



Hundre år med det randomiserte forsøk

REDAKSJONELT

BJØRNDAL A

Sammenliknende erfaring er en absolutt betingelse for den eksperimentelle og vitenskapelige medisin, ellers vandrer legen på måfå og blir offer for tusen illusjoner.

Claude Bernard (1)

Det eldste kontrollerte kliniske forsøk vi kjenner i medisinen er datert 1898. Den danske legen Johannes Fibiger vurderte effekten av serumbehandling ved difteri (2). Pasienter som ble innlagt den ene dagen, fikk serumbehandling, mens de som ble innlagt neste dag ikke fikk noen behandling. Forsøket, som varte et helt år, omfattet 484 pasienter som ble innlagt ved Blegdamshospitalet med difteri. Åtte pasienter døde i behandlingsgruppen mot 30 i den like store kontrollgruppen.

Interessen for å bruke den eksperimentelle design i medisinsk forskning også utenfor laboratoriet våknet imidlertid først for alvor etter den annen verdenskrig. I 1948 presenterte The UK Medical Research Council resultatene av streptomycinbehandling ved tuberkulose bedømt ved hjelp av et randomisert forsøk (3). "Its importance cannot be exaggerated. It opened up a new world of evaluation and control which will, I think, be the key to a rational health service," skrev epidemiologen Archie Cochrane (1909-88) senere om streptomycinforsøket (4). Den britiske statistikeren Austin Bradford Hill (1897-1991) må tilskrives mye av æren for de systematiske bestrebelsene på å etablere pålitelige kontrollgrupper i vurderingen av medisinske behandlingsformer. Han var influert av Ronald A. Fischers arbeider i 1920-årene. Fischer (1890-1962), som arbeidet med landbruksforskning, introduserte eksperimentet som evalueringsform, oppfant variansanalysen og var antakelig den første som brukte uttrykket randomisering (5).

HVORFOR VURDERE EFFEKTEN AV MEDISINSKE HANDLINGER?

På godt og vondt er legeyrket handlingsorientert; leger vil gjerne "gjøre noe" for sine pasienter. Forsøkene på å svare på "Hva kan vi gjøre med det?" utgjør derfor en stor del av kunnskapsvolumet i praktisk medisin.

Når vi griper inn for å påvirke det naturlige forløp av ulike helseplager, er det viktig å være sikker på at vi gjør mer nytte enn skade. Av og til er virkningen så åpenbar og dramatisk at erfaring basert på noen få observasjoner er tilstrekkelig som begrunnelse. Hvis vi ikke får stoppet alvorlige blødninger, dør pasientene. Antibiotika ved meningitt redder liv.

Erfaring er imidlertid en dårlig veiviser når effektene er moderate eller små. Hvis noe hjelper én av ti eller én av 300, er det ikke lett å skille ut effekten av tiltaket fra den naturlige biologiske variasjon og de tilfeldigheter som påvirker alt en pasient kan feile. Ikke sjelden

står mye på spill. I 1981 publiserte Horwitz & Feinstein en pasient-kontroll-studie som tydet på at profylakse med lidokain normaliserte arytmier og reduserte dødeligheten etter hjerteinfarkt (6). Lidokain var allerede i utstrakt bruk for denne indikasjonen. Det tok ti år før ekspertene lot seg overbevise om at denne terapiformen faktisk tar flere liv enn den redder (7, 8). Effekten er i samme størrelsesorden som den positive virkningen av acetylsalisylsyre og streptokinase ved akutt hjerteinfarkt. På 1980-tallet døde årlig trolig 20000-70000 pasienter med hjerteinfarkt pga. behandling med lidokain i USA (7).

Det er først og sist det uetiske ved å utsette pasienter for fare pga. intervensjoner med ukjent effekt som dikterer behovet for pålitelige vurderinger av det vi gjør. Det gjelder enten tiltaket brukes for å forebygge eller behandle, enten "inngrepet" er teknisk eller består i en mellommenneskelig handling. Når alt kommer til alt, er alle forsøk på å bedømme hva som virker, en gjetning vedrørende sannheten. Det er et faglig og etisk imperativ å gjøre eksplisitt hvilket logisk og empirisk grunnlag gjetningen hviler på (9).

HVA ER POENGET MED RANDOMISERING?

Tilfeldig fordeling av personer til henholdsvis intervensjon eller kontroll er viktig av en eneste grunn; det er den eneste effektive måte å sikre sammenliknbarhet mht. prognose og respons på behandlingen. Ingen annen måte å etablere et sammenlikningsgrunnlag på sikrer at kjente (og derfor ofte målbare) og ukjente (ikke-målbare) forhold som virker inn på utfallet, er om lag likt fordelt mellom gruppene (10).

Randomisering er altså et forsøk på å minimalisere seleksjonsskjevhet. Denne design løser ikke problemer som kan oppstå etter etableringen av de grupper som skal sammenliknes; f.eks. systematiske ulikheter i annen behandling som tilbys (utover den som skal vurderes) eller skjevheter i hvordan utfallene blir målt.

Hvordan randomiseringen skjer, er viktig. Det var Thomas Chalmers (1917-95) og medarbeidere som først viste at det ikke å skjule prosedyren for fordeling til intervensjon eller kontroll for dem som rekrutterer pasientene (concealment of allocation), ofte resulterer i en skjevfordeling av prognostiske faktorer og derfor gir effektestimater som er forskjellige fra (og oftere statistisk signifikante) enn de estimerer man får når allokeringprosedyren er skjult (11). Denne feilkilden alene kan ha større effekt enn den vi håper behandlingen har (12).

STRIDEN OM IKKE-RANDOMISERT EFFEKTFORSKNING

Ikke alle deler synet på at randomisering er avgjørende for pålitelige effektvurderinger (13, 14). Kritikerne har flere viktige innvendinger. Systematiske oversikter, bl.a. dem som lages av The Cochrane Collaboration (15), viser at en rekke randomiserte forsøk er dårlig planlagt og gjennomført. I mange undersøkelser har man unnlatt å besvare spørsmål og måle utfall som er viktige for pasientene. Av de skammelig få forsøk som finnes i verden om effekt av schizofrenibehandling, har halvparten en oppfølgingstid på seks uker eller kortere (16). Og et randomisert forsøk kan bare klargjøre at et tiltak har (eller ikke har) effekt. Andre metoder, bl.a. kvalitative design og laboratorieforskning, må brukes for å forklare hvorfor og hvordan noe virker.

Det er også innlysende at spørsmålet må avgjøre hvordan man bør gå frem for å finne svaret. Hvis man lurer på hva innvandrerkvinner tenker om prevensjon, vil ikke et randomisert forsøk gi svaret. Hvis spørsmålet er om staten bør eie sykehusene, er randomisering en praktisk umulighet. Er man på jakt etter sjeldne bivirkninger, må man lete med andre designere.

Det er imidlertid en heftig uenighet i medisinen for tiden om hvorvidt man kan nøye seg med ikke-randomiserte effektvurderinger for å svare på spørsmål som kunne belyses med en randomisert design, men der det enten ikke er gjort eller ikke er gjort godt nok. Etter min vurdering skyldes dette at mange leger og andre handlingsorienterte faggrupper har en overdreven tro på effektene av det de driver med og ikke erkjenner klart nok at systematiske feil kan lede til innføring av uvirksomme (bortkastede ressurser) eller farlige virkemidler. Inntil nylig var det f.eks. helt uklart hva som var det beste legemiddelvalg for å

kontrollere kramper ved eklampsi fordi det ikke fantes noen randomiserte forsøk. 100 års uenighet skyldtes den håpløse oppgave å tolke ukontrollerte pasientmaterialer og ikke-randomiserte sammenlikninger. I 1995 ble dette spørsmålet besvart overbevisende (17), men ikke før bokstavelig talt millioner av kvinner hadde betalt prisen for at ingen fagfolk hadde tatt seg sammen og organisert et tilfredsstillende forsøk.

FORSKNING SKAL OPPLYSER BESLUTNINGER, IKKE DIKTERE DEM

En viktig begrunnelse for å få frem gode effektvurderinger, er å kunne bli ferdig med spørsmålet om hva den medisinske effekten er. Da er det mulig og meningsfullt å gå videre til et annet viktig spørsmål i klinisk medisin; hvordan pasienter selv vil fargelegge kunnskapsbildet med meninger, følelser og opplevelser av hvilke utfall som er viktig for dem.

Så lenge fagfolk faktisk ikke besitter et pålitelig bilde av hva som virker og hvor sterkt det virker, eller når man tror mer på den forskning som er tilgjengelig enn det er grunnlag for, fungerer "kunnskap" tilslørende. Toneangivende kvinnelige gynekologer gav i dagspressen nylig uttrykk for at flere norske kvinner burde bruke østrogen etter menopausen og hevdet at allmennpraktikerens tilbakeholdenhet trolig skyldtes at "de ikke er helt oppdatert på forskningsresultater" (18). Eksemplet illustrerer at diskusjonen om design er noe mer enn en akademisk øvelse. Kunnskap om effekter av østrogensubstitusjon fra kohort- og pasientkontroll-undersøkelser tyder på en positiv effekt på hjerte- og karsykdom (ett unngått tilfelle av hjertesykdom per 50 behandlede etter ti år) og en viss økning i forekomsten av brystkreft (ett tilfelle per 100 behandlede etter 15 år) (19, 20). Ikke bare er dette moderate effekter (som knapt kaller på sterke anbefalinger), det er også upålitelige estimater (21).

Kunnskap er makt og bør deles med andre. Upålitelig kunnskap vanskeliggjør en nødvendig demokratisering av medisinen, der publikum blir en mer aktiv deltaker i medisinske beslutninger (22). Derfor har uenigheten om hva som er pålitelig kunnskap om effekt av helsetjenester, viktige implikasjoner for relasjonen mellom lege og pasient og forholdet mellom fag og politikk.

Arild Bjørndal

LITTERATUR:

1. Bernard CL. Introduction à l'étude de la médecine expérimentale. 1866, gjenoptrykt. London: Garnie-Flammarion, 1966.
2. Wulff HR, Gøtzsche PC. Rationel klinik. Evidensbaserte diagnostiske og terapeutiske beslutninger. København: Munksgaard, 1997.
3. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *BMJ* 1948; 2: 769-82.
4. Cochrane AL. Effectiveness and efficiency: random reflections on health services. London: Nuffield Provincial Hospitals Trust, 1972.
5. Box JF, Fisher RA. The life of a scientist. New York: Wiley, 1978.
6. Horwitz RI, Feinstein AR. Improved observational method for studying therapeutic efficacy: suggestive evidence that lidocaine prophylaxis prevents death in acute myocardial infarction. *JAMA* 1981; 246: 2455-9.
7. Moore T. Deadly medicine. New York: Simon & Schuster, 1995.
8. Hine LK, Laird N, Hewitt P, Chalmers TC. Meta-analytic evidence against prophylactic use of lidocaine in acute myocardial infarction. *Arch Intern Med* 1989; 149: 2694-8.
9. Chalmers I. Unbiased, relevant, and reliable assessments in health care. Important progress during the past century, but plenty of scope for doing better. *BMJ* 1998; 317: 1167-7.
10. Bjørndal A. Det er resultatene som teller, men hvordan skal vi telle resultatene? *Tidsskr Nor Lægeforen* 1994; 114: 3416-7.

11. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983; 309: 1358-61.
12. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998; 317: 1185-90.
13. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; 312:1215-8.
14. Hansen NE, Karle H. Medicinsk teknologivurdering med kommandoetik II. *Ugeskr Læger* 1996; 159: 69-70.
15. Bjørndal A. Hvilke av de helsetjenester vi tilbyr, er effektive? *Tidsskr Nor Lægeforen* 1993; 113: 3669-70.
16. Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ* 1998; 317:1181-4.
17. Which anticonvulsant for women with eclampsia? Evidence from the Collaborative Eclampsia Trial. *Lancet* 1995; 345:1455-63.
18. Vatnelid LS, Hansen F. Du kan holde deg ung med hormoner. Det gjør kvinnelegene. *Dagbladet* 20.8.1998.
19. Petitti DB. Hormone replacement therapy and heart disease prevention: experimentation trumps observation. *JAMA* 1998; 280: 650-2.
20. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 1997; 350: 1047-59.
21. Hulley S, Grady D, Bush T, Furberg C, Herrington D, Riggs B et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. *JAMA* 1998; 280: 605-13.
22. Bjørndal A. Fra pasienttilfredshet til brukermedvirkning. *Tidsskr Nor Lægeforen* 1998; 118: 353.

Publisert: 17. oktober 2018. *Tidsskr Nor Lægeforen*. DOI:

© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no