

Leve den randomiserade studien!

Det är den enda metod som med tillräcklig kraft kan värdera nyttan av medicinska åtgärder

Arets första ledare i New England Journal of Medicine [1] ger en överblick över det förflutna årtusendets medicinska framsteg. Bland dessa tilldelas den randomiserade studien en central roll.

Liten historik

Randomisering i experimentella studier introducerades av sir Ronald Fisher (1890–1962), känd också för utvecklingen av variansanalys och multivariat analys. Metoden innebär att varje patient slumpas till något av de behandlingsalternativ som skall jämföras. På så sätt får man en god balans mellan de olika behandlingsalternativen beträffande faktorer – t ex ålder och komplicerande sjukdomar – som påverkar resultatet av behandlingen. Därmed undviks systematiska fel, bias, i slutsatser avseende vilket behandlingsalternativ som är bäst.

Den randomiserade studien ger unika möjligheter att värdera medicinska åtgärder, t ex läkemedelsbehandling. Noggrant genomförda randomiserade studier ger möjlighet att stävja snedvriden utveckling i klinisk medicin genom att onödiga rutiner eller verkningslösa åtgärder som kan ha ekonomiska vinstmotiv i bakgrunden avvecklas.

Ekonomiska vinstmotiv kan å andra sidan ibland förorsaka en alltför stor skepsis mot en behandling. Ett exempel härvidlag är invasiv utredning och be-

Sammanfattat

- Den randomiserade studien, där patienter i en klinisk prövning fördelas till olika behandlingsalternativ med hjälp av slumpen, har starkt påverkat medicinsk utveckling och dagens läkarroll.
- Genom randomiseringen undviks bias från kända och okända faktorer som påverkar behandlingsresultatet. Detta gör väl grundade slutsatser möjliga.
- Randomiserade studier om mammografiscreening för bröstcancer respektive invasiv behandling av kranskärlarna med kateteringrepp eller kirurgi efter hjärtinfarkt är exempel på studier med stor påverkan på medicinsk praxis.

handling (dvs kranskärlsröntgen, kate-terburna ingrepp och kranskärlskirurgi) efter hjärtinfarkt eller instabil kärlkramp. Denna invasiva strategi har praktiserats främst i USA med god ekonomisk förtjänst, vilket har medfört att den ofta har avvisats av skeptiker till den amerikanska kapitalismen.

Den nyligen publicerade FRISC-studien [2], en mycket väl genomförd randomiserad studie, visar att den invasiva behandlingsstrategin faktiskt är mycket fördelaktig och värd betydande ekonomiska satsningar, som aldrig skulle ha kunnat motiveras i vår ansträngda sjukvårdsekonomi utan beviskraften av en väl genomförd randomiserad studie.

Entusiasmen och pionjäranda kan göra att stora organisationer byggs upp på osäkra grunder. Granskningen via det Nordiska Cochrane-centret i Köpenhamn [3] av de randomiserade studier som ligger till grund för screening med mammografi har visat på metodologiska brister. Detta gäller särskilt de studier som talar för att mammografiscreen-

ing är av värde, medan övriga studier är av bättre kvalitet. Detta skapar stora problem vid värderingen av denna kostnadskrävande verksamhet.

Läkarrollen och randomiserade studier

Ett accelererande antal randomiserade läkemedelsprövningar ger varje år ett nytillskott av läkemedel som bevisats uppskjuta sjukdom och död i stora grupper av befolkningen. Detta ger allt digrare förteckningar över medel mot högt blodtryck, höga blodfetter, benskröhet m m. Läkarrollen förändras: mer tid går åt till att dels hantera all denna profylaktiska medicinering, dels bedöma om den har ett eventuellt samband med en ibland rik symptomflo-

ra. Därmed blir tiden att hjälpa patienter som söker för sjukdomssymtom proportionellt mindre. Detta gör läkaryrket mindre glamoröst.

Svårt uppskatta graden av publiceringsbias

Det kan finnas ekonomiska och and-

Författare

GÖRAN NILSSON

docent, hjärtavdelningen, medicinska kliniken, Centrallasarettet, Västerås

STEFAN SÖRENSEN

docent, Centrum för klinisk forskning, Uppsala universitet/Centrallasarettet, Västerås.

ra intressen att inte publicera en negativ men väl en positiv randomiserad studie, s k publiceringsbias. Detta har lett till berättigade krav på registrering av randomiserade studier [4]. Därvid skulle det protokoll som ligger till grund för en randomiserad studie registreras. Detta förpliktigar till att rapportera resultatet, och att varken justera protokollet i efterhand eller utföra godtyckliga subgruppsanalyser som från början inte var specificerade.

Tendensen att företrädesvis publicera positiva studier gör att många små publicerade studier sammanfogade i s k metaanalyser aldrig riktigt kan mäta sig i beviskraft med en stor randomiserad studie.

Det är svårt att uppskatta hur missledda vi är av publiceringsbias. Det är dock sannolikt att mängden oredovisade data med neutrala eller negativa effekter i många fall överstiger publicerade data med positiva effekter, vilket ger oss en snedvriden verklighetsbild. Den gamla tanken om en Journal of Negative Results är inte dum.

Syftet med randomisering är att skydda från bias, inte bara från kända faktorer – t ex ålder, kön och sjukdomsgrad – utan också från okända faktorer. Randomisering är nödvändig för att balansera sådana okända faktorer. Betydelsen av dessa okända faktorer kan illustreras av dödligheten i hjärtinfarkt. Endast 27 procent av variationen i dödlighet vid denna sjukdom kan förklaras av de många kända faktorer som man vet påverkar dödligheten, t ex ålder, komorbiditet och kliniska fynd vid ankomsten till sjukhus. Resten kan vi bara spekulera om [5].

Om man skall använda dödlighet i hjärtinfarkt som mått på vårdkvalitet måste man därför korrigera inte bara för de 27 procent av dödlighetsvariationen som beror på kända riskfaktorer, vilket är svårt, utan också för de 73 procent av denna variation som beror på okända riskfaktorer, vilket är omöjligt. För att lösa problemet fordras att patienten ran-

domiseras till vård på de sjukhus vars kvalitet skall jämföras, vilket är praktiskt ogenomförbart.

Randomiseringens effektivitet

Vid studium av vetenskapliga artiklar är det viktigt att kontrollera randomiseringens effektivitet genom att notera hur väl centrala baslinjekarakteristika som kön, ålder och sjukdomsgrad överensstämmer mellan de randomiserade grupperna. Om det finns stora skillnader i något baslinjekarakteristikum som har starkt samband med behandlingsresultatet, t ex ålder, manar detta till försiktighet i slutsatsen. Tack vare slumpens effekt är dock detta sällsynt i tillräckligt stora studier med korrekt utförd randomisering. Baslinjekarakteristika med föga samband med behandlingsresultaten är tämligen onödiga att registrera.

Randomisering motverkar bias från kända men inte registrerade faktorer av betydelse för behandlingens resultat. Detta innebär att tillräckligt stora randomiserade studier ofta kan nöja sig med registrering av några få baslinjekarakteristika, och avvara komplicerad och utförlig registrering av många karakteristika. I stället kan man lägga energin på att samla flera fall. De stora trombofysstudierna vid hjärtinfarkt, t ex ISIS 2 [6], bygger på en sådan strategi.

Somliga forskare får dock aldrig riktigt ro i själen utan mycket långa listor av baslinjekarakteristika väl jämförda med signifikantest, vilka ju även vid perfekt randomisering genom slumpens effekt är positiv på 5 procents nivå i ca 5 procent av fallen. Signifikantest är i detta sammanhang tämligen onödiga och kan sällan statistiskt säkerställa likhet i bakgrundsvariablerna, vilket är en annan sak än att statistiskt säkerställa olikhet i dessa variabler. Överambition på detta område kan göra studien onödigt dyr och svår att praktiskt genomföra.

Dålig balans i någon av de viktigas-

te baslinjekarakteristika, t ex ålder, kan väcka misstanke om brister i randomiseringsproceduren, vilket avslöjats vid granskning av mammografistudierna [3]. Justering i efterföljande analyser för olikheter i ett baslinjekarakteristikum, t ex ålder, skapar bristande balans i andra baslinjekarakteristika eftersom sådana karakteristika rimligen inte är perfekt korrelerade till varandra. Rättar man till obalans i ett avseende får man ny obalans i ett annat. Man kommer ut på ett gungfly som gör det omöjligt att dra säkra slutsatser av studien.

Fast regelverk av nöden

Den randomiserade studien fordrar dock ett fast regelverk:

- *Det skall finnas entydigt definierade inklusions- och exklusionskriterier.*
- *Randomiseringen skall vara korrekt och beskriven i detalj, och får inte tillåta någon medveten eller omedveten styrning av patient till någon behandlingsgrupp.* En bra metod är t ex att använda ogenomskinliga kuvert med meddelande om vilken grupp patienten har randomiserats till.
- *Värderingen av utfallet av behandlingen, t ex olika inträffade sjukdomsfall, måste ske blint, dvs utan kännedom om vilken behandlingsgrupp patienten tillhör.* I många fall är det viktigt att värdera vad som är dödsorsak: hjärtinfarkt eller cancer; blindningen är då givetvis väsentlig.
- *Patienten får inte uteslutas från studien efter randomisering.* Vid en jämförelse mellan medicinsk och kirurgisk behandling får man inte utesluta patienter som visar sig inte ta medicin eller får biverkningar av medicin, alternativt är så sjuka att de måste opereras. En användning mot vissa av de randomiserade mammografistudierna är t ex att patienter utesluts efter randomisering på grund av att de haft bröstcancer före

ANNONS

randomiseringen. Misstanke uppkommer då att man inte med lika stor säkerhet kan utesluta tidigare bröstcancer i kontrollgruppen som i interventionsgruppen, med resultat att exkluderade patienter blir fler i sistnämnda grupp – en misstanke som bekräftats vid senare granskning [3].

Randomiserade studier är sällan renodlade jämförelser mellan två olika behandlingar – t ex patienter som tar respektive inte tar en medicin – utan en jämförelse mellan två strategier, där man inkluderar konsekvenser som att patienten finner medicineringen så besvärlig att han avbryter denna eller att han blir så sjuk att han måste opereras för sin åkomma.

- *Det kan ibland vara nödvändigt att i en klinisk prövning avbryta en påbörjad behandling hos vissa patienter.* Kriterier för sådant behandlingsavbrott bör definieras i förväg.

- *Alla patienters öden efter randomisering måste redovisas.* Utfallet för alla patienter som tagits in i studien måste analyseras oavsett om avsedd behandling kommit till stånd eller inte, s k »intention to treat»-analys. Vid läsning av vetenskapliga artiklar är det relativt lätt att kontrollera att så sker. Ett specialfall är när den ursprungliga diagnosen är direkt felaktig, t ex om en patient randomiseras till ett behandlingsalternativ för hjärtinfarkt men senare visar sig ej ha haft hjärtinfarkt utan bröstsmärtor utlösta från matstruten.

Det kan då i vissa situationer vara motiverat att också göra en analys för att utesluta uppenbart felklassificerade fall. Vad som skall betraktas som felklassificerade fall måste dock avgöras av för typ och resultat av behandlingen blindade forskare.

Subgruppsanalyser

Om många subgruppsanalyser görs blir genom slumpvariation alltid enstaka analyser positiva [7]. Subgruppsana-

” Syftet med randomisering är att skydda från bias, inte bara från kända faktorer – t ex ålder, kön och sjukdomsgrad – utan också från okända faktorer. Randomisering är nödvändig för att balansera sådana okända faktorer. Betydelsen av dessa okända faktorer kan illustreras av dödligheten i hjärtinfarkt. Endast 27 procent av variationen i dödlighet vid denna sjukdom kan förklaras av de många kända faktorer som man vet påverkar dödligheten, t ex ålder, komorbiditet och kliniska fynd vid ankomsten till sjukhus. Resten kan vi bara spekulera om.

lyser bör begränsas till några kliniskt viktiga variabler som har ett starkt samband med utfallet av behandlingen, och bör anges i det ursprungliga protokollet.

Vid studium av dödlighet efter hjärtinfarkt är ålder, diabetes och tidigare hjärtinfarkt exempel på sådana variabler.

Det är viktigt att skilja på i förväg planerade (ad hoc) och icke planerade (post hoc) subgruppsanalyser. De icke planerade analyserna genomförs ofta för att påvisa positiva resultat, där den planerade analysen har utfallit negativt. De icke planerade analyserna kan utgöra underlag för hypoteser som testas i kommande studier, men de bör inte utnyttjas för behandlingsrekommendationer.

Den statistiska kraften hos en studie är vanligen beräknad för hela patientgruppen, varför subgrupperna ofta innehåller för få patienter för att medge en meningsfull analys. Analys av resultaten för samtliga patienter ger därför ofta en bättre bild av behandlingseffekten i subgruppen än analysresultat i själva subgruppen. Det klassiska exemplet är en studie av streptokinas- och aspirinbehandling vid hjärtinfarkt [6], där subgruppsanalys visade signifikant positiv effekt av dessa läkemedel i samtliga subgrupper, med undantag för att aspirin inte hade någon effekt hos personer födda i tvillingarnas och vågens stjärnbilder. Aspirinets effekt hos hjärtinfarktpatienter födda i nämnda stjärnbilder bedöms med all rimlighet bättre genom analys av hela patientmaterialet än genom subgruppsanalys!

Den randomiserade studiens begränsning

Randomiserade studier ger i många fall en fast grund för kliniskt medicinskt handlande, men man måste känna dess begränsning:

- *Medicinsk utveckling gör att studier-na snabbt blir inaktuella.* Om tio år har sannolikt medicinering och invasiv me-

todik vid akut koronarsjukdom ändrats så mycket att balansen mellan för- och nackdelar av olika alternativ förskjutits kraftigt. Mammografin kan ju t ex ha genomgått tekniska förbättringar som gör att värdet är större nu än då metoden introducerades.

• *Den strategi som visar sig bäst i den randomiserade studien är sannolikt inte fördelaktig för alla patienter med en sjukdom.* I vissa fall är de ogynnsamma utfallen uppenbara, t ex när patienten får en hjärnblödning efter trombolys vid hjärtinfarkt, eller då kirurgisk behandling slutar med allvarliga komplikationer. I andra fall är inte ogynnsamma effekter uppenbara, men troliga. Sammantaget förbättrad överlevnad med betablockad efter hjärtinfarkt kan t ex förmodas bero på att vissa får sitt liv förlängt och att andra får det förkortat, men att den första kategorin överväger.

Möjligheterna att förbättra urvalet till en behandling är viktiga och gör att den randomiserade studien ofta inte utgör det slutliga svaret på värdet av en behandling.

• *Viktiga kliniska frågor kan av praktiska skäl inte lösas med randomiserade studier.* Vi kan t ex inte randomisera vissa människor att hålla en kost och andra en annan för att studera framtida

överlevnad och sjuklighet. Vid observationsstudier kan ju en viss kost vara kopplad till många andra hälso- eller ohälsosfaktorer. Detta ger mycket stor osäkerhet om effekterna av olika sorters kost, en osäkerhet som ofta kommer bort i medicinsk rapportering.

Observationsstudier har visat att kvinnor som medicinerar med östrogen insjuknar mindre ofta i kranskärslssjukdom än andra kvinnor [8]. En randomiserad studie visar å andra sidan att östrogen inte har något värde när det gäller sekundär prevention av kranskärslssjukdom [9]. Diskrepansen beror rimligen på att kvinnor som tar östrogen skiljer sig från dem som inte tar östrogen beträffande andra faktorer – t ex kost och motion – av betydelse för kranskärslssjukdom.

En stor pågående randomiserad primärpreventiv studie av östrogen och kranskärslssjukdom [10] kommer inom några år att ge värdefull kompletterande information i denna fråga.

Samtantaget är välgjorda randomiserade studier av utomordentligt värde för evidensbaserad klinisk medicin, men de har många begränsningar. Vardagslivets kliniska frågeställningar täcks endast till en mindre del av den fasta kunskap som fås av sådana studier. Kliniska rimlighetsbedömningar blir då nödvändiga och gör att läkaryrket fortfarande i högsta grad är en konst, för vilken kliniskt omdöme är minst lika

viktigt som evidensbaserade vårdprogram.

Referenser

1. The editors. Looking back on the millenium in medicine. *N Engl J Med* 2000; 342: 42-9.
2. FRISC II investigators. Invasive compared with non-invasive treatment in unstable coronary-artery disease: FRISC II prospective randomised multicentre study. *Lancet* 1999; 354: 708-15.
3. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000; 355: 129-34.
4. Horton R, Smith R. Time to register randomised trials. *Lancet* 1999; 354: 1138-9.
5. Normand SLT, Glickman ME, Sharma RGVK, McNeil BJ. Using admission characteristics to predict short-term mortality from myocardial infarction in elderly patients. *JAMA* 1996; 275: 1322-8.
6. ISIS-2 Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction. *Lancet* 1988; ii: 349-60.
7. Yusuf S, Wittes J, Probstfield J, Tyroler H. Analysis and interpretation of treatment effects in subgroups of patients in randomised clinical trials. *JAMA* 1991; 266: 93-8.
8. Stampfer M, Colditz GA. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. *Prev Med* 1991; 20: 47.
9. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women: Heart and estrogen progestin replacement study (HERS) research group. *JAMA* 1998; 280: 605-13.
10. Women's Health Initiative Study Group. Design of the women's health initiative clin-

Inbjudan till skribenter

Läkartidningen kommer under hösten år 2000 och våren 2001 att publicera temanummer om smärta och smärtbehandling samt genusperspektiv i medicinen. Om du har material som du tror kan vara lämpligt för något av dessa temanummer skicka in en kort sammanfattning på cirka 250 ord av ditt tilltänkta bidrag till:

Josef Milerad, Läkartidningen, Box 5603, 114 86 Stockholm eller per e-post till redaktionen@lakartidningen.se

Ange på kuvertet eller e-postmeddelandet att det rör sig om bidrag till temanumren.

Josef Milerad
docent, medicinsk chefredaktör