J et al, red. *Cochrane Handbook for Systematic Reviews of Interventions*, version 6.3. Cochrane, 2022. <http://www.training.cochrane.org/handbook> (16. jun 2023).
13. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-25.
14. Hayes-Larson E, Kezios KL, Mooney SJ, Lovasi G. Who is in this study, anyway? Guidelines for a useful Table 1. *J Clin Epidemiol*. 2019;114:125-132.
15. Hernán MA, Sauer BC, Hernández-Díaz S et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70-75.
16. Smith VA, Coffman CJ, Hudgens MG. Interpreting the results of intention-to-treat, per-protocol, and as-treated analyses of clinical trials. *JAMA*. 2021;326(5):433-434.
17. Chua H, Feng S, Lewnard JA et al. The use of test-negative controls to monitor vaccine effectiveness: a systematic review of methodology. *Epidemiology*. 2020;31:43-64.
18. Stowe J, Andrews N, Kirsebom F et al. Effectiveness of COVID-19 vaccines against Omicron and Delta hospitalisation, a test negative case-control study. *Nat Commun*. 2022;13(1):5736.
19. Westreich D, Hudgens MG. Invited commentary: beware the test-negative design. *Am J Epidemiol*. 2016;184(5):354-6.
20. Vandenbroucke JP, von Elm E, Altman DG et al. The strengthening the reporting of observational studies in epidemiology

(STROBE): explanation and elaboration. PLoS Med. 2007;4(10):e297.

21. 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.