

Registerstudie endast ett steg i en vetenskaplig process

Genererar hypoteser
snarare än prövar dem



ADAM TAUBE, professor, institutionen för informationsvetenskap, Uppsala universitet adam.taube@dis.uu.se

Under de senaste två decennierna har ett stort antal medicinska patientregister byggts upp i Sverige. Ambitionen är att kontinuerligt kunna följa respektive verksamhet och få ett underlag för kvalitetsgranskning och utvärdering. Man får då möjlighet att t ex se i vilken utsträckning de olika klinikerna följer givna behandlingsrekommendationer, studera långtidsöverlevnaden efter viss behandling och konstatera om biverkningsfrekvensen varierar mellan olika slags åtgärder, såväl medicinska som kirurgiska.

Stora förhoppningar ställs nu på dessa register såväl från politiskt/sjukvårdsadministrativt håll som från själva professionen. En fråga av intresse är då bl a hur man kan jämföra resultaten vid olika kliniker inom samma bransch.

I en katalog [1] presenteras ett 60-tal olika register, och för vart och ett av dem påpekas, med något varierande formuleringar, »att varje klinik erhåller även egna data, samt möjligheter att jämföra egna resultat med landets genomsnitt«. Ytterligare ett steg i utvecklingen antyds sedan när det skrivs att: »Registren kommer de närmaste åren att ha en central roll när hälso- och sjukvården i en helt annan omfattning än hittills börjar redovisa resultat öppet för att tillgodose befolkningens krav på insyn och valmöjlighet.«

Det finns alltså goda skäl att diskutera vad för slags information som kan uppnås när data från ett register används för att jämföra resultaten vid olika kliniker.

Vi ska här närmare granska möjligheterna att via ett operationsregister göra jämförelser mellan olika kliniker, och vi kommer att göra detta under den förenklade (och ibland kanske till och med orealistiska) förutsättningen att det för varje opererad patient går att ange ett realistiskt och invändningsfritt mått som indikator på behandlingens framgång – det må vara överlevnaden efter ett operativt ingrepp, ett livskvalitetsmått eller något annat. Problematiken kring möjliga relevanta mått i detta sammanhang är så omfattande att den inte kan tas upp här.

I den vetenskapliga metodarsenalen utgör den randomiserade kliniska prövningen (RCT; randomised clinical trial) en grundansats när det gäller att jämföra olika behandlingar. Även om registerstudier och randomiserade kliniska prövningar inte har helt likartade syften ska vi här, för att få en utgångspunkt, betrakta operationer utförda vid två skilda kliniker, A

»Men just därför är det viktigt att besinna att ett registerfynd inte ger hela sanningen ...«

och B, som två olika »behandlingar« och först söka klargöra precis hur en RCT skulle kunna genomföras så enkelt som möjligt för att undersöka huruvida det är någon systematisk skillnad i resultat mellan dem. Vi kan sedan ställa detta mot situationen när man i stället för en RCT använder registerdata. Av Fakta framgår för- och nackdelar med registerstudier respektive RCT.

Uppläggning av en randomiserad klinisk prövning

Om operationsresultaten på klinik A och klinik B ska jämföras är utgångspunkten det sammantagna flödet av presumtiva patienter från de bägge klinikernas upptagningsområden. Dessa individer ska alltså vara i behov av den aktuella operationen och villiga att genomgå densamma. De må utgöra den initiala populationen, och de är utgångspunkten i Figur 1.

Första steget (I) blir att för varje potentiell patient bedöma – centralt – om en operation bör företas eller inte, och, i enlighet med i förväg specificerade inklusionskriterier, om vederbörande är lämplig att ingå i en RCT. Vid en RCT föreligger det för det mesta en strävan att renodla den studerade patientgruppen så mycket som möjligt för att underlätta den äsyftade jämförelsen mellan resultaten av de bägge behandlingarna. Förmodligen finns bland de granskade personerna t ex somliga med s k multipla åkomor eller i så hög ålder att de inte är lämpliga att inkluderas i en prövning. Resultatet blir till slut att det skapas en grupp av »personer lämpliga för operation och inklusion i RCT«.

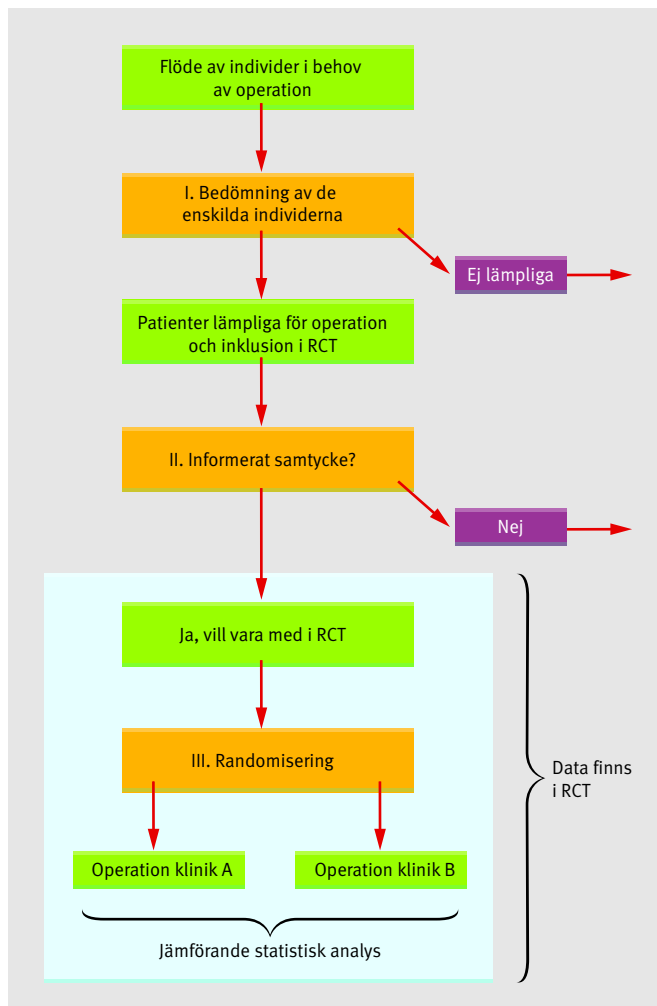
Nästa steg (II) blir att de personer som funnits lämpliga för operation tillfrågas om de verkligen vill delta i en randomiserad klinisk prövning, dvs om de vill ge s k informerat samtycke. Detta är en skyldighet rent juridiskt. De som svarar »ja« kommer att utgöra grundvalen för den randomiserade prövningen. De utgör alltså en grupp individer som har behov av operationen, dessutom är villiga att låta sig opereras och är lämpliga som deltagare i en RCT.

Sista steget (III) blir att dessa personer lottas till endera av

SAMMANFATTAT

I Sverige finns en stor mängd patientregister, och det är viktigt att veta vilken information de kan erbjuda vid jämförelse mellan kliniker. Därför ställs här en tänkt studie av ett operationsregister för jämförelse mellan två kliniker mot en randomiserad klinisk prövning (RCT) av samma frågeställning. En RCT möjliggör prövning av hypotesen att klinikerna ger likvärdiga resultat. Om en statistiskt signifikant skillnad erhålls är den direkt tolkbar. Slutsatsen gäller dock den begränsade kategori patienter som sällats ut som lämpliga för en RCT och där alla givit sitt informerade samtycke. En registerstudie avspeglar verkligheten (»hur det brukar gå«) vid respektive klinik med

de patienter man där är van att hantera. Informerat samtycke behövs inte eftersträvas. Om en statistisk skillnad mellan klinikernas resultat erhålls är orsakerna därtill inte självklara. De kan sammanhånga med olika patienturval för operation, olika operationsskicklighet eller skillnader i prognos-påverkande faktorer mellan klinikernas patienter. Registerstudier blir därför hypotesgenererande. De ger inga slutliga svar, men de kan innebära ett nog så viktigt steg i den vetenskapliga processen. Med en adekvat planerad RCT är det möjligt att få mer kunskap då en registerstudie indikerat att det kan finnas stora skillnader mellan två kliniker.



Figur 1. Flödet bakom en randomiserad klinisk prövning.

två (ej nödvändigtvis likstora) grupper, den ena för att opereras på klinik A och den andra på klinik B för att sedan jämföras statistiskt. Med hjälp av signifikansanalys testas så nollhypotesen att resultaten i de bägge grupperna är likvärdiga, dvs att det inte finns någon systematisk skillnad mellan de bägge klinikerna. Detta görs på grundval av det utfallsmått (primary endpoint) som från början har angivits i det försöksprotokoll som ligger till grund för hela prövningen.

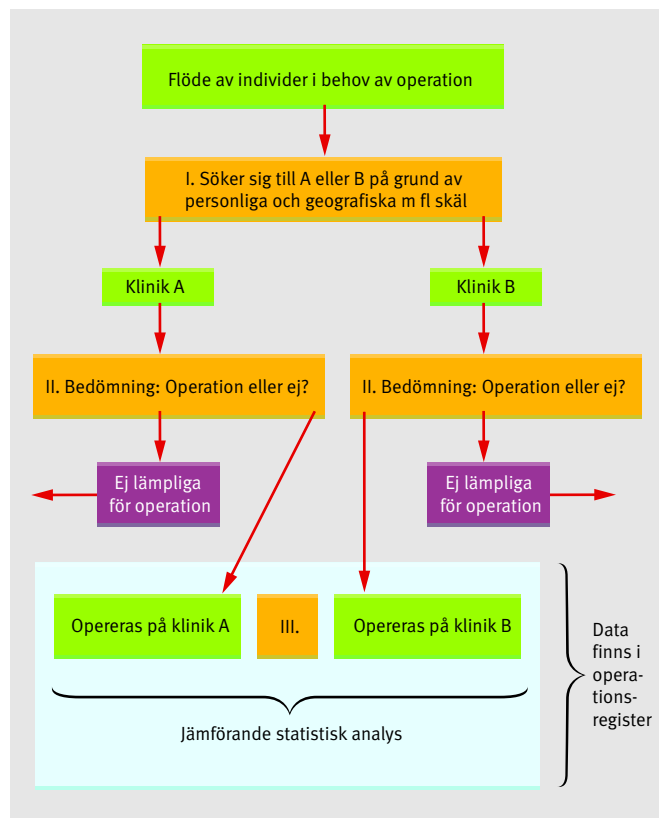
Slutsatser som kan dras vid en RCT

De två grupper som slutligen jämförs i en RCT kommer alltså från en population som i sig inte behöver avspejla det ursprungliga flödet av individer i behov av operation särskilt väl. Det beror på vilka och hur många individer som sorterats bort på grundval av de i förväg uppsatta inklusionskriterierna och hur kravet på informerat samtycke utfallit.

I gengäld är de individer som slutligen opererats på klinik A respektive B allokerade enligt slumpens lagar. Detta innebär att inga systematiska skillnader mellan dessa grupper kan förväntas snedvräta analysresultatet – grupperna är »jämförbara«.

När en signifikansanalys mellan de bägge grupperna leder till en statistiskt signifikant skillnad (dvs » $P < 0,01$ «) är den naturliga tolkningen att det verkligen föreligger en systematisk skillnad mellan klinikernas resultat.

En möjlig förklaring skulle kunna vara skillnad i operations-



Figur 2. Flödet bakom en registerstudie.

FAKTA. Registerstudie och randomiserad klinisk prövning

Jämförelse mellan två kliniker med avseende på operationsresultat

Registerstudie

Fördelar

- Avspeglar verkligheten, dvs »hur det brukar gå« vid respektive klinik
- Inget informerat samtycke och inga speciella inklusionskriterier behövs – alla som opererats kan ingå
- Grundas på information från den ordinarie sjukvården

Nackdel

- Resultatet inte entydigt tolkbart, eftersom en skillnad mellan klinikerna kan bero på
- a skillnad i operationskicksighet
 - b skillnad i urvalet av patienter lämpliga för operation
 - c skillnad i riskpopulation för respektive klinik
 - d slumpen (kan dock förkastas när P-värdet är litet)

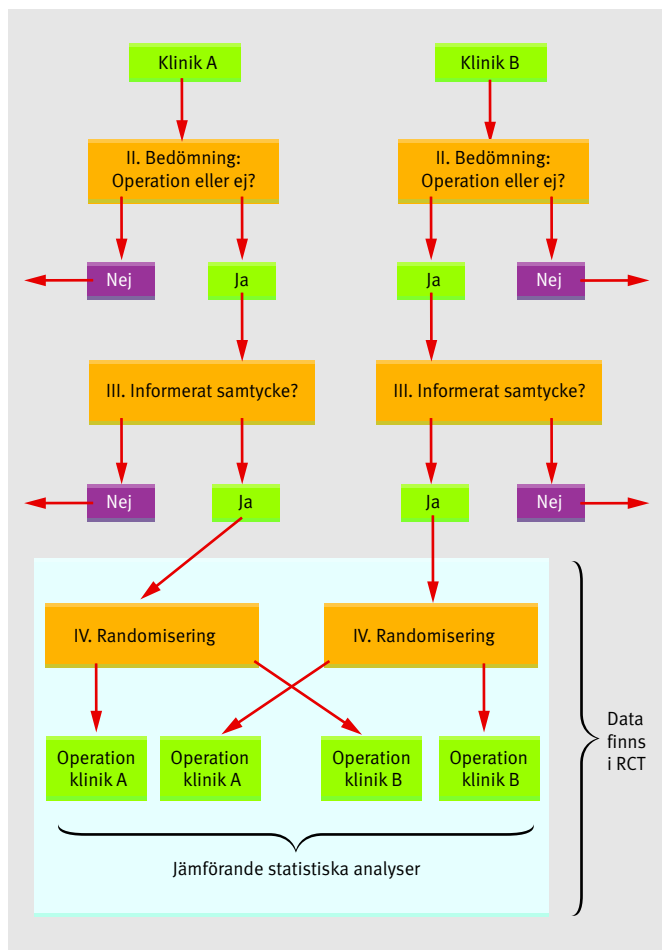
Randomiserad klinisk prövning

Fördelar

- Resultatet entydigt tolkbart, eftersom en skillnad mellan klinikerna endast kan bero på
 - a reell skillnad mellan klinikerna
 - b slumpen (kan dock förkastas om P-värdet är litet)
- Prövningen kan planeras så att flera separata hypoteser (om t ex operationsskicklighet och patienturval till operation) kan testas i samma sammanhang

Nackdelar

- Slutsatserna gäller en selekterad population på grund av (ofta hårda) inklusionskriterier och nödvändigheten att begära informerat samtycke
- En kostsam och besvärlig extra belastning utöver ordinarie sjukvård



Figur 3. Randomiserad klinisk prövning med beaktande av två faktorer – operationsresultat och urval till operation.

skicklighet, vilket i så fall bör verifieras i en ny kontrollerad studie. Den enda alternativa tolkningen – att slumpen skulle ställt till med denna skillnad – kan sorteras bort, eftersom det magiska P-värdet var så litet.

Det är på så sätt möjligt att vid en RCT genomföra en statistisk signifikansanalys och, när man kunnat förkasta nollhypotesen, få ett tolkbart resultat.

Slutsatsen av en statistiskt signifikant skillnad vid en RCT kan alltså betraktas som ett statistiskt bevis på att resultaten är systematiskt bättre vid den ena kliniken än den andra. Det bör emellertid understrykas att detta gäller de speciella patienter som kommit att till slut ingå i den randomiserade prövningen. Som redan antytts behöver dessa inte vara helt representativa för hela mängden potentiella operationsfall vid klinikerna.

Uppläggning av en registerstudie

Utgångspunkten för uppläggning av en registerstudie är givetvis densamma som vid en uppläggning av en RCT – flödet av potentiella patienter från klinikernas upptagningsområden (Figur 2). Men det första steget (I) innebär här att de enskilda individerna söker sig till, eller snarare hamnar på, klinik A eller B på grundval av geografiska skäl, remisser etc. Fördelningen av potentiella patienter på de bägge klinikerna är alltså inte slumpmässig.

Detta kan i princip (men inte nödvändigtvis) medföra att de individer som hamnar på klinik A respektive B skiljer sig åt i olika avseenden, dvs grupperna blir kanske inte helt »jämför-

bara». Men på de bägge klinikerna får man då i gengäld behandla den mix av patienter som man är van att handskas med.

Nästa steg (II) blir att respektive klinik bedömer för varje enskild patient huruvida operation bör företas eller ej. Denna bedömning görs alltså av olika bedömare vid klinik A respektive B.

Sista steget (III) blir en statistisk jämförelse mellan operationsresultaten vid klinik A och klinik B. De två patientgrupper som nu ger basen för denna analys kan skilja sig åt såväl i fråga om prognospåverkande faktorer som bedömningar av om operation varit lämplig eller ej.

Slutsatser som kan dras vid en registerstudie

De två grupper som till slut jämförs vid en registerstudie representerar var och en operationsresultaten på den mix av patienter man brukar ha vid respektive klinik. Om en statistiskt säkerställd skillnad ($P < 0,01$) erhålls vid jämförelse mellan dessa bägge kategorier kan förklaringen att den uppkommit av en slump förkastas, precis som vid en RCT. Däremot återstår en mängd andra möjliga förklaringar:

- Det föreligger verkligen en systematisk skillnad i operationsskicklighet mellan de bägge klinikerna.
- Vid den ena kliniken är man mycket skickligare på att välja ut de patienter som är lämpliga för operation.
- De patienter som hamnat vid respektive klinik A och B skiljer sig från början åt i fråga om prognos, de kan inte betraktas som »jämförbara».

Dessa förklaringar kan alla vara relevanta på en gång. Det skulle t ex kunna inträffa att man på den ena kliniken är skickligare att operera (a) men att man på den andra i stället är duktigare på att sortera ut de för operation lämpligaste patienterna (b). Detta skulle i slutändan till och med kunna leda till att ingen skillnad i slutresultaten från de bägge klinikerna uppkommer. Naturligtvis kan skillnader i fråga om prognospåverkande faktorer föreligga mellan de patienter som söker sig till de olika klinikerna (c). I den mån dessa faktorer är kända kan deras effekt studeras och elimineras med t ex regressionsanalys. Är de inte kända kan deras eventuella effekter inte åtgärdas.

Trots alla dessa tolkningssvårigheter går det inte att komma ifrån att registerstudier kan ge en viktig indikation på »hur det brukar gå» vid de bägge klinikerna med de patienter som där brukar behandlas.

Jämförelse mellan RCT och registerstudier

Det kan alltså konstateras att en randomiserad klinisk prövning ger hållbarare slutsatser än en registerstudie när operationsresultat vid två kliniker jämförs. På grund av de inklusionskriterier som forskarna alltid måste ställa upp inför en RCT gäller dessvärre dessa mer hållbara slutsatser endast speciellt selekterade patienter. Man har på denna begränsade population kunnat göra en hypotesprövning enligt konstens alla regler och fått möjlighet att eventuellt förkasta nollhypotesen och därmed på mycket goda grunder hävda att det verkligen föreligger en systematisk skillnad mellan klinikerna. Det är emellertid inte absolut säkert att detta resultat direkt kan generaliseras till alla patienter vid klinikerna, och man har inte heller någon information om skickligheten att välja ut lämpliga patienter.

En studie på grundval av ett operationsregister kan ge grova indikationer om hur »resultaten brukar bli» vid de bägge klinikerna med de patienter som där handhas och med de bedömningar som där görs. I den meningen ger registerdata en bild närmare verkligheten än vad som fås från en RCT. Något infor-

merat samtycke behöver dessutom inte påverka utvalet. Några »bevis« kan knappast uppnås på grundval av registerdata, men många iakttagelser kan göras som ger hypoteser att närmare penetrera, inte bara på grundval av andra registermaterial utan även i framtida RCT av olika slag.

En registerstudie blir på så sätt hypotesgenererande i stället för hypotesprövande.

Ytterligare steg för att gå vidare

Den RCT som inledningsvis skisserades (Figur 1) skulle nog inte genomföras i praktiken – den gavs här en så förenklad form som möjligt för att ge en principiell bakgrund till bedömningen av registerstudier. Det finns därför anledning att fråga: Om en registerstudie givit en tydlig fingervisning om att det kan föreligga en systematisk skillnad mellan resultaten från de två klinikerna – hur kan detta sedan verifieras (eller eventuellt falsifieras) med större tillförlitlighet än vad registerdata kan ge?

Antag att flödena till respektive klinik sker så som beskrivits i steg I och II i Figur 2. Vi låter detta ge underlag för den RCT som schematiskt återges i Figur 3. Nästa steg (III) blir då att till de vid respektive klinik framsorterade patienterna ställa frågan: »Vill ni delta i en klinisk prövning där det lottas om ni ska opereras vid denna klinik eller på den andra?» För dem som då svarar »ja« följer nästa steg (IV), så att de randomiseras till operation på antingen klinik A eller klinik B. Resultatet blir till slut fyra grupper, så som framgår nederst i Figur 3.

Det är välbetänkt att lägga upp det hela som ett tvåfaktorsförsök så att de slutliga fyra grupperna i Figur 3 blir lika stora. Om så sker gäller att av de patienter som opererats vid klinik A har hälften bedömts vid denna klinik och hälften vid den andra kliniken. Samma sak gäller dem som opererats vid den andra kliniken. Hypotesen att ingen systematisk skillnad i operationsresultat finns mellan de bägge klinikerna kan så testas genom en jämförelse mellan de två kategorierna patienter opererade vid klinik A respektive klinik B.

För test av hypotesen att ingen systematisk skillnad föreligger beträffande urvalet av patienter lämpliga för operation gäller motsvarande – hälften av de patienter som bedömts vid klinik A har opererats vid denna klinik och övriga vid den andra. Motsvarande gäller givetvis för den kategori patienter som bedömts vid klinik B.

Till slut kan på likartat sätt göras ett test av huruvida man vid klinikerna får likvärdiga resultat på de patienter man själv valt ut till operation jämfört med resultaten hos dem som sållats fram av den andra kliniken.

Det är säkerligen inte att vänta – eller ens meningsfullt – att varje registerbaserat »fynd« i fråga om skillnader mellan kliniker följs upp på detta mera sofistikerade sätt. Men just därför är det viktigt att besinna att ett registerfynd inte ger hela sanningen och att det, för intressanta resultat från registerstudier, i princip är möjligt att gå vidare för att nå större visshet. Registerstudier ger inte det slutliga svaret, men de kan innebära ett nog så viktigt steg i den vetenskapliga processen.

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

Kommentera denna artikel på www.lakartidningen.se

REFERENSER

1. Nationella kvalitetsregister inom hälso- och sjukvården 2005. Stock-

holm: Sveriges Kommuner och Landsting; 2005. ISBN 91-7164-074-6.

Dela med dig av dina erfarenheter.

Kommentera artiklarna
i Läkartidningen på
www.lakartidningen.se

Utmanande
saklig

Läkartidningen