

Hva kan vi si om kausalitet?

Det er en vanlig oppfatning at statistikk kun er egnet for å vise assosiasjoner og korrelasjoner, men statistiske metoder er nødvendige for å etablere kausale sammenhenger i rammene av kunnskapsbasert medisin. I de senere år har det vært økende interesse for fagfeltet kausal inferens, der man utforsker metoder for å oppdage kausale forhold. Dette bør kliniske forskere benytte seg av.

Mats Julius Stensrud
m.j.stensrud@medisin.uio.no
Odd O. Aalen

> Se lederartikkel side 1424

Klinisk medisin er et handlekraftig fag, og det forventes vanligvis at legen har en formening om tiltak når hun møter pasienten. Hvis en behandling settes i gang, bør den begrunnes – legen bør ha en kausal forklaring på hvorfor intervensjonen virker. For å begrunne intervensjoner kan det legges vekt på kausale, mekanistiske forklaringer.

Kausalitet i medisinen

I medisinen har for eksempel en mekanistisk forståelse forsvart bruken av ebolamedikamentet ZMapp (1). Dette består av tre monoklonale antistoffer som nøytraliserer ebolaviruset ved å binde seg til tre av virusets glykoproteiner (2). Den mekanistiske forklaringen tilsier at ZMapp gir passiv immunitet. Til tross for at forklaringen virker plausibel, er det stor skepsis til å bruke medikamentet (3), selv etter at dyremodeller har vist effekt (3).

Skepsisen skyldes at mekanistiske forklaringer nesten aldri er tilstrekkelige i medisin og biologi. Derimot finnes det flere tilfeller der mekanistisk resonnering har hatt fatale følger. Før 1990-årene ble det anbefalt å legge spedbarn på maven. Mageleie skulle blant annet hindre kvaling ved oppkast, og barna skulle puste lettere. Statistiske data viste imidlertid at mageleie øker risikoen for plutselig, uventet spedbarnsdød betraktelig (4).

Gullstandarden for å etablere kausale forhold i medisinen er derfor randomiserte, kontrollerte forsøk, som gjerne utføres etter at man har funnet plausible mekanistiske forklaringer. Det randomiserte forsøket lener seg alene på statistikk. For ZMapp finnes det ingen resultater fra randomiserte, kontrollerte forsøk med mennesker (3, 5).

Vi kan altså snakke om to årsaksbegreper – det mekanistiske og det empiriske. Grunnen til at statistikk ble listet av *New England Journal of Medicine* som ett av de 11

viktigste bidragene til medisinen de siste tusen år (6), er den viktige posisjonen empiriske årsakssammenhenger har.

En kjent suksesshistorie om behandling som ble rettfærdiggjort av empiriske forsøk alene, er C-vitamin mot skjorbuk. I 1747 ga skipslegen James Lind (1716–94) ulike matvarer til besetningen. Noen fikk appelsiner og sitroner, andre fikk eddik, sider, saltvann eller en eliksir med vitriol (7). Bare de som fikk sitrusfrukter unngikk skjorbuk, men det fantes ingen mekanistisk forståelse som begrunnet virkningen.

Kausale slutninger fra observasjonsstudier

Et grunnleggende problem er at randomiserte, kontrollerte forsøk vanligvis er vanskelig å gjennomføre. De krever gjerne store ressurser, mye tid og etiske avveininger. I situasjoner hvor randomiserte, kontrollerte forsøk ikke finnes, er observasjonelle studier det beste alternativet. Det er blitt utviklet flere nye metoder som kan gjøre det mulig å trekke kausale slutninger selv når man bare sitter på observasjonelle data. Disse metodene blir stadig lettere å anvende for forskere uten tung matematisk bakgrunn.

Strengte forutsetninger må være oppfylt for å kunne trekke kausale slutninger. For det første må vi ha en klar formening om hva kausalitet er. I flere tilfeller er det heller ikke opplagt at kausalitet eksisterer. Blant annet er kvantefysikken basert på at individuelle hendelser ikke har noen kausal årsak (8).

I medisinen er det nyttig å definere kausalitet med utgangspunkt i *kontrafaktiske situasjoner*. En kontrafaktisk situasjon betegner et scenario som ikke har hendt i virkeligheten, men som potensielt kunne ha skjedd (9). Den kontrafaktiske definisjonen av kausalitet kan illustreres med ebola som eksempel.

La oss betrakte ebolasmittede David i to forskjellige situasjoner: Enten får han ZMapp 1. juni eller så får han ingen behandling 1. juni. I virkeligheten kan David bare plasseres i én av disse situasjonene. Det kontrafaktiske scenarioet betegner den situasjonen som David *ikke* tilhører.

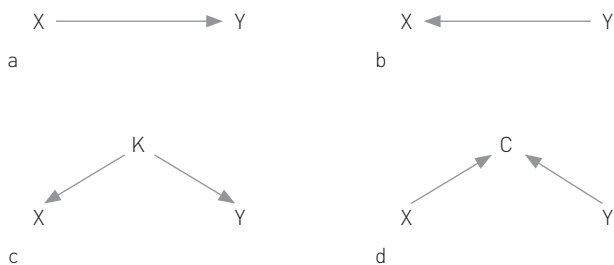
Hvis han *fikk* ZMapp 1. juni, er altså det kontrafaktiske scenarioet at han *ikke fikk* behandling. Nå kan vi definere en kausal effekt av ZMapp: Hvis sannsynligheten for å overleve ville vært annerledes dersom David ikke fikk ZMapp 1. juni, har ZMapp en kausal effekt på eboladødeligheten. Generelt sier vi at en variabel har kausal effekt hvis følgende kriterium er oppfylt: Dersom variabelen (f.eks. ebolabehandling) ble endret, mens alt annet var likt, ville utfallet (f.eks. død) blitt forandret (10).

Med en forståelse av kausalitet som er direkte knyttet til kontrafaktiske scenarioer, finnes det metoder for å manipulere observasjonelle data til å lage fiktive populasjoner, *pseudopopulasjoner*, med individer som går gjennom randomiserte forsøk (11). For at manipulasjonene skal kunne gjøres, kreves det tre forutsetninger (11–13).

For det første må behandlingen være *veldefinert*, dvs. alle i behandlingsgruppen utsettes for samme prosedyre, inngrep eller medikament. Når et ebolamedikament skal sammenliknes med placebo, er det for eksempel viktig at alle i behandlingsgruppen får samme preparat i sammenliknbare doser. I et randomisert forsøk er det lett å sørge for at kriteriet er oppfylt, men i observasjonelle studier er det ikke alltid opplagt. Dessuten er det ofte regnet som et prinsipielt poeng at et reelt kontrafaktisk alternativ til behandlingen eksisterer. For ebolamedikamentet ville det vært enkelt: Vi kunne gi placebo.

For å finne den kausale effekten av overvekt på hjerteinfarkt er det vanskelig. Å være normalvektig er det kontrafaktiske alternativet til å være overvektig, men hvordan skal vi tenke oss at individer manipuleres til å bli normalvektige? Ved å trene? Spise? Røyke sigaretter? Ved fedmeoperasjon?

For å lage pseudopopulasjonen antar vi også at alle faktorer som bestemmer utfallet av behandlingen, enten er kjent for forskerne eller er tilfeldig fordelt mellom individer i behandlingsgruppene. Det skal altså ikke finnes umålte konfunderende variabler. På engelsk kalles denne antakelsen «conditional exchangeability», og dessverre er den ikke testbar. Derfor er forskernes fagkun-



Figur 1 Hvis X og Y er assosiert, kan dette svare til fire grunnleggende typer kausale forhold. Disse forholdene kan enkelt illustreres med såkalte rettede asykliske grafer (DAG-er). a) Det kan være en kausal virkning av X på Y. b) Det kan være en kausal virkning av Y på X. c) Det kan være en konfunderende variabel K som påvirker både X og Y. d) Det kan være en «collider»-effekt, noe som betyr at en variabel C som bestemmer seleksjon til studien, er påvirket både av X og av Y

skap om faktorer som påvirker utfallet avgjørende – forskerne må anta at alle faktorer som påvirker utfallet enten er inkludert som variabler i studien eller er tilfeldig fordelt mellom behandlingsgruppene. Eventuelt kan forskeren lage en sensitivitetanalyse for å vurdere effekten av ukjente konfunderende variabler (14, 15).

Det siste kriteriet, *positivitet*, innebærer at alle risikogrupper har en sannsynlighet større enn null for å havne i alle behandlingsgruppene. Med andre ord – i hver risikogruppe skal det være folk som blir behandlet og andre folk som ikke blir behandlet.

RAMME 1

Rettet asyklisk graf

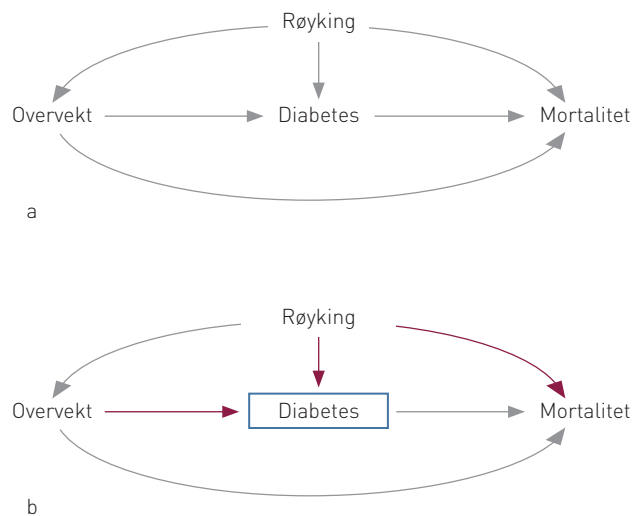
Det viktigste poenget er at hver pil i en rettet asyklisk graf illustrerer et kausalt forhold: Dersom det er en pil fra knutepunkt X til knutepunkt Y, har X en kausal effekt på Y. Formelt er det tre punkter som må være oppfylt i rettede asykliske grafer (16):

- Dersom det ikke er en pil mellom knutepunktene X og Y, er det ingen kausal effekt mellom X og Y
- Alle felles årsaker til to knutepunkter er også tegnet inn (selv om disse årsakene ikke nødvendigvis er målt av forskeren)
- Alle knutepunkter er en årsak til sine etterfølgende knutepunkter (dvs. knutepunkter som blir pekt på)

Kausale forhold kan visualiseres

For å konkretisere antakelsene og strukturen i kausale studier, er det blitt lagt vekt på grafiske representasjoner. Et viktig redskap er *rettede asykliske grafer* (directed acyclic graph, DAG) (ramme 1) (16). Slike grafer er nyttige for å forstå og visualisere forhold som kan gjøre det vanskelig å trekke direkte kausale konklusjoner (17). Noen eksempler er vist i figur 1. Grafene består av knutepunkter med stokastiske variabler (16), og knutepunktene er forbundet med piler. Hver pil illustrerer et kausalt forhold: Pilen peker fra en årsaksfaktor til et utfall. Sammen danner pilene *stier* gjennom knutepunktene. Stiene er asykliske fordi det ikke er mulig å starte ved et knutepunkt og følge piler slik at man ender i samme knutepunkt. Det betyr at en årsak ikke kan være årsaken til seg selv. For eksempel er det umulig at diabetes er årsaken til at et individ utvikler diabetes.

Kausale grafer gjør det dessuten mulig å stille kritiske spørsmål til konklusjoner fra tidligere epidemiologiske studier. Et relevant eksempel er *fedmeparadokset*. Det henviser til observasjoner om at overvekt disponerer for en rekke sykdommer (18, 19), men at hos pasienter som er rammet av slike sykdommer, er overvekt assosiert med lavere dødelighet (20). Nylig ble fedmeparadokset tilsynelatende brukt som et argument for potensielt positive virkninger av overvekt i en kommentar i Tidsskriftet (21). I rammene av kausal inferens er det klart at dette er en risikabel slutning (22, 23), og det kan illustreres med rettede asykliske grafer.



Figur 2 a) Asyklisk rettet graf (DAG) som illustrerer sammenhengen mellom overvekt, diabetes og mortalitet. b) Firkanten rundt diabetes indikerer at bare individer med diabetes er inkludert i undersøkelsen. De røde pilene viser en ikke-kausale sti mellom overvekt og mortalitet som dannes når kun mennesker med diabetes studeres. Det er viktig å vurdere denne ikke-kausale stien når assosiasjonen mellom overvekt og mortalitet skal tolkes hos denne gruppen

Her bruker vi effekten av overvekt ved diabetes som eksempel, noe som nylig ble studert av Preston & Stokes (23). Figur 2 viser en rettet asyklisk graf som inkluderer effekten av overvekt på dødeligheten. Andre velkjente forhold kommer også frem i figuren: Sigaretttrøykere tenderer til å ha lavere kroppsvekt, og røyking er forbundet med økt dødelighet. Disse to kausale forholdene fører til problemer når vi skal studere effekten av overvekt hos diabetikere. Problemet er ikke intuitivt opplagt, men med et kort matematisk argument viser Preston & Stokes at den negative assosiasjonen mellom røyking og kroppsvekt blir forsterket når vi utelukkende studerer personer med diabetes (23). Derfor blir det galt å sammenlikne dødelighet hos normalvektige diabetikere med dødelighet hos overvektige diabetikere som har ulike røykevaner.

Et annet viktig poeng er at lav kroppsvekt kan skyldes sykdom. For eksempel kan sykdom gi dårlig appetitt og økt metabolsk aktivitet. Røyking kan være årsak til slike sykdommer. Hos diabetikere kan dessuten dårlig kontroll over blodsukker nivået gi vektnedgang. Dette fenomenet kalles gjerne *invers kausalitet* (24). Når vi studerer syke, for eksempel mennesker med diabetes, kan andelen med sykdomsassosiert vektapp bli større enn i den vanlige befolkningen. Det gir økt trøbbel med *invers kausalitet*. I vårt eksempel kan personer som har lav vekt på grunn av bakenforliggende sykdom, maskere en negativ effekt av overvekt.

Disse problemene blir ofte favnet under

sekkebetegnelsen seleksjonsskjevheter, men nå har begrepet «collider bias» fått rotfeste (25). Begrepet er appellerende fordi en «collider» kan defineres enkelt og presist fra en rettet asyklisk graf: Hvis to piler peker mot samme knutepunkt, er knutepunktet en «collider» (fig 1d). «Collider» skiller seg fra konfundering, en annen type seleksjonsskjevheter som oppstår når to piler peker ut fra samme knutepunkt (fig 1c). Det finnes ingen enkel norsk oversettelse av «collider», men begrepet sikter til *kollisjonen* mellom to piler som peker mot samme knutepunkt.

I figur 2b er diabetes en «collider» fordi piler fra både røyking og overvekt peker mot sykdommen – de kolliderer i diabetes. Den røde stien i figur 2b viser den problematiske ikke-kausale stien som oppstår mellom overvekt og dødelighet når vi studerer mennesker med diabetes. Figur 2b illustrerer et viktig poeng: Forskeren trenger *ikke* ha en intuitiv forståelse for problemene en «collider» skaper – det holder å identifisere den i rettede asykliske grafer. I mange studier vil subtile «collidere» være til stede. Ofte er det vanskelig å oppdage dem ved å resonnerer, men de kan likevel identifiseres i rettede asykliske grafer. Når de er oppdaget, er det mulig å gjøre egnede analyser.

Preston & Stokes' arbeid tyder altså på at fedmeparadokset for diabetes oppstår fordi diabetes er en «collider». I en analyse av ikke-røykere viste Preston & Stokes dessuten at fedmeparadokset var fraværende (23). På samme måte er det holddepunkter for at fedmeparadokset ved kardiovaskulær sykdom i hvert fall delvis kan forklares av «collidere» (22, 26). I det hele tatt er det sannsynlig at assosiasjoner som er observert i flere epidemiologiske studier er påvirket av «collidere», for eksempel det at røyking ser ut til å beskytte mot Alzheimers sykdom (27, 28).

Konklusjon

De siste årene har det vært en enorm vekst i tilgjengelig datamateriale, og norske helse-

registre er i verdensklasse. Kliniske kvalitetsregistre er også under utvikling. Fremtidsrettede klinikere og forskere har derfor en unik sjanse til å forbedre sine kliniske studier ved å bruke nye metoder for kausal inferens.

Mats Julius Stensrud (f. 1989)

er dr.philos., har en mastergrad i statistikk fra Oxford og studerer medisin. Han er tilknyttet Avdeling for biostatistikk ved Universitetet i Oslo som forsker og underviser. Forfatter har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Odd O. Aalen (f. 1947)

er ph.d. i matematisk statistikk og professor i biostatistikk ved Avdeling for biostatistikk ved Universitetet i Oslo. Han har spesiell interesse for overlevelsesanalyse/kausale inferens. Forfatter har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Litteratur

- Hampton T. Largest-ever outbreak of Ebola virus disease thrusts experimental therapies, vaccines into spotlight. *JAMA* 2014; 312: 987–9.
- Olinger GG Jr, Pettitt J, Kim D et al. Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques. *Proc Natl Acad Sci U S A* 2012; 109: 18030–5.
- Joffe S. Evaluating novel therapies during the Ebola epidemic. *JAMA* 2014; 312: 1299–300.
- Markestad T. Sovestilling og krybbedød i Norge. *Tidsskr Nor Lægeforen* 1992; 112: 1427–9.
- McCarthy M. US signs contract with ZMapp maker to accelerate development of the Ebola drug. *BMJ* 2014; 349: g5488.
- Marcia A, Kassirer P, Relman A. Looking back on the millennium in medicine. *N Engl J Med* 2000; 342: 42–9.
- Harvie DI. Limeys: the true story of one man's war against ignorance, the establishment and the deadly scurvy. London: The History Press, 2002.
- Zeilinger A. The message of the quantum. *Nature* 2005; 438: 743.
- Höfler M. Causal inference based on counterfactuals. *BMC Med Res Methodol* 2005; 5: 28.
- Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; 15: 615–25.
- Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health* 2006; 60: 578–86.
- Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. *Int J Epidemiol* 1986; 15: 413–9.
- Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974; 66: 688–701.
- Thabane L, Mbuagbaw L, Zhang S et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol* 2013; 13: 92.
- VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology* 2010; 21: 540–51.
- Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol* 2008; 8: 70.
- Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology* 2006; 17: 360–72.
- Adams KF, Schatzkin A, Harris TB et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N Engl J Med* 2006; 355: 763–78.
- Flegal KM, Kit BK, Orpana H et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA* 2013; 309: 71–82.
- Carnethon MR, De Chavez PJ, Biggs ML et al. Association of weight status with mortality in adults with incident diabetes. *JAMA* 2012; 308: 581–90.
- Samdal GB, Meland E. Er det overvekt eller vektfokus som er helseskadelig? *Tidsskr Nor Legeforen* 2014; 134: 2247–8.
- Banack HR, Kaufman JS. Does selection bias explain the obesity paradox among individuals with cardiovascular disease? *Ann Epidemiol* 2015; 25: 342–9.
- Preston SH, Stokes A. Obesity paradox: conditioning on disease enhances biases in estimating the mortality risks of obesity. *Epidemiology* 2014; 25: 454–61.
- Flanders WD, Augestad LB. Adjusting for reverse causality in the relationship between obesity and mortality. *Int J Obes* 2008; 32 (suppl 3): S42–6.
- Cole SR, Platt RW, Schisterman EF et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol* 2010; 39: 417–20.
- Lajous M, Banack HR, Kaufman JS et al. Should patients with chronic disease be told to gain weight? The obesity paradox and selection bias. *Am J Med* 2015; 128: 334–6.
- Chang CC, Zhao Y, Lee CW et al. Smoking, death, and Alzheimer disease: a case of competing risks. *Alzheimer Dis Assoc Disord* 2012; 26: 300–6.
- van Duijn CM, Hofman A. Relation between nicotine intake and Alzheimer's disease. *BMJ* 1991; 302: 1491–4.

Mottatt 20.3. 2015, første revisjon innsendt 13.5. 2015, godkjent 27.5. 2015. Redaktør: Siri Lunde Strømme.