

Att utvärdera hälso- och sjukvårdsinsatser

Epidemiologiska och sociologiska synpunkter på val av utvärderingsmetoder

Denny Vågerö

Om vi vill angripa det breda spektrum av frågeställningar som kan belysa resultatet av hälso- och sjukvårdsinsatserna måste vi använda ett brett spektrum av undersökningsmetoder. De skillnader som då finns mellan experimentella och icke-experimentella metoder är relativa, framhåller i nedanstående artikel sociologen Denny Vågerö, socialmedicinska enheten, Hud- dinge sjukhus.

Metodiken är emellertid särskilt svår när det gäller att värdera resultatet av mer allomfattande hälso- och sjukvårdsinsatser. Samtidigt är detta ett centralt problem då mer övergripande teorier "vägleder" en hel del av den forskning som bedrivs med mer precis metodik.

likna det naturvetenskapliga experimentet. I det senare håller man alla variabler som ingår i försöket och som mäts under kontroll. Experimentet består sedan i att man inför en ny faktor och studerar någon beroende variabel som förändras till följd av denna faktors införande. I ett randomiserat försök (tex av en ny behandlingsmetod inom en grupp människor) kan man givetvis inte ha alla relevanta faktorer som kan påverka behandlingsresultatet under kontroll. Genom randomisering kan man dock beräkna hur stor variation i behandlingsresultatet som kan uppnås pga "slumpen" och hur sannolikt det är att det behandlingsresultatet jag erhållit är en effekt av mitt försök.

En utveckling av det randomiserade försöket är det randomiserade blindtestet, tex där man prövar ett läkemedel och en grupp får ett kemiskt verksamt preparat, medan den andra gruppen får ett placebo-preparat (placebo = jag behagar), dvs ett preparat utan känd kemisk effekt. Skillnaden mellan grupperna brukar man kalla "real-effekten" och den tolkas som beroende på läkemedlet.

I det randomiserade försöket kan ju deltagarna vara medvetna om att ett försök pågår och detta antas påverka behandlingseffekten. I det randomiserade blindtestet är deltagarna ej medvetna om att de utsätts för experiment, medan det faktum att doktorn både distribuerar medicinen respektive placebo-preparatet och ingår i utvärderingsteamet kan påverka värderingen av behandlingseffekten.

För att lösa detta problem har man konstruerat det s k dubbelblindtestet, där varken patienter eller doktor känner till vilken grupp som får placebo respektive kemiskt verksamt medicin. Här fastställer man realeffekten av en behandling, menar man.

Vad skall utvärderas? En rad av olika tillämpningsområden kan tänkas från generella utvärderingar av typen medicinens roll i dödlighetens nedgång från 1850 till 1970 (typ McKeown, (1)) till mera specifika problem av typen vilken rehabiliteringseffekt på hjärtinfarktpatienter har en ny behandlingsmetod (traditionell behandling + samtalsgrupp bestående av hjärtinfarktpatienter och specialutbildad kurator) jämfört med den tidigare traditionella behandlingsmetoden.

Finns det någon gemensam metodik, med vilken man kan angripa två så olikartade frågeställningar?

Nedan skall göras en genomgång och klassificering av några utvärderingsmetoder, med försök till angivelse och tillämpningar för var och en av dessa klasser av metoder.

Experimentella metoder och randomiserade försök

Det randomiserade försöket är ett försök att efter-

Dock, även här finns problem såsom påpekas i tex Lindahl (2), Gustafsson (3). En omfångsrik litteratur diskuterar vad placeboeffekten egentligen är. Den är ju också en reell effekt. Den har tolkats som effekten av socialt stöd eller doktors psykologiska stöd till patienten, alltså effekten av ett slags psykologisk mobilisering hos patienten (4). Om man nu genom dubbelblindtestet opererar bort denna effekt och särskiljer den andra sk realeffekten, uppstår ju ett annat problem, nämligen *generaliserbarheten av resultaten* av studiet av den medicin eller behandlingsmetod som har utvärderats i dubbelblindtest. Flera problem finns i själva verket här.

För det första: Är realeffekten densamma i laboratoriesituationen som i verkliga livet? Miettinen diskuterar detta i en studie (5) och föreslår som komplement till randomiserade försök, retrospektiva studier av medicinens effekt när dessa redan introducerats på marknaden. Dessa försök kan då givetvis inte vara randomiserade, utan hör till nästa stora grupp av utvärderingsmetoder som kan kallas för icke-experimentella utvärderingsmetoder.

För det andra: Utanför laboratoriet har vi troligen alltid att göra med en placeboeffekt. Det är dessutom möjligt att realeffekt och placeboeffekt interagerar, så att den totala effekten blir större än summan av de två. I synnerhet om vi antar att det finns en interaktion mellan kemisk "realeffekt" och psykologiskt grundad "placeboeffekt", så är det tveksamt hur generaliserbart resultatet från det sofistikerade dubbelblindtestet är.

Dubbelblindtestet förekommer mig veterligt enbart i läkemedelstester. Enkla randomiserade försök har förekommit i alla möjliga olika sammanhang, tex vid utvärdering av behandlingsmetoder av olika slag, kirurgiska ingrepp, psykiatriska samtal etc. Det finns även exempel på att man använder randomiserade försök i organisationsstudier. Om det gäller den typ av problem som McKeown brottades med eller andra mer övergripande utvärde-

Det är möjligt att realeffekt och placeboeffekt interagerar så att den totala effekten blir större än summan av de två

I icke-experimentella metoder är valet av kontrollgruppen en strategisk fråga

ringar, så är det dock knappast möjligt att lösa dessa med randomiserade eller experimentella metoder.

Jämförande icke-experimentella metoder

Det är mycket vanligt att man i medicinsk litteratur hittar jämförelser mellan en försöksgrupp och en kontrollgrupp, där man har utsatt försöksgruppen för någon form av ingrepp eller påverkan, medan kontrollgruppen ej har påverkats. Skillnader mellan försöksgruppen och kontrollgruppen vid en viss tidpunkt tolkas som effekten av denna påverkan. Den svaga punkten i de här studierna är ofta frågan om jämförbarheten mellan försöksgruppen och kontrollgruppen. Om de skiljer sig från varandra från början med avseende på någon variabel som är relaterad till den utfallsvariabel som studeras, så blir hela studien problematisk, ja möjligen ifrågasatt. I komparativa, icke-randomiserade eller icke-experimentella metoder är alltså valet av kontrollgrupp och kontrollgruppens jämförbarhet med försöksgruppen en strategisk fråga.

Många studier redovisas i tabellform där två grupper jämförs med avseende på sjukdom och icke-sjukdom. Om frekvensen sjuklighet är högre i den ena gruppen än i den andra, har man en korrelation och den tolkas då ofta som om ett orsakssamband förelåg. Detta är tveklaktigt eftersom sambandet kan vara skenbart, "spurious" och stört av ej kontrollerade variabler, confounders.

Ett alternativ till att vid studier av enkla samband diskutera confounding och värdera den ena efter den andra variabeln för att undersöka om de har stört sambandet är en simultan analys av ett större antal variabler. En samtidig kontroll av flera potentiella confounders kan ske genom att man testar data mot en multivariat funktion eller en multivariat modell. Enligt Rothman (6) så är nackdelarna med multivariat analys att man behöver konstruera en funktionell modell som beskriver förhållandet mellan behandling, confounders och utfallsvariabel. Dessutom kräver användandet av en multivariat

.....
Kan vi säga att en viss effekt gäller generellt bara därför att den gäller mellan två grupper som vi har undersökt
.....

modell att vissa statistiska antaganden är uppfyllda, vilket enligt Rothman sällan är fallet. Han föreslår därför ett alternativ till en konventionell multivariat analys, nämligen en kombination av stratifiering och multivariat analys.

Även denna grupp av metoder har ett generaliserbarhetsproblem. Kan vi säga att en viss effekt gäller generellt bara därför att den gäller inom en viss grupp eller mellan två grupper som vi har undersökt? Vid varje generalisering från en studie av en viss klinik och med en viss grupp patienter av en viss behandlingsmetods effekt uppstår ett problem, tex om vi vill säga att en viss behandlingseffekt är densamma vid alla kliniker i Sverige. Vi måste tex ha en viss kunskap om sjukhussystemet som helhet och eventuella skillnader i patienttillströmning för en sådan slutsats.

Självklarhetsbevis och beprövad erfarenhet

Det har ofta hävdats, tex av Lindahl (2), att man visst inte behöver bevisa allting med randomiserade kontrollförsök. Vissa saker är självklara. Om vi går till tandläkaren och tandvärken upphör efter behandlingen, så anser vi att det är en behandlingseffekt utan att vi har gjort någon kontroll utav confounders osv. Vilken status har denna typ av självklarhetsbevis? Hesslow (7) visar på att det självklara ibland har varit det felaktiga. Dock, vill vi väl gärna hålla med om att vissa saker tycks självklara eller uppenbara utifrån direkt erfarenhet.

Självklarhetsbevis eller den direkta erfarenheten har givetvis spelat en viktig roll i framväxten av den kliniska erfarenheten, "den beprövade erfarenheten". I kritiken av McKeown har man tex ofta utgått från klinisk erfarenhet (8, 9). Den senare kan då vara att många olika slag från osystematisk till mer eller mindre systematiserad och formaliserad som mångårig praxis. I det senaste fallet ges den ett tungt värde och tex i det mödosamt uppbyggda amerikanska utvärderingssystemet PSRO (Profes-

sional Standard Review Organization) utgör denna praxis den standard gentemot vilken nya metoder utvärderas.

Det finns emellertid också en kritik som är omvänd den Lindahlska och argumenterar att alltför få behandlingsmetoder och behandlingsparadigmer är prövade genom randomiserade försök och följaktligen att den beprövade erfarenheten ej vetenskapligt visats vara hållbar (10).

Historiska analyser och utvärdering av ingrepp på aggregerad nivå

Om vi vill studera effekten av den nya lagstiftning som 1948 introducerade National Health Service i England och gjorde om dess sjukvårdssystem i grunden, så har vi ett helt annat utvärderingsproblem än de som vi kan lösa med hittills nämnda metoder. Om vi vill värdera effekten av Huddinge sjukhus tillkomst på sjukligheten i Huddinge sjukhus upptagningsområde, så behöver vi åter andra metoder. I båda dessa fall talar vi om ingrepp på en annan nivå i samhällssystemet än den "enkla" individ- eller kliniknivån. Vi studerar aggregat av kliniker, sjukdomar, patientgrupper och befolkningsgrupper och försöker värdera kontentan av en viss insats. De metoder som hittills använts för denna typ av frågeställningar är väl snarast hämtade från sociologi, statsvetenskap eller historia. Det är kanske mer fråga om beskrivning av hur man tänker sig att en viss förändring skulle kunna påverka sjukligheten i en hel befolkning. Ett slags rimlighetsbedömning. Sådana beskrivningar tjänar dock ofta som referensramar för studier av effekter på lägre nivå, inom delsystem. Därigenom får de också den betydelsen att de "väljer" vad som tycks vara rimliga hypoteser att testa i studier med mer precision och säkrare metodologi, liksom att de "väljer" vilka studier som är av intresse att utföra.

Problemet att utvärdera insatser på högre aggre-

.....
Det kanske mer är fråga om beskrivning av hur man tänker sig att en viss förändring skulle kunna påverka sjukligheten i en hel befolkning
.....

gationsnivå är kanske mest diskuterat i samband med att man försöker värdera effekter av socialpolitik. Man har där genomgående funnit att nödvändig förutsättning för sådana värderingar är att man har tidsserier av "social indicators" och genom tidsserieanalys och/eller internationella jämförelser kan man möjligen dra slutsatser. McKeowns studie är ju egentligen ett exempel på detta, där han studerar mortalitetens förändring över tiden.

Sammanfattning

Om vi vill angripa ett brett spektrum av frågeställningar som alla har att göra med resultat av hälso- och sjukvårdsinsatser måste vi använda ett brett spektrum av metoder. Metoderna tycks vara mer sofistikerade och precisa när vi tittar på förändringar på lägre nivå, tex för en viss definierad patientgrupp och med en väldefinierad insats. Skillnaden mellan experimentella och icke-experimentella metoder är relativ, vill jag hävda.

När vi försöker lösa mer generella problemställningar, värdera resultatet av mer allomfattande hälso- och sjukvårdsinsatser, är metodiken svårare. Vid studier av effekten av socialpolitiska eller sjukvårdspolitiska förändringar på en hel population, är i allmänhet problemet att de data som skulle möjliggöra en tidsserieanalys eller en icke-randomiserad jämförelse saknas eller är av dålig kvalitet. Det kan följaktligen vara svårt att testa eller hitta kriterier för att falsifiera teorier om förändring av detta slag. Dock, måste detta anses centralt då sådana mer övergripande teorier eller referensramar "vägleder" en hel del av den forskning som bedrivs med mer precis metodik.

REFERENSER

1. *McKeown, T* (1976): *The Role of Medicine: Dream, Mirage or Nemesis?* London: The Nuffield Provincial Hospitals Trust Fund.
2. *Lindahl, O & Lindwall, L* (1978): *Vetenskap och beprövad erfarenhet, Natur och Kultur.*
3. *Gustafsson, R* (1980): *Kunskap på fel sätt? Om professionell kunskapstillämpning inom sjukvården. Expertinläga till Tema-forskningen om Hälso- och sjukvården i samhället, Universitetet i Linköping. Psykologiska institutionen, Stockholms Universitet.*
4. *Shapiro, A K* (1964): *Factors contributing to the Placebo effect, Am Journal of Psychotherapy, 43.*
5. *Miettinen* (1978): *Post-marketing drug research. Unpublished, Harvard.*
6. *Rothman, K* (1977): *Epidemiological methods in clinical trials, Cancer, 39.*
7. *Hesslow, G* (1979): *Medicinsk vetenskapsteori. Studentlitt. Lund.*
8. *Beeson, P B* (1977): *McKeown's The Role of Medicine. A clinician's reaction. Milbank Memorial Fund Quarterly. Summer.*
9. *Godber, Sir G* (1977): *McKeown's The Role of Medicine. Comments from a former Chief Medical Officer. Milbank Memorial Fund Quarterly. Summer.*
10. *Armstrong, D* (1975): *The changing Basis of Medical Knowledge. Unpublished. Dissertation. Bedford College Social Research Unit, London University.*