

Sten Thelander, Statens beredning för medicinsk utvärdering, Stockholm (thelander@sbu.se)

Observationsstudien är central vid flera medicinska frågeställningar

Men det finns en risk för systematiska fel

II Den mångåriga, men ofruktbara, diskussionen om vilken studiedesign som är »bäst« blommade upp i somras, då *New England Journal of Medicine* i samma nummer publicerade två artiklar som inte fann några väsentliga skillnader då behandlingseffekt studerats i randomiserade kontrollerade studier eller i prospektiva kohortstudier [1, 2]. Studierna leddes av en kritisk ledare, skriven av två kända brittiska medicinska metodologer och statistiker [3]. En ledare publicerades också något senare i *British Medical Journal*, i vilken den välgjorda randomiserade studien med rätta framhölls som den mest tillförlitliga [4].

Stora skillnader länder emellan

Sannolikt måste man se diskussionerna i ljuset av de stora skillnader som föreligger mellan USA och övriga industriländers respektive vårdorganisationer. Det amerikanska fragmenterade systemet med starka vinstintressen i alla led, och med en ofta mycket bristfällig vårdkontinuitet, har gjort det allt svårare att genomföra stora randomiserade studier. Federala satsningar på randomiserade undersökningar inom ramen för verksamheten inom National Institutes of Health är omfattande, men otillräckliga, och inte sällan förenade med generaliserbarhetsproblem. Detta gäller speciellt studier av långtidsbehandlingar för kroniska sjukdomar. Läkemedelsstudier är oftast industrifinansierade och har huvudsyftet att ge dokumentation som underlag för kontrollmyndigheternas bedömning av nya preparat.

I ett försök att kompensera för dessa svårigheter har man utvecklat det som brukar kallas »outcome research« [5]. Denna syftar till att utifrån patientadministrativa databaser dra slutsatser om behandlingars effektivitet. Speciellt har man försökt använda de data som samlats in som underlag för att debitera försäkringsbolagen. Dessa data är ju inte alls insamlade i syfte att dokumentera behandlingseffekter. Det är inte slumpen som avgör vilken behandling en viss patient får, utan en mängd mer eller mindre okända faktorer som man sedan försöker korrigera för i avancerade statistiska modeller. Såväl inom som utom USA råder en stark skepsis mot denna typ av studier och de slutsatser som kan dras av dessa [6-8].

Av betydligt större värde är de patientdatabaser som byggts upp i åtskilliga länder, inklusive USA [9]. Dessa initi-

SAMMANFATTAT

Icke-experimentella studier, ofta sammanfattade som observationsstudier, är centrala för många medicinska frågeställningar.

Kvaliteten på observationsstudier har blivit allt bättre.

Hur välgjorda observationsstudier än är kan man inte bortse ifrån risken för systematiska fel.

Observationsstudiernas viktigaste frågeställningar rör riskvärderingar.

Resultaten från observationsstudier bör bekräftas av randomiserade studier.

Randomiserade studier bör ofta kompletteras med observationsstudier för att öka generaliserbarheten av resultaten.

Serie: Evidensbaserad medicin

eras ofta lokalt/regionalt eller av specialistföreningar. I Sverige finner man motsvarande register inom Socialstyrelsens MARS-databas (Medical Access and Result System). Eftersom data oftast rapporteras på klinisknivå är det dock svårt att använda informationen mer än som en jämförelse mellan olika vårdenheter. Det är däremot inte möjligt att dra slutsatser om olika behandlingars inbördes effekt och säkerhet.

I Storbritannien finns flera högklassiga register, men när man i en färsk studie försökte använda rutindata från sjukvården i Skottland var informationen genomgående otillräcklig för bestämda slutsatser [10, 11]. Detta innebär att förutsättningen för att patientdataregister skall kunna användas som komplement till randomiserade studier är att de utformas med

II Fakta

Studiedesign

Kontrollerad studie. En viss behandling ges till en grupp patienter, en annan behandling eller ingen behandling alls ges till en annan patientgrupp som man på olika sätt försöker få så likartad den första gruppen som möjligt. Detta kan ske genom matchning för olika egenskaper eller genom fördelning baserad på besöksdag, födelsedag, vårdavdelning eller liknande. Alla dessa sätt att fördela patienter kan potentiellt påverkas av behandlaren, så att grupperna i något viktigt avseende skiljer sig åt.

Randomisering, dvs slumpfördelning till de jämförda grupperna, är det enda sättet att undvika systematiska fel, »bias«, i fördelningen. Randomiseringen gör dock inte med någon automatik grupperna lika, men ju större antal patienter som ingår, desto mindre är risken att slumpen gör grupperna signifikant olika i väsentliga avseenden.

Observationsstudier kallas ibland också epidemiologiska studier, men det mest generella draget hos denna studiekategori är att den inte är experimentell. Huvudkategorierna bland observationsstudier är kohort- respektive fall-kontrollstudier. Båda dessa kan vara antingen prospektiva eller retrospektiva.

Kohortstudier identifierar en population som utsatts för den exponering man vill studera och jämför med en annan kohort som inte exponerats. Sedan följs de båda kohorterna över tid, och utfallet i de olika antalet sjukdomsfall eller dödsfall jämförs.

Fall-kontrollstudien identifierar en grupp individer som utvecklats en sjukdom och som utsatts för en viss exponering. Sedan försöker man matcha dessa med en annan grupp som inte exponerats men i övrigt är så lik den identifierade gruppen som möjligt. Förekomsten av sjukdom hos dessa studeras därefter. Om färre individer i kontrollgruppen utvecklar sjukdom kan detta tala för att en identifierad faktor i fallgruppen medför ökad risk för en viss sjukdom.

stor detaljrikedom, och att avidentifierade data på den enskilda patientnivån blir tillgängliga.

Skepsis mot observationsstudiers värde

Inom den lösligt organiserade rörelse som brukar kallas evidensbaserad medicin (EBM) finns en utbredd skepsis mot observationsstudiernas värde när det gäller att dra slutsatser om behandlingseffekt [12-14]. Problemet med denna kritik är att den ofta drar alla typer av observationsstudier över en kam. I de klassiska studierna från 1970-talet jämfördes de randomiserade studierna med fallserier eller historiska kontroller [15-17]. Då var det föga överraskande att sådana studietyper visade större behandlingseffekter än dem man såg i de randomiserade studierna. Det är ett genomgående resultat från olika forskargrupper att när observationsstudien är utformad som en prospektiv kohortstudie, företrädesvis baserad på de patienter som uppfyllde randomiseringskriterier men som av olika anledningar inte randomiserades, finner man inga systematiska skillnader i effektens riktning eller storlek [1, 2, 18-20].

Begränsningar i den randomiserade studiedesignen

Den experimentella modellen innebär att man försöker hålla allt utom det man önskar studera så standardiserat och enhet-

ligt som möjligt. Detta gör att möjligheterna att upptäcka eventuella skillnader ökar; dock på bekostnad av möjligheterna att generalisera till andra grupper än de som studerats. Man brukar säga att den interna validiteten är hög, men att den externa validiteten kan vara begränsad. Försök har gjorts att skapa modeller för att integrera information från randomiserade studier och observationsstudier i så kallad »cross-design synthesis«. Dessa teoretiskt tilltalande försök har inte gett några praktiskt användbara tillämpningar.

När väl en effekt påvisats i en randomiserad studie, eller i en metaanalys av flera studier, kan även effekten i mer selekterade populationer undersökas med samma metod. Dessa stora, enkla, randomiserade studier har med stor framgång tillämpats inom bl a onkologi och kardiologi. Flera tiotusentals patienter har randomiserats till två eller tre behandlingsgrupper. I dessa grupper har patienter med mycket varierande egenskaper inkluderats. Förutom den studerade behandlingen har övriga behandlingar givits enligt respektive kliniks egna rutiner. Med så stora grupper, och förutsatt att prognostiskt intressanta subgrupper definierats före studien, har undersökningar av skillnader mellan dessa subgruppers behandlingsresultat kunnat göras.

Finansieringsproblem, logistiska svårigheter och trögheter i forskningsorganisation har dock gjort att de flesta behandlingar i sjukvården inte undersökts i sådana gigantiska studier.

För den behandlande klinikern är därför en återkommande fråga i vilken utsträckning som kliniska behandlingsresultat är tillämpliga för en viss patient. Ålder, kön och annan samtidig sjukdom, liksom behandlingen av denna, är faktorer som kan vara annorlunda än de som kännetecknade deltagarna i de randomiserade undersökningarna. Även om dessa faktorer är desamma kan den aktuella sjukdomen vara lindrigare eller allvarligare än hos de studerade patienterna. En behandling kanske är mindre verksam vid långt gånget sjukdom. Eller, tvärtom, kan sjukdomen vara alltför beskedlig för att behandlingen skall kunna ge någon specifik effekt [21]. Allvarliga men ovanliga komplikationer av behandlingen kanske inte längre är acceptabla om den form av sjukdomen som den aktuella patienten har är förhållandevis godartad.

Välgjorda observationsstudier kan ge värdefull kompletterande information om behandlingseffekter i mindre selekterade patientpopulationer än de som studeras i randomiserade prövningar. Väsentlig kritik har framförts mot vissa uttorkningar av paradigmat evidensbaserad medicin [22, 23]. Denna kritik gäller dock bara de, lyckligtvis, få som mer eller mindre öppet förnekar behovet av andra studietyper än den randomiserade, kontrollerade prövningen.

Observationsstudiernas viktigaste användningsområden

De huvudsakliga frågeställningarna vid observationsstudier är förekomsten av riskfaktorer antingen i miljö eller som följd av en oavsiktlig eller avsiktlig exponering för en faktor som kan utgöra en risk [24]. Upptäckten av allvarliga men ovanliga biverkningar av läkemedel eller toxiner kan i princip ske enbart med hjälp av observationsdata. Dessa kan vara i form av obligatorisk eller spontan rapportering av oönskade reaktioner på behandling, där sedan farmakoepidemiologer söker efter ett mönster. Av särskild vikt är symtom som inte hör till läkemedlets farmakologiska effekter och som inte brukar ses vid den sjukdom som behandlas.

Identifieringen av riskfaktorer i omgivningen är svårare. Små riskökningar kan lika väl bero på okända bakomliggande faktorer, »confounders«, som på en faktisk effekt. Reproducerade riskökningar i andra populationer kan öka trovärdigheten, men om samma studiemetod tillämpas kan samma »confounder« vara verksam i alla. Ett dos-effekts samband ger

ett starkt stöd för att ett orsakssamband föreligger. Likaså är stora riskökningar ofta svårare att förklara med »confounders«.

Det finns ingen enighet om vad som skall menas med små respektive stora riskökningar. Vissa brukar ange en oddsratio på över 2 eller 3 för att man skall kunna fästa någon större vikt vid en observationsstudie [14]. Andra avvisar denna typ av gränssättning och vill i stället väga ihop resultatet med patofysiologisk rimlighet och liknande [22, 23]. Det finns inte heller någon enighet om värdet av att göra metaanalyser av observationsstudier [25-27]. Antalet publicerade metaanalyser av observationsstudier ökar dock mycket snabbt. En stor försiktighet är dock att rekommendera vad gäller tolkningen av sådana metaanalyser, kanske speciellt baserade på fall-kontrollstudier [27]. I flera jämförande studier har man funnit betydligt större effektstorlekar i fall-kontrollstudier än i kohortstudier med samma frågeställning.

Utmaning

Det är i allmänhet betydligt svårare att bedöma observationsstudier än randomiserade behandlingsstudier. Inte sällan används sofistikerade statistiska metoder för att försöka korrigera för »confounders«. Sådana metoder ingår sällan i den kliniska forskarutbildningen, och aldrig i den medicinska grundutbildningen. Om man kan vara tveksam till hurtiga påståenden om att evidensbaserad medicin kan läras in på någon dags kurs vore motsvarande påstående om klassisk epidemiologi enbart löjeväckande. Det finns däremot ett stort behov av ökade kunskaper och färdigheter i utformningen speciellt av kohortstudier inom den kliniska forskarvärlden. Det förefaller helt klart att många viktiga kliniska frågeställningar inte kan besvaras enbart med hjälp av den randomiserade prövningen [28-30].

Det är en stor utmaning att tillämpa eller utveckla metoder för att fördjupa och bredda den information om behandlingseffekter som ursprungligen påvisats med hjälp av den randomiserade studieuppläggnings.

Referenser

1. Benson K, Hartz A. A comparison of observational studies and randomized controlled trials. *N Engl J Med* 2000; 342: 1878-86.
2. Concato J, Shav N, Horwitz R. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000; 342: 1887-92.
3. Pocock S, Elbourne D. Randomized trials or observational tribulations? *N Engl J Med* 2000; 342: 1907-9.
4. Barton S. Which clinical studies provide the best evidence? The best RCT still trumps the best observational study. *BMJ* 2000; 321: 255-6.
6. Byar D. Problems with using observational databases to compare treatments. *Stat Med* 1991; 10: 663-6.
8. Sheldon T. Please bypass the PORT. *BMJ* 1994; 309: 142-3.
9. Black N. High-quality clinical databases: breaking down barriers. *Lancet* 1999; 353: 1205-6.
10. Lewsey J, Murray G, Leyland A, Boddy F. Comparing outcomes of percutaneous transluminal coronary angioplasty with coronary artery bypass grafting – can routine health service data complement and enhance randomized controlled trials? *Eur Heart J* 1999; 20: 1731-5.
12. Kunz R, Oxman A. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998; 317: 1185-90.
14. Sackett D, Struss S, Richardson S, Rosenberg W, Haynes B. Evidence Based Medicine - How to practice and teach EBM. 2nd ed. London: Churchill Livingstone, 2000.
19. McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999; 319: 312-5.
20. MacLehose R, Reeves B, Harvey I, Sheldon T, Russell I, Black A. A systematic review of comparisons of effect sizes derived from ran-

domised and non-randomised studies. *Health Technol Assess* 2000; 4: 1-168.

21. Ioannides J, Lau J. The impact of high-risk patients on the results of clinical trials. *J Clin Epidemiol* 1997; 50: 1089-98.
22. Feinstein A, Horwitz R. Problems in the »evidence« of »evidence-based medicine«. *Am J Med* 1997; 103: 529-35.
23. Vandenbroucke J. Observational research and evidence-based medicine: what should we teach young physicians? *J Clin Epidemiol* 1998; 51: 467-72.
26. Shapiro S. Is meta-analysis a valid approach to the evaluation of small effects in observational studies. *J Clin Epidemiol* 1997; 50: 223-9.
27. Egger M, Schneider M, Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; 1998: 140-4.
29. Naylor D. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet* 1995; 345: 840-2.
30. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; 312: 1215-8.

Fullständig referenslista kan erhållas från författaren.