

Några problem i utvärdering av sociala interventioner och utfallsstudier

haluk soydan & bo vinnerljung

Med ökat intresse i studier av interventioners effekter har i större utsträckning frågor kring utfall blivit tydliga. Denna artikel beskriver några av de problem som författarna mött i utvärderingsverksamhet. Metodproblem i utfallsstudier kan vara mycket allvarliga och påverka analys och slutsatser på ett oberäkneligt sätt.

Inledning

Utfall kan definieras som avsedda och icke-avsedda förändringar hos enheter till följd av påverkan från sociala interventioner. Enheter som är föremål för interventioner och sociala program kan vara personer, familjer, grupper, bostadsområden, institutioner och organisationer.

Inom verksamhetsområden som socialtjänst, hälso- och sjukvård, kriminalvård har kravet på empiriska kunskaper om interventioners effekter ökat i Sverige

Haluk Soydan är professor i socialt arbete och arbetar som forskningsledare vid Centrum för utvärdering av socialt arbete.

Bo Vinnerljung är docent i socialt arbetet och arbetar som forskningsledare vid Centrum för utvärdering av socialt arbete.

under 1990-talet. Internationellt är detta inget nytt. Särskilt i USA har beslutsfattare och policy-ansvariga sedan 1960-talet ställt krav på professionella grupper att ta reda på och redovisa utfall av sociala interventioner och program (se Albaek 1988). Uppföljning av insatser, avläsning av insatsers effekter, informationsåterföring till insatsansvariga har också blivit ett kvalitetskriterium för professionellt arbete. Sammantaget har detta skapat en omfattande forskning kring utvärdering som främst under 1990-talet började sprida sig till Europa. Den amerikanska utvärderingsforskningen har använt sig av en rad olika designmodeller, från randomiserade kontrollerade försök till kvalitativa studier (Rossi & Freeman 1993, Oakley 2000).

Identifiering, kartläggning och analys av sociala interventioners utfall är förknippade med en lång rad kunskapsteoretiska, metodologiska och mättekniska utmaningar. Syftet med föreliggande artikel är att lyfta fram några vanliga metodologiska problem som genereras vid studier av sociala interventioners utfall.

Efter att ha berört några begreppsliga aspekter av utfall i sociala interventioner kommer vi mer specifikt att behandla fyra teman:

- 1) I utvärderingsstudier fokuseras ofta, och många gånger uteslutande, på ett utfallsmått. Vi kommer att argumentera för nödvändigheten av mätning och analys av flera utfallsvariabler i utvärderingar av sociala interventioner.
- 2) När en experimentell design inte kan användas i utvärderingsstudier och när särskilda grupper som t ex adopterade barn följs upp och jämförs med normalpopulationen kan en snedfördelning av specifika problem i undersökningsgruppen skapa problem i utfallsanalysen.
- 3) Tidens roll i utfallsstudier är av kritisk betydelse. Vi kommer att diskutera några besvärande problem med tidsfaktorn i utfallsstudier.
- 4) Att klienter inte fullföljer påbörjade sociala program – att de blir drop-outs, eller avvisade från programmet – skapar avgörande bortfallsproblem i utvärdering av sociala interventioner. Vi kommer att belysa problemet med avbrutna insatser i utvärderingsstudier.

Utfall i sociala interventioner – några begreppsliga aspekter

Innan dessa fyra teman diskuteras vill vi fästa uppmärksamheten på mångsidigheten av begreppet »utfall« i utvärderingar av sociala interventioner för att ge ett perspektiv på de frågor som mer ingående belyses i artikeln.

När sociala interventioner genomförs vill man naturligtvis åstadkomma en positiv förändring hos enheten (personen, gruppen, organisationen osv.). Exempel är målsättningen om att få personer att sluta med missbruk eller att återfalla i brott etc. I en stadsdel vill man kanske minska arbetslösheten och/eller minska kriminaliteten. När en social intervention genomförs aktualiseras minst tre centrala aspekter vad gäller förändringen hos enheten: 1) Har sociala interventionen genererat ett eller flera utfall? 2) Är effekterna avsedda och icke-avsedda? 3) Är effekterna resultat av den genomförda sociala interventionen och/eller externa faktorer? Frågan om flera utfall behandlas i nästa avsnitt. I det följande vill vi beröra de två senare aspekterna.

Avsedda effekter av sociala interventioner är lätta att identifiera i den bemärkelsen att sociala program vanligtvis brukar ange målet/målen för interventionen. Icke-avsedda effekter kan vara såväl icke önskvärda som önskvärda men är inte självklara och omedelbart synliga för utvärderaren. De kan spåras på flera olika sätt, exempelvis genom mätning av vad som rimligen är relevanta utfallsvariabler men som inte är angivna som direkta mål för interventionen. I KrAmi-utvärderingen (se annan

artikel i denna tidskrift) omfattas flera avsedda effekter: förbättrad arbetsmarknadsförankring och egenförsörjning samt avståndstagande från kriminalitet. Förändringar i programdeltagarnas familje- och sociala relationer, och förändringar i hälsotillstånd är icke-avsedda utfall i samband med exponering för KrAmi-programmet. Förbättrade familjerelationer och hälsoförhållanden är inte en del av KrAmi-programmets uttalade mål, men identifierades tack vare mätinstrumentets (Addiction Severity Index) förmåga att rutinmässigt mäta dessa två utfallsvariabler.

Ett centralt problem i utvärdering av sociala interventioner är frågan om ett givet utfall är en effekt av själva interventionen. Utfall som utvärderaren mäter hos enheter som exponeras för en intervention kan vara såväl interventionseffekter som effekter av externa faktorer. Det centrala problemet för utvärderaren är att om möjligt särskilja dessa två grupper av utfall. Det klassiska metodologiska greppet för att klara detta är användning av randomiserade experiment med kontrollgrupp (i andra hand kvasiexperimentella studier med jämförelsegrupp som uppstått utan forskarens medverkan) med för- och eftermätningar (se Boruch 1997).

Med hjälp av design och olika statistiska metoder försöker forskaren att kontrollera en mängd oönskade effekter av olika art: selektionseffekter, bortfall, statistiska effekter etcetera. Aspekter av selektionseffekter och, bortfall diskuteras särskilt i denna artikel. Vi vill dock nämna ytterligare två som försvårar tillvaron för forskare som arbetar med utfallsstudier. Individens livshistoria är ett »mätproblem«. Inter-

ventioner i behandlings- och socialt förändringsarbete löper ofta under en längre tid. Samtidigt händer också en rad andra saker i dessa individers liv. Någon eller några av dessa livshändelser som äger rum utanför interventionen kan ha avgörande betydelse för de utfall som utvärderaren är intresserad av, exempelvis att personer får barn, förälskar sig eller att de får en ny omvälvande livssyn. Lika »besvärligt« för utvärderingen är ett annat känt problem som brukar benämnas »personlig mognad«. Inom brottspreventionsområdet till exempel har det visat sig väsentligt att åtskilja interventionseffekter från individens personliga mognad över tid som bland annat har samband med biologisk ålder (refMcGuire 1995).

Nödvändigt att arbeta med flera utfallsmått

Sociala interventioner genererar i regel flera utfall samtidigt eller i en given tidssekvens, medan forskare, praktiker och beslutsfattare tenderar att intressera sig för ett utfall i taget. Ett typiskt exempel är forskningen om brottsåterfall. Medan metaanalyser huvudsakligen sysslar med utfallsvariabeln 'brottsåterfall' genererar interventioner som avser att begränsa eller förhindra brottsåterfall även andra (positiva) utfall.

Frågan om vad som är adekvata – och rimliga – mått på utfall av insatser är ständigt närvarande i utvärderingsstudier. Inom området insatser för asociala barn/ungdomar är återfall i brott (»recidivism«) det traditionella utfallsmåttet när olika behandlingsprogram jämförs med varan-

dra. Framförallt gäller detta i metaanalyser eller systematiska forskningssynteser (se t ex Lipsey & Wilson 1998; Lipsey 1992, 19956). Huvudskälet till denna starka fokusering på återfall är förmodligen det starkt uttalade intresset bland politiker och andra beslutsfattare för interventioner som leder till reducerad brottlighet. Inom detta område kom först i början 1980-talet kvalificerade litteraturoversikter (se t. ex. Gendreau & Ross 1979; Ross & Gendreau 1980). Under 1990-talet började metaanalyser publiceras, särskilt pionjärinsatser av Mark Lipsey (se t. ex. Lipsey 1992, 1995). Även om dessa metaanalyser fokuserar på återfall i brott som den huvudsakliga utfallsvariabeln finns möjligheter i underlaget till systematiska synteser av andra utfallsvariabler. Eftersom kunskapsdatabasen för Lipseys metaanalyser rutinmässigt registrerar ett stort antal variabler inklusive olika utfallsvariabler var det möjligt att studera några av dessa i en särskild studie (Wilson, Lipsey and Soydan 2002). Denna metaanalys bygger på 142 studier av interventioners effekter på ungdomsbrottlighet. Forskarna arbetar i analysen inte bara med utfallsvariabeln »återfall i brott« utan även med andra resultat som »anpassning till skolan«, »reducering av beteendeproblem« och attitydförändringar.

Andra studier har använt betydligt mer sammansatta mått på »hur det gått för klienterna«. Ibland har detta tvingat fram frågor om hur mycket som är rimligt att kräva för att en genomgången behandling ska anses som framgångsrik. Ett bra exempel är det svenska SWEDATE-projektet, som fann att om man bara hade kravet »drogfri« senaste 6 månaderna, då hade

det gått bra för 51 procent av 387 institutionsbehandlade missbrukare ett år efter behandling. Om man dessutom ställde krav på att inget annat missbruk (t ex av alkohol) skulle förekomma, sjönk siffran till 37 procent. När forskarna dessutom lade till kraven »inte mer än två månaders institutionsvård under uppföljningsperioden« och »ingen kriminalitet« blev det bara 22 procent »lyckade fall« kvar. Efter att ha lagt till ytterligare en rad krav (inkomst från arbete eller utbildning, ordnat boende, inga allvarliga psykiska problem vid uppföljningen, inget behov av socialt stöd m m (alla vanliga mål för behandling av missbrukare) återstod till sist bara 10 procent som klarade sig bra (Berglund et al. 1991).

Ett annat exempel på kombinationer av utfall är Maja Andersons (1976) klassiska registeruppföljning av Gustav Jonssons och Anna-Lisa Kälrvstens 222 Stockholmspojkar (1964) och 100 pojkar som varit på institutionen Skå i sin barndom (Jonsson, 1967). Hennes rapport är fortfarande en lärorik läsning idag, 25 år efter att den publicerades. Uppföljningen genomfördes när »pojkar« var 20–34 år. Anderson konstruerade sina utfallskriterier efter vad hon tolkade var den dominerande uppfattningen om en önskvärd utveckling i Sverige på 1960-talet. Hon använde en sjugradig skala på social anpassning. Längst ner på skalan (1-2) fanns de som ej levde upp till samhällets normer för att man »klarar sig«. De varken arbetade eller studerade, en del var omhändertagna av samhället eller var till och från föremål för samhällets åtgärder. Mitt på skalan (3-4) fanns de som klarade sig »fast lite knackigt« (i Ander-

son, 1976, s 22). De hade brister i vissa avseenden men det fanns också positiv information i registren om deras liv. Högst upp (5-7) fanns de som lyckats bra åtminstone i något avseende, exempelvis skaffat sig ett yrke. Likaledes fanns en liknande skala för självförsörjning med tre grupper. »Självförsörjande« var män som inte registrerats för kriminalitet eller fylleri (förutom lindriga förseelser) och som ej varit ideligen sjukskrivna och som 1971 hade en minimiinkomst. »Någon brist« var de som hade en registerindikation på problem i något av dessa avseenden. »Ej självförsörjande« var en grupp som klart inte uppfyllde de tre villkoren för »självförsörjande«. Även om kategoriseringarna innebar en del kvalitativa tolkningar som inte kan spåras i detalj i rapporten (jfr Börjeson & Håkansson, 1990), framstår Andersons sammansatta utfallsmått både rimliga och trovärdiga även idag.

David Magnussons och Håkan Stattins (med fleras) banbrytande forskning om barns utveckling över tid, särskilt risk för asocialitet i tonåren och i vuxen ålder, har långtgående implikationer för frågan om utfallsmått i utvärderingar av insatser för barn och ungdom. Deras arbeten baseras på flera decenniers studier av longitudinellt material, främst den s k IDA-undersökningen i Örebro. I den redan klassiska artikeln »Antisocial development: A holistic approach « (Stattin & Magnusson 1996) sammanfattas några centrala fynd. Författarna visar att anpassningsproblem över tid tenderar att samlas och förstärkas hos en liten grupp med dystert prognos (problemgravitering och -aggregering). De som är kriminella tenderar också att vara

de samma som använder droger etc. Ett illustrativt exempel finns i en studie av mobbade barn/ungdomar (Andershed, Kerr & Stattin 2000). Forskarna fann att det till stor del är samma barn som mobbar andra på skolgården, som bär vapen på stan och utövar våld i gatumiljön.

Extremgruppen med multiproblem är så dominerande i longitudinella material, att om de exkluderas i uppföljningar får enskilda riskfaktorer under barns uppväxt ett svagt prediktivt värde för senare utveckling (Stattin & Magnusson, 1996). Detta gäller för flera utfall, till exempel brottsbelastning och missbruk. Det är rimligt att anta att denna grupp även har en stark påverkan på utfall av interventioner riktade mot asociala ungdomar.

Av detta följer för det första att utvärderingar – när det är möjligt – bör arbeta med flera olika utfallsvariabler och med aggregeringar av dessa utfallsvariabler för att realistiskt kunna gradera resultatet av en genomgången behandling, exempelvis av asocialt beteende hos ungdomar. För det andra bör utvärderingar av insatser för barn och ungdomar med kvasiexperimentell design inte begränsas till att hålla kontroll av enskilda (risk-) variabelers distribution mellan undersöknings- och jämförelsegruppen, utan forskaren bör även söka efter variabelmönster, d v s leta efter individer som kan antas tillhöra extrema riskgrupper.

Snedfördelningar

Många sociala insatser är svåra, nära nog omöjliga, att utvärdera med experimentell design. Inom barnavården gäller det exem-

pelvis fosterhemsplaceringar och adoptioner. Dessa insatser har dessutom mycket långtgående syften över tid, ofta att ge utsatta barn en trygg uppväxt under många år. Mycket av den »utvärderande« kunskap vi har om exempelvis långvarig fosterhemsvård kommer från studier av före detta fosterbarn i vuxen ålder (se översikt i Vinnerljung, 1996.). Man brukar här tala om »globalutfall«, vilket ävengäller adoptionsforskningen. I dessa studier jämförs vuxna före detta fosterbarn och adopterade med jämnåriga i normalpopulationen (se t ex Vinnerljung 1995), alternativt med andra grupper som haft jämförbara utgångslägen i tidig barndom (t ex Vinnerljung, 1996; Vinnerljung & Ribe, 2001). Ofta undersöks förekomst av relativt sällsynta oönskade utfall, såsom mortalitet, psykisk sjukdom och allvarlig kriminalitet (t ex Hjern, Lindblad & Vinnerljung, 2002).

Flera adoptionsforskare har under senare år uppmärksammat problemet med sneda fördelningar (»skewed distributions«). I en diskussion om tillsynes motstridiga resultat i adoptionsforskningen – högre förekomst i kliniska populationer av adoptivbarn än icke-adopterade men avsevärt mindre skillnader i populationsbaserade undersökningar – visade Jeffery Haugaard (1998) på hur dessa resultat teoretiskt kunde förklaras av en snedfördelning inom adoptionsgruppen jämfört med normalpopulationen. Om det fanns en liten grupp med mycket dåliga utfall inom adoptionsgruppen, kunde en stor del av skillnaderna jämfört med en normalgrupp förklaras av detta fenomen, exempelvis om forskarna huvudsakligen använde olika former av medelvärden i sina

analyser. Minst två studier har visat att det verkar finnas empiriskt stöd för denna hypotes (Sharma et al., 1998; Miller et al., 2000). Skillnaderna mellan adoptivbarn avseende exempelvis skolprestationer, skolproblem, psykiskt välmående, drogbruk m m var i dessa undersökningar störst bland de fem procent adoptivbarn som hade sämsta värden i mätningarna. Det fanns med andra ord en tydlig »puckel« av mycket belastade individer bland adoptivbarn i ena ändan av den antagna normalfördelningskurvan. Problemet återkommer i olika typer av social forskning som använder sig av sällsynta, negativa utfall – exempelvis mortalitet – som indikationer på insatsens långsiktiga effekter. Eftersom forskaren då arbetar med dikotoma data (förekomst/ej förekomst) blir det svårt att veta vad resultaten av en jämförelse med exempelvis normalbefolkningen säger oss. Är det något allmänt om risker för undersökningsgruppen, exempelvis före detta samhällsvårdade barn? Eller speglar resultaten istället att en liten del av undersökningsgruppen har en rejält dålig livssituation, medan skillnaderna för resterande delen av gruppen inte är så stora jämfört med andra jämnåriga?

Detta problem pekar på att studier av sällsynta utfall har sina uppenbara begränsningar i utfallsstudier, även de som syftar att undersöka »globala utfall«. Ofta behövs det data om hela undersökningsgruppen för att kunna säga något mera konklusivt. Erfarenheterna från adoptionsforskningen pekar på betydelsen av analysmetoder där även fördelningen av utfall i undersöknings- och en jämförelsegrupper undersöks och analyseras.

Tiden ställer till med problem

Sociala problem och beteendestörningar som blir föremål för interventioner genereras, utvecklas och får sitt »naturliga« förlopp i ett tidsperspektiv. Alkoholism, drogproblem, kriminalitet, skolk etc uppstår inte över en natt utan utvecklas över en tid. På samma sätt pågår behandling av sociala problem och beteendestörningar över tid, ibland i flera år. I studier av utfall efter avslutad intervention uppstår då en rad frågor som handlar om tid: när uppnås vad man kan kalla för »en effekt« i samband med en intervention? När är det bäst att avläsa en sådan effekt? Under eller efter behandlingen? Om svaret är »efter behandlingen«, uppstår en ny fråga: strax efter avslutad behandling eller först efter en tid? Om det senare gäller— ska det vara efter sex månader, ett år, tre år, tio år efter behandlingen? Med andra ord : hur beständig över tid är behandlingens effekter?

Den kloka forskaren inser att det inte finns ett standardiserat och receptboksbaserat svar på frågan när en intervention kan tänkas ha effekt, om hur länge interventionens effekter varar och när mätningar av utfall ska ske. Man kan kanske prata om interventioners kontextualitet i bemärkelsen att det finns en mängd skillnader när man jämför olika interventionsområden, som till exempel alkoholbehandling, brottsprevention och försörjningsstöd. Observera att med kontextualiteten avses inte här att refererar till attvarje individuell intervention är unik och därmed kontextuell med den konsekvensen att det inte går att generalisera till andra liknande interventions-situationer. Snarare är det fråga om skillna-

der av en annan karaktär. Beakta exemplet med att interventionen »blindtarmsoperation« har en radikal engångseffekt som är omedelbar vad gäller upphörande av den akuta inflammationen och som dessutom är beständig i tiden under individens livsförlopp. Blodstryckspatienten, ceteris paribus, måste däremot medicineras dagligen eftersom blodstryckskontrollerande preparat verkar under begränsad tid (exempelvis ett dygn) även om effekten är nästan omedelbar. Det finns med andra ord en mångfald av samverkansmönster mellan intervention och utfall med tiden som mellanliggande faktor. Här följer fler exempel som antyder vikten av dessa problem.

Ett färskt exempel visar hur snabbt effekterna av behandling av alkoholproblem interventioner eroderas. I en nyligen utgiven volym redovisar Statens beredning för medicinsk utvärdering metaanalyser om intervention mot riskfylld alkoholkonsumtion (s. k. sekundär prevention av alkoholproblem). I denna studie har man kommit fram till slutsatsen att det är »osäkert om insatser ger effekter efter mer än två år eller om man då måste upprepa interventionen« (SBU 2001:I s 55). Om denna slutsats är riktig och kan bekräftas av ytterligare studier är det givet att sekundära preventiva interventioners positiva utfall har i genomsnitt en beständighet på högst två år och att interventionerna måste upprepas efter denna period om man vill upprätthålla preventiva effekter.

Ett annat exempel kommer från drogbehandlingsområdet. I det tidigare nämnda SWEDATE Projektet, som genomfördes mellan 1981 och 1983, följde forskarna upp effekterna av drogbehandlingsprogram

hos drygt 1100 missbrukare åtta år efter insatsen. Ett för vår diskussion intressant resultat var att hela 75,2 procent av populationen hade tolv eller flera registrerade återfall i brott eller var avlidna. Dessutom var de tidigare noterade skillnaderna mellan olika behandlingsprogram nu utjämnade (Bergmark et al, 1996). Återigen ser vi här hur effekter sakta men säkert eroderas med tiden.

I en annan och betydligt mindre studie (Nyström & Soydan 1997) exponeras ett överraskande resultat. I denna utfallsstudie av ett socialt rehabiliteringsprogram för kriminella och arbetslösa, KrAmi, följdes 29 programdeltagare upp för andra gången 1996, dvs fyra år efter avslutat eller avbrutet program. I den första uppföljningen studerades två grupper, de som fullgjorde programmet och de som var drop-outs (Lindberg & Soydan 1993). Denna visade bättre resultat för gruppen som hade fullföljt programmet. Fyra år senare hade den första gruppen fortfarande positiva utfall i termer av arbete, bostad, familj, missbruk och självförtroende. Detta resultat hade alltså bestått i fyra år efter avslutat program. Men därutöver hade ett flertal i drop-out gruppen också förbättrat sin situation sedan första uppföljningen, vilket kan tolkas på olika sätt – inte minst i termer av social mognad som vi tidigare nämnt.

Sammanfattningsvis menar vi att det finns reella problem vad gäller när och hur interventioners effekter gör sig gällande och hur beständiga de är. Det återstår dock att systematiskt kartlägga typer av problem som empiriska studier visar och att utveckla strategier som kan hantera dessa problem. En rimlig »basrekommendation«

är att utvärderingar av program/insatser som har långtgående ambitioner om effekters beständighet över tid, måste innehålla upprepade mätningar vid flera tillfällen efter behandlingens avslutning. Samma gäller utvärderingar av program/insatser som enbart på ideologisk (ej empiriskt baserad) grund påstår att effekterna av behandlingen har hög beständighet över tid. Det är viktigt att komma ihåg att vissa interventioner med goda intentioner – och mycket resurser – faktiskt också kan ha en skadlig inverkan över lång tid, (se t ex McCord 1978).

Drop-outs

Drop-out – avbrutna insatser – är ett mycket vanligt problem i socialt arbete (se diskussion i Vinnerljung, Sallnäs & Kyhle-Westermark 2001). Deltagare hoppar av behandlingsprogram i förtid, blir avvisade från familjehem och institutioner, avviker från behandlingshem etc.

Inom psykiatri och psykoterapi är problemet välkänt sedan länge. Joel Fisher konstaterade i en klassisk översiktsartikel 1978 (med titeln Does anything work?) att upp till 50 procent av psykoterapiklienter kom aldrig tillbaka efter första behandlingssamtalet, och upp till 80 procent föll bort före sex planerade samtal. Hans slutsats var närmast brutal: – Thus, psychotherapy appears to reach – and hold – only a very small percentage of those people who may need it (Fisher 1978 s 221). Även senare forskningsöversikter över såväl vuxen- som barnpsykiatrisk terapeutisk behandling har funnit att höga dropout-siffror är snarare

regel än undantag (se t ex Garfield, 1994; Kadzin, 1994a, 1994 b).

Problemet är dock avsevärt mindre uppmärksammat i socialt arbete, trots tydliga bevis om problemets omfattning i behandlingsforskningen. I en översikt av öppenvårdsinsatser för föräldrar som misshandlat sina barn redogör Corcoran (2000) för flera studier där 70-80 procent av föräldrarna aldrig genomförde behandlingen. Det svenska SWEDATE-projektet som utvärderade institutionsvård för närmare 1.200 drogmissbrukare fann att hela 60 procent genomförde aldrig behandlingen. För vuxna missbrukare var siffran 70 procent (Berglund et al, 1991). Detta är inte en extrem siffra. En norsk utvärdering av ett Phoenix-house-projekt för närmare 150 narkomaner fann att hela 87% av de manliga deltagarna avbröt behandlingen i förtid (Ravndal & Vaglum 1995). I SBU:s rapport (2001) om behandling av alkohol- och narkotikaproblem redovisas retentionssiffror (kvarstannande i behandling) för behandling av narkotikamissbrukare på mellan nära 0 och 100 procent. Just kvarstannande i behandling används till och med som »ett primärt mått på behandlingseffekt« (vol. 1, s 21), trots att drop-out-processen rimligtvis måste spegla en slags anrikning av den ursprungliga undersökningsgruppen (jfr Edwards & Rollnick 1997).

När Patricia Chamberlain och hennes medarbetare i USA startade försök med fosterhemsplaceringar av kriminella tonåringar i kombination med bland annat familjeterapi¹ var deras första fråga: är

1 S k Multi-dimensional treatment foster-care. Metoden beskrivs utförligt i Hansson 2001.

det överhuvudtaget möjligt att genomföra familjehemsplaceringar av denna grupp, utan höga sammanbrottsiffror?² (Chamberlain & Reid 1998; Hansson 2001 m fl; jfr Baker 1989)? I en nyligen publicerad svensk studie undersökte Vinnerljung med medarbetare en nationell kohort av 13-16-åringar som placerades i dygnsvård 1991. Ungdomarna följdes i akter under fem år, eller till 18-årsdagen. Några exempel från resultaten: i »vanliga familjehem«, (ej släkt med barnet) havererade 41-51 procent av placeringarna i förtid, siffran beroende på val av definition av sammanbrott. I en grupp små enskilt ägda institutioner var sammanbrottsfrekvensen 52 procent. Den vanligaste placeringen i kohorten var att en tonåring med asocialt beteende sattes i ett »vanligt fosterhem«. 57-67 procent av dessa placeringar slutade med sammanbrott (siffran återigen beroende på val av definition). Resultaten för samma målgrupp i institutionsvård var inte mycket bättre: 40-50 procent, beroende på definition och ägandeform. (Vinnerljung, Sallnäs & Kyhle-Westemark, 2001). De enda grupperna med jämförelsevis låga sammanbrottsiffror (<20 procent) var släktinghem³, jourfosterhem och s k §12-hem (tidigare benämnda ungdomsvårdsskolor).

2 Med sammanbrott avses att placeringen »spricker«, d v s avslutas på ett uppenbart oavsiktligt sätt. Vanligaste skälet är att fosterhemmet/institutionen inte vill fortsätta med placeringen, eller att barnet/den unge rymmer. För diskussion om definitioner och för översikt av forskning om sammanbrott i barn- och ungdomsvården se Vinnerljung, et al, 2001).

3 Fosterhem där fosterföräldrarna är släkt med barnet, vanligtvis mormor och morfar.

I utvärderingar och till och med i översikter av utvärderingsresultat är det dessvärre inte ovanligt att drop-out-frekvenser ej redovisas eller förs bort från diskussionen, ibland med motivet »data saknas för gruppen«. Detta är självfallet mycket otillfredsställande och leder sannolikt till felaktiga slutsatser i många fall. Om vi exempelvis ska utvärdera ett enskilt behandlingshem för tonåringar där majoriteten av placeringar aldrig blir genomförda – ungdomarna avvisas av personalen efter disciplinkonflikter, rymmer etc. – blir det grovt missvisande att i en uppföljning bara redovisa hur det går för den minoritet som fullföljer behandlingsprogrammet. Låt oss konstruera ett hypotetiskt exempel: säg att vi jämför två behandlingshem, och i analysen bara tar med de som fullföljt behandlingen. Resultaten visar inga skillnader mellan programmen. Hem A har avvisat 50 procent av eleverna under behandlingens gång (t ex med motivet »de passade inte in i gruppen«) medan hem B har en drop-outfrekvens på 20 procent. Denna stora, avgörande skillnad kommer inte att synas i resultaten om utvärderaren bara tar med de som fullföljt behandlingen. Då gynnas hem A medan slutsatsen »ingen skillnad« blir orättvis och felaktig för hem B.

Vi vet också att drop-out/sammanbrott är en selektionsprocess, där de som avbryter eller avvisas från pågående behandling är mer problembelastade än de som stannar kvar (se t. ex. Kadzin, 1997, Newton et al, 2000 för exempel från barn- och ungdomsvården; Edwards & Rollnick, 1997 om interventioner för alkoholmissbrukare). Men även efter kontroll för detta, visar en

mycket stor metaanalys av insatser för sociala barn/ungdomar att behandlingsprogram med höga drop-out-siffror tenderar tydligt att visa sämre resultat (Lipsey & Wilson, 1998). I ljuset av den kunskapen kan man faktiskt se ungdomar som »röstar med fötterna« och avviker från olika behandlingsprogram som välinformerade konsumenter.

Hur bör det vara? I den »klassiska« utvärderingsforskningen där sk randomiserade experimentella studier är den »gylene« designen, gäller följande vedertagna regel: once randomized – analyzed (Boruch 1997). Det finns två huvudsakliga motiv för denna fundamentala regel. Det ena är vetenskapligt baserat, och berör utvärderingens validitet. Gruppen som avbryter behandlingen skiljer sig sannolikt från den som fullföljer, alltså speglar dropouts en selektionsprocess som inte kan antas vara densamma i experimentgruppen som i jämförelsegruppen. Särskilt självklart blir detta om jämförelsegruppen består av individer på en väntelista där man har slumpmässigt valt ut vilka som ingår i experimentgruppen. Det andra motivet är policy baserat och berör utvärderingens anknytning till »real-world«-förhållanden. I verkligheten utanför studien är det sannolikt att en behandling som erbjuds, föredras av policyskäl eller som ingår i lagstiftade insatser, kommer att avvisas av en del klienter och andra kommer att avvika eller avvisas från behandlingen. Det som händer under mer kontrollerade former i en utvärderingsstudie kan då sägas vara bevis på vad som skulle hända om dessa behandlingsprogram skulle implementeras i stor skala i ett land eller region, utanför

studien. Drop-out speglar med andra ord ett verkligt förhållande, inte ett av forskarna konstruerat, och det är därför av central vikt att även omfattningen av avbrutna behandlingar studeras och analyseras i en undersökning av behandlingens effektivitet.

Vi upprepar: den klassiska regeln säger att utfall ska studeras, analyseras och redovisas för alla individer som inledningsvis placeras i en behandlingsgrupp (experimentgrupp) och likaledes för alla fall som från början ingår i en jämförelsegrupp. Detta, menar vi, bör även gälla i utvärderingsstudier oavsett designform, så långt som det är praktiskt möjligt. Om det inte är möjligt att följa denna huvudregel bör forskaren klart redovisa detta och motivera varför det inte har gått. Han/hon bör även försöka bedöma konsekvenserna av att dessa data saknas, och slutligen diskutera hur detta påverkar studiens generaliserbarhet.

Avslutning

Den svenska inomvetenskapliga debatten om utvärderingar av sociala interventioner finns många kritiska röster. Ett fåtal menar till och med att det inte går att göra meningsfulla utvärderingar av exempelvis socialt arbete, eftersom metodproblemen är oöverstigliga. Även om vi anser att sådana resonemang i princip är ovetenskapliga (hur kan vi så säkert veta att det aldrig kommer att gå?), menar vi att det är viktigt att öppet och konkret diskutera de reella metodproblem som finns. I denna artikel har vi försökt belysa aspek-

ter av några konkreta problem, och pekat på ställningstaganden och utvägar ur de dilemma som uppstår. Vi vill betona att det främsta syftet med denna artikel har varit att uppmärksamma problem som sällan diskuteras i handböcker om utvärderingsmetodologi. De slutsatser som vi drar av det som presenterats i artikeln är:

Visst – det finns problem med enskilda utfallsmått. I verkligheten åstadkommer praktiskt socialt arbete flera utfall genom en given intervention. Även forskarna själva mäter i många undersökningar flera utfallsvariabler, utan att sedan explicit arbeta med dessa i analysen, och vi bör när det är möjligt arbeta med mer sammansatta kriterier på utfall

Var vaksam mot förekomst av sällsynta fenomen som utfallsmått i uppföljningar. De kan spegla snedfördelningar som vilseleder både forskaren och konsumenten av den gjorda genomförda utvärderingen. Sök i stället efter utfallsmått som ger information om alla samtliga studieenheter i både undersöknings- och jämförelsegruppen. Undersök dessutom om en liten multi-belastad problemtyngd extremgrupp har ett högt förklaringsvärde för resultaten.

Frågan om hur man ska hantera tidsfaktorn på ett rimligt sätt i uppföljningar har inte ett givet svar. Här bör varje forskare vara öppen för såväl kliniska erfarenheter som konsensusuppfattningar i den internationella forskningen.

Drop-outs och avbrutna behandlingar är ett stort och reellt problem i utvärderingar av sociala interventioner, och är sannolikt större ju svårare problem som interventionen avser att påverka (se t ex sammanställ-

ning i SBU, 2001). Grundprincipen »once randomized – analyzed« – överförd även till studier med annan design än det randomiserade experimentella försöket – bör vara utgångspunkten. Det gäller både för analys och för diskussion om de resultat som undersökningen redovisar.

Vi tackar Annika Puide och Marie Sallnäs som lämnat värdefulla kommentarer på en tidigare version av denna artikel.

Referenser

- Albaek, E. (1988) Fra sandhed till information. Evalueringsforskning i USA – förr och nu. Viborg: Akademisk Forlag.
- Andershed H., Kerr, M. & Stattin, H. (under tryckning) Bullying in school and violence in the streets: are the same people involved? Journal of Scandinavian Studies in Criminology and Crime Prevention.
- Anderson, M (1976) Hur går det för 50-talets Stockholmspojkar? En uppföljning av 222 vanliga skolpojkar och 100 Skå-pojkar. Stockholm: Monografier utgivna av Stockholms kommunalförvaltning, nr 38.
- Baker, J. (1989) Therapeutic foster parent: professionally or emotionally involved parent? Child and Youth Services, vol 12, s 149–157.
- Berglund, G., Bergmark, A., Björling, B., Grönbladh, L., Lindberg, S., Oscarsson, L., Olsson, B., Segraeus V., & Stensmo C. (1991) The SWEDATE-project: interaction between treatment, client background and outcome in a one-year follow-up. Journal of Substance Abuse Treatment, vol 8, s 161–169.
- Bergmark, A., Lindberg, S., Olsson, B. & Oscarsson, L. (1996) A long-term follow-up of residentially treated drug abusers. Paper presenterat vid 22:nd Annual Alcohol Epidemiological Symposium of the Kettil Nruun Society for Social and Epidemiological Research on Alcohol, Edinburgh June 3–7, 1996.
- Boruch, R. (1997) Randomized experiments for planning and evaluation. A practical guide. Thousand Oaks, Ca: Sage Publ Inc, Applied Social Research Methods series vol 44.
- Börjeson, B. & Håkansson, H. (1990). Hotade, försummade, övergivna. Är Familjehemsplacering en möjlighet för barnen? Stockholm: Rabén & Sjögren.
- Chamberlain P & Reid J (1998) Comparison of two community alternatives to incarceration for chronic juvenile offenders. Journal of Consulting and Clinical Psychology, 66:4, s 624–633.
- Corchoran, J. (2000) Family interventions with child physical abuse and neglect: a critical review. Children and Youth Services Review 22: 7, s 563–591.
- Edwards, A. & Rollnick, S. (1997) Outcome studies of brief intervention in general practice: the problem of lost subjects. Addiction 92:12, s 1699–1704.

Haluk Soydan & Bo Vinnerljung: Problem i utvärdering av...

- Fisher, J. (1978) Does anything work? *Journal of Social Service Research*, 1:3, s 215–243.
- Garfield, S. (1994) Research on client variables in psychotherapy. I Bergin, A. & Garfield, S. (red) *Handbook of Psychotherapy and Behavior Change*. New York: Wiley & Sons.
- Gendreau, P. & Ross, B. (1979) Effective correctional treatment: Bibliotherapy for cynics. *Crime and Delinquency* 25: 463–489.
- Hansson K (2001) Familjebehandling på goda grunder. En forskningsbaserad översikt. Stockholm: Gothia/CUS.
- Haugaard, J. (1998) Is adoption a risk factor for the development of adjustment problems? *Clinical Psychology Review* 18:1, s 47–69.
- Hjern, A., Lindblad, F. & Vinnerljung, B. Suicide, psychiatric illness and social maladjustment in intercountry adoptees in Sweden. *The Lancet*, 360, s 443–448.
- Jonsson, G. (1967) Delinquent boys, their parents and grandparents. *Akadem. avhandl. Acta Psychiatrica Scandinavica*, vol 43, 1967, suppl 195.
- Jonsson, G. & Kälvesten A-L. (1964) 222 Stockholmspojkar. Stockholm: Almqvist & Wiksell.
- Kazdin, A. (1994a) Methodology, design and evaluation in psychotherapy research. I Bergin, A. & Garfield, S. (red) *Handbook of Psychotherapy and Behavior Change*. New York: Wiley & Sons.
- Kazdin, A. (1994b) Psychotherapy for children and adolescents. I Bergin, A. & Garfield, S. (red) *Handbook of Psychotherapy and Behavior Change*. New York: Wiley & Sons.
- Kazdin, A. (1997) Practitioner review: psychosocial treatments for conduct disorder in children. *Journal of Child Psychology and Psychiatry* 38: 2, s 161–178.
- Lipsey, M. (1992) Juvenile delinquency treatment: A metaanalytic inquiry into the variability of effects in T. D. Cook, H Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, F. Mostellar, red., *Meta-analysis for explanation. A Casebook*. New York: Russell Sage Foundation.
- Lipsey, M. (1995) What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? I J. McGuire, ed., *What works: Reducing reoffending. Guidelines from research and practice*. Chichester: John Wiley & Sons.
- Lipsey, M. & Wilson, D. (1998) Effective intervention for serious juvenile offenders: a synthesis for Research. I Loeber, R. & Farrington, D. (red) *Serious & violent juvenile offenders. Risk factors and successful interventions*. Thousand Oaks, Ca: Sage Publications.
- Lindberg, O. & Soydan, H. (1993). Att bli människa på nytt: En studie av socialt förändringsarbete vid KrAmi i Örebro. Högskolan i Örebro: Forskargruppen i socialt arbete, Rapport 2.
- McCord, J. (1978) A thirty-year follow-up of treatment effects. *American Psychologist*, March 1978, s 284–289.
- McGuire, J. Red. (1995) *What Works. Reducing reoffending. Guidelines from research and practice*. Chichester: John Wiley & Sons.
- Miller, B., Fan, X., Christensen, M., Grotevant, H. & van Dulmen, M. (2000) Comparisons of adopted and nonadopted adolescents in a large, nationally representative sample. *Child Development* 71:5, s 1458–1473.
- Newton R, Litrownik A & Landsverk J (2000) Children and youth in foster care: disentangling the relationships between problem behaviors and number of placements. *Child Abuse & Neglect*, 24:10, s 1363–1374.
- Nyström, S. & Soydan, H. (1997) Förändringars beständighet och social värdighet en uppföljning av KrAmi-verksamheten i Örebro. *Socialvetenskaplig tidskrift*. 1: 22–42.
- Oakley, A. (2000) *Experiments in Knowing. Gender and method in the social sciences*. New York: The New Press.
- Ravndal, E. & Vaglum, P. (1995) Psykiske belastninger og frafall blant stoffemisbrukere i behandling. Oslo: Sosial- og Helsedepartementet – Rusmiddeldirektoratet.
- Ross, R. & Gendreau, P. (red) (1980) *Effective correctional treatment*. Toronto: Butterworths.
- Rossi, P. H. & Freeman, H. E. (1993) *Evaluation. A systematic approach*. Newbury Park: SAGE Publications.
- Rutter M., Giller H. & Hagell A. (1998) *Antisocial*

- behavior by young people. Cambridge: Cambridge University Press.
- SBU (2001) Behandling av alkohol- och narkotika-problem. En evidensbaserad kunskapssammanställning, volym 1+2. Stockholm: SBU.
- Sharma, A., McGue, M. & Benson, P. (1998) The psychological adjustment of United States adopted adolescents and their nonadopted siblings. *Child Development* 69:3, s 791–802.
- Stattin, H. & Magnusson, D. (1996) Antisocial development: A holistic approach. *Development and Psychopathology*, 8, 617-645.
- Vinnerljung, B. (1995) Mortalitet bland fosterbarn som placerats före tonåren. *Socialvetenskaplig Tidskrift* 2:1, s 60–72.
- Vinnerljung, B. (1996) Fosterbarn som vuxna. Akad avhandling. Lund: Arkiv Förlag.
- Vinnerljung, B. & Ribe, M. (2001) Mortality after care among young adult foster children in Sweden. *International Journal of Social Welfare* 10:3, s 164–173.
- Vinnerljung, B., Sallnäs, M. & Kyhle-Westermark, P (2001) Sammanbrott vid tonårsplaceringar – om ungdomar i fosterhem och på institution. Stockholm: CUS/Socialstyrelsen.
- Wilson, S. J., Lipsey, M. & Soydan, H. (2002) Are mainstream programs for juvenile delinquency less effective with minority youth than majority youth? A meta-analysis of outcomes research . *Research on social work practice* (utkommer)

Summary

Some problems in evaluation of social interventions and outcome studies

The aim of this article is to discuss methodological problems that frequently occur in outcome studies of social interventions. Over the years we have observed a number of problems that disturb outcome studies. Four problems that evaluation researchers should be careful about are focused on.

1. Researchers (as well as practitioners who use results of those studies) often focus on only one outcome variable in intervention studies. However, social interventions usually generate a set of outcomes in a given point of measurement. Also, researchers often measure several outcomes without necessarily analyzing those registered outcomes. It seems more

efficient to work with a multi-outcome approach trying to understand impact of social interventions.

2. Sometimes when experimental design in evaluation is not an option, follow up studies are used instead (e.g. for adopted children). Comparisons with the normal population may then lead to problems in the final analysis as specific groups may have skewed distributions. We draw the reader's attention to these problems especially when using rare phenomena as outcome variables.

3. Social and behavioral problems emerge and develop over a time span rather than instantly. Interventions also take time. A

crucial question is then raised: when does an effect occur in the course of an intervention and when should effects be measured? We find no self-evident guidance in the literature and recommend openness to clinical experiences and consensus among researchers.

4. Dropouts are a very common phenomenon in social work practice. Participants

in social intervention programs leave the program at an early stage, are expelled. This problem affects outcomes of the intervention program. We draw the reader's attention to the importance of including all individuals who once were allocated in treatment as well as control or comparison groups.