

Virker MST?

*Kommentarer til en systematisk forsknings-
oversikt og meta-analyse av MST*

.se.fis
.dk.no



TERJE OGDEN
Forskningssjef
Terje.Ogden@atferd.unirand.no
KRISTINE AMLUND HAGEN
Forsker
k.a.hagen@atferd.unirand.no

Norsk senter for studier av
problematferd og innovative
praksis, Unirand, Universite-
tet i Oslo

I denne artikkelen kommenterer vi den systematiske forskningsoversikten og meta-analysene av MST fra Nordic Campell Center. Studien vekker oppsikt, – og har dermed bidratt til fornyet diskusjon om bruk av MST og evidensbaserte metoder. For ifølge Nordisk Campell Center er det ikke noe ...”der tyder på, at MST er bedre for de unge end alternative behandlingsformer. Der er dog heller ikke tegn på, at MST er dårligere”. Men kan det være at det er selve meta-analysen som er prematur, – og at den derfor egner seg dårlig som grunnlag for å vurdere hvor effektiv MST er i arbeidet med alvorlige atferdsproblemer, og da særlig i de nordiske landene?

Multisystemisk terapi (MST) har blitt introdusert i Norge som en familie- og nærmiljøbasert behandlingsmetode i arbeidet med alvorlige atferdsproblemer blant ungdom. Siden 1999 er MST etablert i alle fylker, og mer enn 1600 ungdommer og deres familier har gjennomgått behandlingsprogrammet.

Bakgrunn:

- Multisystemisk terapi (MST) er et hjemmebasert, intensivt og familiebasert tilbud til ungdom mellom 13–17 år med alvorlige atferdsvansker. Behandlingen foregår i ungdommens familie og nærmiljø, og både skole/arbeidsgiver, politi, andre hjelpetjenester og familie trekkes inn. Gjennom MST får foreldrene hjelp til å styrke sin oppdragerpraksis og bli styrket i å utøve sin foreldrerolle på linje med andre tenåringsforeldre. Målet med MST er å øke familiens samhold og omsorg for hverandre, gjenopprette kontakt med positive venner, bedre ytelsene i skole/arbeidsliv, øke deltakelse i positive fritidsaktiviteter, styrke kontakten mellom familien og omgivelsene og hjelpe familien til å løse problemer selv i fremtiden.
- Norske evalueringer av MST (Ogden & Halliday-Boykins 2004 og Ogden & Amlund-Hagen 2006) viser at MST kan forebygge plasseringer utenfor hjemmet, og redusere problematferd i signifikant grad sammenlignet med andre barnevernstiltak.
- En systematisk forskningsoversikt fra Nordisk Campell Center, utarbeidet av Julia Litell (2005), bygget på åtte evalueringer av MST fra USA, Canada og Norge. Studien kom til det motsatte resultat av de primæranalysene den bygget på (www.sfi.dk/sw24214.asp).
- I denne artikkelen kommenterer Terje Ogden og Kristine Amlund-Hagen den systematiske forskningsoversikten og meta-analysene av MST fra Nordic Campell Center. De hevder Litell stiller urimelige høye standardiseringskrav til design, variabler og prosedyrer i undersøkelser som skal inngå i en meta-analyse, noe som medførte at bare åtte primærstudier ble inkludert, mens det vanligvis tas med 50–200 studier.
- Campbell-samarbeidet, som Nordisk Campbell Center er en del av, er et internasjonalt nettverk av forskere, praktikere og beslutningstakere. Formålet er å utarbeide, oppdatere og formidle systematiske forskningsoversikter over effekten av sosialt arbeids metoder og resultater (www.nc2.net).

Den norske evalueringen ble foretatt 6 (Ogden & Halliday-Boykins 2004) og 24 måneder (Ogden & Amlund Hagen 2006) etter oppstart av behandlingen. Den viste at MST forebygger plasseringer utenfor hjemmet og reduserer problematferd i signifikant grad, sammenlignet med andre barnevernstiltak. MST har også blitt evaluert av programutvikleren Scott Henggeler og hans kolleger i USA, og disse underbygger at MST gir gode resultater for målgruppen (for en oversikt, se Curtis m.fl. 2004). En replikasjonsstudie med mindre oppmuntrende resultater ble utført i Canada (Lescheid & Cunningham 2002).

I en protokoll fra Campbell Collaboration presenterer Littell, Popa og Forsythe (2005) en systematisk forskningsoversikt og en meta-analyse av resultatene fra 8 eva-

lueringer av MST i USA, Canada og Norge¹. Forskningsoversikten kommer til det motsatte resultatet av primærstudiene som meta-analysen bygger på, - og konkluderer med at det er prematurt å trekke konklusjoner om MSTs effektivitet sammenlignet med andre tjenester. Resultatene fra Littell og kollegers meta-analyse har vakt oppsikt blant forskere såvel som praktikere i de nordiske land, og har skapt usikkerhet om MST virker. Vi på vår side mener altså at det er meta-analysen som er prematur, og dermed egner seg dårlig som grunnlag for å vurdere hvor effektiv MST er i arbeidet med alvorlige atferdsproblemer, særlig i de nordiske landene. The Cochrane Collaboration, som har utviklet retningslinjer for hvordan meta-analyser bør utføres og som gjentatte ganger blir sitert av Littell og kolleger (2005), skriver følgende: "If you identify and suspect that important diversity or heterogeneity is present in your review, there are several options open to you. [...] one option is that of not performing a meta-analysis. An unwise meta-analysis can lead to highly misleading conclusions (The Cochrane Collaboration).

I denne artikkelen presenterer vi noen hovedtemaer fra den systematiske forskningsoversikten (Littell m.fl. 2005) og den påfølgende diskusjonen mellom Littell og Henggeler m.fl. Disse er publisert i *Children and Youth Services Review* med følgende titler: "Lessons from a systematic review of effects of Multisystemic Therapy" (Littell 2005a), "The Littell paper: Methodological critique and meta-analysis as a Trojan horse" (Henggeler Schoenwald, Borduin & Swensson 2005) og "The case for Multisystemic Therapy: Evidence or orthodoxy?" (Littell 2005b).

Hvilke undersøkelser er med i meta-analysen?

Littell m.fl. (2005) har stilt bestemte metodiske krav til undersøkelser som skal inngå i meta-analysen. De skal ha benyttet et kvantitativt gruppedesign med randomisert fordeling av deltakerne på tiltaksalternativene (randomised controlled trials: RCT), samt anvendt standardiserte undersøkelsesinstrumenter, slik at pre-post resultater skal kunne sammenlignes på tvers av undersøkelsene. Littell har inkludert alle MST-undersøkelser som omfatter ungdom med alvorlige sosiale, emosjonelle og/eller atferdsproblemer og ikke tatt hensyn til at resultatene varierer i forhold til målgrupper (for eksempel var resultatene dårligere når ungdommene hadde alvorlig psykiatrisk tilleggsproblematikk). Utvelgelseskriteriene medførte at 8 studier ble inkludert, 6 fra USA, én fra Canada og én fra Norge. Littell m.fl. (2005) stilte imidlertid ikke krav til at studiene skal ha vært publisert i tidsskrifter med peer-review, slik det ofte anbefales i retningslinjer for forskning om evidensbaserte metoder (Flay m.fl. 2005) og som dermed gir artikler et faglig kvalitetsstempel. Utvalgsprosedyren medførte at denne meta-analysen i motsetning til andre inkluderer den upubliserte canadiske MST-studien (Lescheid & Cunningham 2002), noe Littell (2005a) også hevder kan ha fått betydning for at hennes konklusjoner avviker fra de som er trukket i andre meta-analyser.

Hva viser meta-analysen av MST?

Mens sju av de åtte primærstudiene konkluderte med at det var signifikante forskjeller i favør av MST på én eller flere resultatvariabler, så fant Littell m.fl. (2005) i sin meta-analyse at selv om dataene hadde en tendens til å favorisere MST, så var ikke forskjellene statistisk signifikante(1). Forklaringen var ifølge Littell m.fl. (2005) at effektene ikke var konsistente på tvers av studiene. I meta-analysen reanalyserer Littell m.fl. resultatene variabel for variabel på tvers av de åtte studiene, til sammen 21 enkeltanalyser. Hver enkelt analyse av sentrale resultatvariabler omfattet mellom to og fem studier. Med andre ord: Ingen av analysene omfatter mer enn fem undersøkelser, og i de fleste tilfellene er det faktisk bare tre, eller to. I dette meget begrensede utvalget har også studiene høy heterogenitet, noe som kan forventes av studier som skiller seg fra hverandre på viktige områder som i valg av undersøkelsesgruppe, observasjonsperiode, implementering og design. Med et så lite antall primærstudier kreves det stor homogenitet i design og resultater for å kunne påvise statistisk signifikante forskjeller(2). Vanligvis inneholder meta-analyser langt flere undersøkelser (gjerne 50–200 studier). Med bare åtte undersøkelser er det vanskelig å avgjøre om MST er signifikant bedre enn tilbudet i sammenligningsgruppen. Littell m.fl. (2005) skriver da også selv at: "Since statistical power is low, we cannot conclude that MST is not more effective than other services" (s.22)(3).

Den canadiske MST-studien (Lescheid & Cunningham, 2002) veier tungt i meta-analysen fordi den hadde mange deltakere (N=409) og utgjør 21–45 prosent av undersøkelsespopulasjonen i meta-analysene den inngår i. Samtidig rapporterer den klart mer negative resultater sammenlignet med de andre evalueringene. Sundell (2005) antar at dersom denne analysen hadde vært ekskludert fra meta-analysen så ville det har vært flere signifikante forskjeller mellom MST og sammenligningsgruppene. Dette bekreftes også av meta-analysen fra Curtis, Ronan og Borduin (2004) som inkluderer 6 av de 8 undersøkelsene som Littell m.fl. (2005) bygger på. Til tross for at det er flere likheter i utvalget av undersøkelser mellom Curtis m.fl. (2004) og Littell m.fl. (2005), så kommer de til svært forskjellige konklusjoner. Curtis m.fl. (2005) fant at MST signifikant reduserer problematferd med effektstørrelse Cohens $d = .87$ i demonstrasjonsundersøkelser (efficacy) og Cohens $d = .26$ i undersøkelser fra vanlig praksis (effectiveness). Den canadiske studien skiller seg også fra andre MST-studier ved at den omfattet ungdom med færre problemer, tiltakene i kontrollgruppen hadde en høyere kvalitet enn i USA og implementeringskvaliteten var lavere (Lescheid & Cunningham 2002; Sundell 2005). I en kommentar til meta-analysen skriver Sundell (2005) at det er uklart om den skal generaliseres til MST som intervensjon, til målgruppen, til konteksten eller graden av programintegritet. Han påpeker at det er viktig å kunne avgjøre om manglende effekter skyldes svakheter ved behandlingen eller svakheter ved implementeringen (f.eks. dårlige forberedelser, eller manglende behandlingsintegritet).

Hvem skal telle med i analysene?

Littell m.fl. (2005) foretar en kvalitetsrangering av MST-studiene, og tillegger 'intention to treat' (ITT) analyser høyest kvalitet. I deres meta-analyse er det bare den canadiske MST-studien som har gjennomført slike analyser. ITT-analyser omfatter alle som tilfredsstillter inklusjonskriteriene og som har vært forespurt om å delta i undersøkelsen. Begrunnelsen for dette er at ungdom som ikke deltar (refusers) eller faller ut (drop-outs) kan ha mer alvorlige problemer enn de som fullfører, og at en dermed kan komme til å overvurdere effekten av et program. En alternativ forklaring kan være at familiene finner ut at problemene ikke er så alvorlige at de har behov for MST. Analyser av 'treatment on the treated' (TOT, av andre omtalt som 'totally treated'), som bare inkluderer deltakere som faktisk har gjennomgått behandlingen, rangeres lavere. Men for den som ønsker å orientere seg om hvor effektiv for eksempel MST er, er det kanskje mest interessant å vite hvordan de som faktisk deltar i programmet kommer ut og da er TOT-analyser mer hensiktsmessige. Problemet med ITT-analyser er at de ofte undervurderer effekten av et program og overser hvordan et slikt design påvirker de såkalte 'never-takers' (de som avslår eller faller ut). 'Never takers' kan lett bli demoraliserte fordi de ikke klarer å gjennomføre behandlingen og kan derfor trekke resultater i negativ retning. Dette er en situasjon som ikke blir aktuell for deltagerne i kontrollgruppen, og de to gruppene blir derfor sammenlignet på ulikt grunnlag. For som Jo (2003) skriver: "This negative psychological effect [of demoralization or discouragement] would not occur for never-takers assigned to the control condition, since the treatment is never offered" (p. 4). Vår oppfatning av at ITT-analyser kan være et dårlig alternativ i RCT-analyser deles av flere statistikk-eksperter (Jo 2003; Jo & Muthen 2003).

Kritikk av primærstudiene i MST

På et mer detaljert nivå påpeker Littell m.fl. (2005) på det de oppfatter som forskningsmetodiske svakheter ved MST studiene som 1) inkonsistent rapportering av antall saker i ulike publikasjoner fra samme prosjekt, 2) inkonsekvent parvis kobling mellom MST og sammenligningssaker, 3) uklare randomiseringsprosedyrer, 4) ustANDARDISerte observasjonsperioder i oppfølgingsstudiene, og 5) subjektiv definisjon av behandlingsavslutning. Littell m.fl. hevder blant annet at selektivt frafall (dropout) i MST-studiene har ført til at resultatene blir mer positive enn dersom dette hadde vært håndtert på andre måter. Gjennomgående, mer enn antyder Littell at avvikene som påpekes er gjennomført for å oppnå bedre behandlingsresultater og at de undergraver det eksperimentelle designet, resultatenes nøyaktighet og dermed undersøkelsenes troverdighet. Kritikken fra Littell m.fl. (2005) tilbakevises av Henggeler m.fl. (2005) som forklarer og begrunner sine prosedyrer. Slik vi ser det, er det vanskelig å etterprøve alle argumentene i denne debatten. Enda vanskeligere synes vi det er å fastslå

hvilken betydning Littells innvendinger har for resultatene. Så vidt vi kan bedømme, er det forskerne som har gjennomført studiene som har den beste kunnskapen om utvalgenes størrelse, og at de har håndtert frafall som i en hvilken som helst randomisert evalueringstudie. Det er også et åpent spørsmål om endringene i forskningsprosedyrer ville hatt noen betydning for behandlingseffekten. Littell er da også mindre opptatt av hvilke konsekvenser kritikken har for de praktiske behandlingsresultatene enn av avvikene fra det hun oppfatter som korrekte prosedyrer for gjennomføring og rapportering av behandlingseffektstudier.

Forhold som modererer effekten av MST

Uenigheten mellom Henggeler m.fl. (2005) og Littell m.fl. (2005) dreier seg blant annet om hvor mye en skal ta hensyn til forhold som virker differensierende eller modererende på resultatene. Henggeler m.fl. (2005) mener det er viktig å skille mellom ulike faser av programutviklingen når en skal vurdere behandlingseffektivitet, og at en må ta hensyn til variasjoner i behandlingsintegritet mellom utprøvningssteder.

Manglende differensiering mellom forskningsbaserte og praksisbaserte studier

Behandlingseffektstudier kan beskrives som ”efficacy”- eller ”effectiveness”-studier, der efficacy-studier illustrerer hva et program kan gi av resultater under optimale betingelser (vanligvis med viktige bidrag fra programutvikleren), mens effectiveness-studier evaluerer hvordan program virker i ordinær praksis (der programutvikleren har en perifer rolle). Henggeler (2004) kommenterer at det fremdeles gjenstår mye før en vet hvordan man skal implementere forskningsbaserte metoder i vanlig praksis og oppnå tilnærmedesvis så gode resultater som i programutviklerens demonstrasjonsstudier. Littell m.fl. (2005) skiller imidlertid ikke mellom ”efficacy”- og ”effectiveness”-studier i sin meta-analyse, - og mener at skillet mellom disse er uklart i MST-forskningen. Derfor vil ethvert forsøk på å forklare variasjoner ut fra en slik inndeling etter deres mening, være spekulative. Petrosino og Hoydan (2005) fant at når evaluatøren, og i enda større grad programutvikleren, hadde stor innflytelse på design og implementering av behandlingen, så ble det konsistent rapportert om høyere effektstørrelser enn i andre evalueringer. Lipsey (1995) stiller spørsmål om hvordan en skal fortolke dette resultatet: “A cynical view might attribute this (meaning: better results in cases where the researcher determines design and implementation) to some biasing or ‘wish fulfilling’ influence researchers have on the outcomes of studies they control. I see another interpretation as plausible ... When a researcher is closely involved in treatment design ... there is likely to be a high level of treatment integrity...” (s. 75).

En kan velge å fortolke det som om programutvikleren påvirker resultatene gjennom sine forventninger, eller at forskjellene skyldes ekstra innsats for å sikre høy grad

av program- og behandlingsintegritet. Etter vår oppfatning har Littell og hennes kolleger (2005) valgt den 'kyniske forklaringen' og sett bort fra at forskjellene kan dreie seg om høyere implementeringskvalitet (Wilson 2004).

Betydningen av behandlingsintegritet

Behandlingsintegritet i MST målt med "treatment adherence measure scores" (TAMS, Henggeler & Borduin 1992) har i flere undersøkelser vist seg å samvariere med behandlingsresultatene slik at høye skårer har vært forbundet med mer positivt behandlingsutfall (Henggeler 2004). Littell m.fl. (2005) mener imidlertid at TAM ikke diskriminerer mellom MST og andre intervensjoner og at det som måles ikke er unikt for MST, slik som for eksempel engasjement, behandlingsdeltakelse og terapeutisk allianse. For oss virker det som Littell m.fl. kritiserer TAM-skalaen for manglende diskriminerende egenskaper ut fra antagelser og uten empirisk underlag. Etter vår oppfatning består alle program av en kombinasjon av generelle komponenter av typen som Littell (2005b) refererer til og unike programkomponenter. MST forutsetter for eksempel at terapeutene behersker generelle prosessferdigheter som er felles for de fleste behandlingsprogram, i kombinasjon med mer spesifikke MST-komponenter som for eksempel arbeid for økt kontakt mellom familie og omgivelser. Vi synes derfor det er rimelig at TAM-skalaen inneholder en kombinasjon av disse felles og unike elementene, – og mener at den er et godt mål på behandlingsintegritet i MST.

Forskjeller i behandlingseffekt mellom utprøvssteder (site-differences)

Littell m.fl. (2005) hevder at data ikke støtter hypotesen om at MST er mer effektivt på noen behandlingssteder enn på andre. De viser til at i to av tre MST-undersøkelser med flere utprøvssteder, rapporteres ikke resultater fra det enkelte utprøvssted. De mener også at det er galt å påberope forskjeller mellom behandlingssteder i ettertid (post hoc) og etterlyser flernivåstudier ("multilevel studies") som kan redegjøre for forskjeller i behandlingseffekt mellom terapeuter og behandlingssteder i den samme undersøkelsen (såkalt nesting effects). Dette kravet virker imidlertid noe urealistisk da multi-nivå-studier forutsetter et høyt forholdstall mellom de ulike nivåene (dvs. at man innenfor studien måtte hatt 20–30 utprøvssteder med flere terapeuter i hver). Littell m.fl. (2005) hevder også at resultater fra det enkelte utprøvssted ("site effects") er uinteressante i en meta-analyse som har som formål å aggregere resultatene fra all tilgjengelig forskning. Det virker motsigende for oss at Littell både etterlyser mer avanserte analyser av variasjoner mellom utprøvssteder, samtidig som hun avviser at det er relevant i meta-analyser.

Effekter av hvor lenge programmet hadde vært i bruk ("evaluating programs before they mature")

Resultatene fra MST-studier viser ifølge Henggeler m.fl. (2005) at det første året er særlig krevende for mange MST-program når de implementeres på nye steder og uten programutviklerens direkte medvirkning. Evalueringer av program rett etter oppstart kan dermed føre til at en undervurderer programmets potensiale. Tilsvaret fra Littell (2005) kretser mest rundt den canadiske studien, der hun, i motsetning til Henggeler m.fl., hevder at data ikke gir grunnlag for å konkludere med at programmets "modenhet" påvirker resultatene i positiv retning.

Oppsummert hevder Henggeler og kolleger (2005) at det er viktig å differensiere mellom evaluering i ulike faser av programutviklingen. Best resultater oppnås i studier der det er høy grad av kontroll med behandlingsintegritet, høy grad av programintegritet på utprøvningsstedet, og der en har en viss erfaring med bruk av programmet. Littell på sin side avviser at MST-forskningen dokumenterer dette. Mye av diskusjonen og uenigheten mellom Henggeler og Littell dreier seg videre om hvilken betydning en skal tillegge undersøkelsen fra Canada. Henggeler m.fl. (2005) legger vekt på at undersøkelsen ikke er publisert, og at resultatene ikke er så ensidig negative som Littell hevder. Littell (2005) på sin side mener at den canadiske studien er den første store, uavhengige replikasjonsstudien av MST, og at den derfor ikke har samme grad av 'bias' som undersøkelsene gjennomført av Henggeler og hans kolleger. Det er også den eneste evalueringen som er gjennomført etter prinsippet om "intention to treat". Hun kommenterer imidlertid ikke at denne studien vektet uforholdsmessig mye i meta-analysene i kraft av utvalgsstørrelsen (N=409). Ingen av de øvrige MST-studiene har over 200 deltakere.

Oppsummering

Ifølge Littell m.fl., (2005) var hovedhensikten med den systematiske forskningsoversikten å sette fokus på det store spriket som kan være mellom forskning og praksis i arbeidet med meta-analyser. Videre var hensikten å argumentere for økt grundighet i gjennomføring, rapportering og syntetisering av kontrollerte evalueringsstudier.

Oppsummert mener vi at Littell og kolleger stiller urimelige høye standardiseringskrav til design, variabler og prosedyrer i undersøkelser som skal inngå i en meta-analyse. For eksempel vil design, deltakerkjennetegn og utvalg av resultatvariabler i den enkelte studie styres av skiftende behov, interesser og muligheter. Videre mener vi at meta-analysen omfattet for få undersøkelser. Dette kunne ha vært kompensert for ved for eksempel å inkludere undersøkelser av andre program som retter seg mot den samme målgruppen. Slik kunne en ha økt den statistiske styrken i analysene, – og satt et bredere fokus på forskningen om familiebaserte behandlingsprogram. Med så få

undersøkelser, og den noe spesielle måten å analysere dataene på, fører det til at konklusjonen framstår som unyansert og dårlig samsvarende med primærstudiene som analysen bygger på.

Littell m.fl. (2005) hevder det er prematurt å konkludere om hvor effektiv MST er. Med så få og heterogene studier i meta-analysen mener vi, med støtte i rådene fra The Cochrane Collaboration, – at det kanskje er riktigere å si at det var prematurt å gjøre en meta-analyse av MST.

Littell og kolleger (2005) gir da også uttrykk for at det trengs flere, uavhengige studier for å bekrefte eller avkrefte om MST er signifikant mer effektiv enn andre tiltak. Vi tror det er viktig å skille mellom forskningskritikk og kritikk av MST som behandlingsmetode, og mener kritikken av forskningspublikasjonenes kvalitet ikke uten videre gir grunnlag for å si noe om behandlingsprogrammets effektivitet. Selv om en tar Littell og kollegers (2005) kritikk for pålydende, så er det fremdeles uklart om de påpeker svakheter ved forskningen eller ved implementeringen. Vi synes også det er en logisk brist i enkelte resonnementer. For eksempel skriver Littell og kolleger (2005) at studiene i meta-analysen har utilfredsstillende metoder for innhenting av data, uegnet kontroll på randomisering og behandlingsperiode, i tillegg til svakheter ved dataanalysene. Men hvis MST-forskningen faktisk er slik som Littell og kolleger (2005) beskriver, – undergraver de samtidig sin egen meta-analyse: En meta-analyse blir aldri bedre enn de data, metoder, og studier den bygger på. Når effektene av behandlingen antas å variere for mye fra studie til studie hevder The Cochrane Collaboration at det er uhen-siktsmessig å kombinere effektene av ulike studier (The Cochrane Collaboration).

Selv om kritikken av MST primært gjelder forskningsresultatene, så skinner det igjennom at Littell m.fl. også er kritiske til selve behandlingsprogrammet. I sin drøfting av implikasjoner for praksis gir de uttrykk for at det er viktig å erkjenne grensene for hvilke resultater som kan oppnås med korttids-, individ- og familiefokuserte intervensjoner, uansett hvor godt design og gode intensjoner de bygger på. Videre skriver de at mer robuste, langsiktige intervensjoner, kombinert med konsistent økonomisk, utdanningsmessig, medisinsk og terapeutisk støtte for ungdom og familier, kan være nødvendig for å oppnå varige forbedringer i ungdommens og familiens fungering. Disse synspunktene er åpenbart ikke basert på de empiriske resultatene fra den systematiske forskningsoversikten, og gir kanskje et innblikk i forfatterens faglige grunnsyn. En kan undre seg om disse synspunktene har vært en premiss for forskningsoversikten eller en konklusjon. Kommentarene viser at Littell m.fl., relativt uavhengig av forskningsdataene, stiller seg tvilende til MST. Vi finner ikke empirisk dekning for disse synspunktene i den systematiske forskningsoversikten, - og tviler på om denne resepten er egnet for norske og nordiske forhold. Dette fordi støtten de etterlyser i mange år har vært viktige elementer i det praktisk kliniske arbeidet med alvorlige atferdsproblemer i Norden, uten at det ser ut til å ha gitt ønsket resultat (Storvoll 1997). Og det er kanskje derfor at en i Norden har etterlyst et behandlingstilbud organisert etter de faglige prinsippene

som MST representerer (forskningsbasert, multisystemisk, hjemmebaserte tjenester, brukermedvirkning osv.)

Forskningsoversikten kommer på et tidspunkt hvor diskusjonen i Norden går høyt om verdien av evidensbaserte metoder generelt og MST spesielt. Samtidig er det etablert et nordisk Campbell-senter som ønsker å presentere forskningsoversikter over aktuelle behandlingsprogram og metoder. I denne sammenhengen må forskningsoversikten over MST fortone seg forvirrende. De fleste har fått med seg at MST har vist gode resultater i kontrollerte evalueringstudier, og forventer nok at forskningsoversikten i noen grad skal gjenspeile det. Imidlertid presenterer det nordiske Campbell-senteret forskningsoversikten under overskriften ”Tvil om effekten av MST”, og informerer uten forbehold om at metoden er verken bedre eller andre enn andre metoder (www.nc2.net). Noe mer tilbakeholdne er da Littell og kolleger (2005) som til tross for sine negativt avventende konklusjoner påpeker at det ikke er evidens for at noen kjent intervensjon er mer effektiv enn MST. Basert på de norske evalueringresultatene (Ogden & Halliday-Boykins 2004; Ogden & Amlund Hagen 2006) vil vi hevde at MST virker tilfredsstillende i norsk sammenheng, – sammenlignet med andre barnevernstiltak for ungdom med alvorlige atferdsproblemer dersom programmets kjernekomponenter implementeres i tråd med retningslinjene fra programutvikleren.

Endnotes

- ¹ p-verdien på resultatvariabelen 'rearrestasjoner' var: 0.0504 noe mange ville ha bedømt som en signifikant forskjell i favør av MST på 5% nivået.
- ² Ifølge Lipsey (2000) har testen (Q-test) som brukes i slike analyser svak 'power' og en trenger minst 20 separate studier for å kunne si noe om gjennomsnittresultater.
- ³ Lav 'statistisk power' øker risikoen for såkalte type II feil, det vil si at virkelige forskjeller ikke oppdages med statistisk hypoteseprøving (Sundell 2005).

*) En systematisk review inneholder klare kriterier for inkludering og ekskludering av undersøkelser, en eksplisitt søkestrategi, en systematisk koding og analyse av inkluderte undersøkelser og en meta-analyse (hvis mulig) for å beregne den gjennomsnittlige effektstørrelsen (pooled effect size ES). En meta-analyse er en forskningsmetode der en oppsummerer resultatene fra to eller flere primærstudier for å beregne den gjennomsnittlige effekten av disse (også omtalt som en analyse av analysene).

Referanser

- Borduin, Charles M., Barton J. Mann, Lynn T. Cone, Scott W. Henggeler, Bethany R. Fucci, David M. Blasko & Robert A. Williams (1995) Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, 63, 569-578.
- Borduin, Charles M. & Scott W. Henggeler (1990) A multisystemic approach to the treatment of serious delinquent behavior. In: McMahon, Robert J. & Peters, Ray D. (eds.) *Behavior disorders of adolescence: Research, intervention, and policy in clinical and school settings*. New York, Plenum Press.
- Curtis, Nicola M., Kevin R. Ronan, & Charles M. Borduin (2004) Multisystemic treatment: A meta-analysis of outcome studies. *Journal of Family Psychology*, 18, 411-419.
- Fly, Brian. R., Anthony Biglan, Robert F. Boruch, Felipe G. Castro, Denise Gottfredson, Sheppard Kellam, Eve K. Moscicki, Steven Schinke, Jeffrey C. Valentine, & Peter Ji (2005) Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6, 151-171.
- Henggeler, S.W., Schoenwald, S.K., Borduin, C.M., Rowland, M.D., Cunningham, P.B. (1998) *Multisystemic Treatment of Antisocial Behavior in Children and Adolescents*. New York: The Guilford Press.
- Henggeler, S.W., Borduin, C.M. (1992) *Multisystemic therapy adherence scales*. Unpublished instrument. Department of Psychiatry and Behavioral Science, Medical University of South Carolina.
- Jo, Booil (2003) Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics*, 27, 385-409.
- Jo, Booil & Muthen, Bengt O. (2003) Longitudinal studies with intervention and noncompliance: Estimation of causal effects in growth mixture modeling. In: S. Reise & N. Duan (Eds) *Multilevel modeling: Methodological advances, issues, and applications*. (pp. 112-139). Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Leschied, Allan & Cunningham, Allison. (2002) Seeking effective interventions for serious young offenders: Interim results of a four-year randomised study of Multisystemic Therapy in Ontario, Canada. Centre for Children and Families in the Justice System, London, Canada.
- Lipsey, Mark (2000) *Practical meta-analysis*. Seminar ved CUS, Sosialstyrelsen, September 2000, ref. fra Sundell, K. (2005).
- Lipsey, Mark (1995) Her sitert fra Antonio Petrosino (2004). Presentation prepared for Social methods seminar, 9 december, 2004, Centre for evaluation of social services, National Board of health and welfare, Stockholm.
- Littell, Julia H. (2005a) Lessons learned from a systematic review of multisystemic therapy. *Children and youth services review*, 27, 445-463.
- Littell, Julia H. (2005b) The case for Multisystemic Therapy: Evidence or orthodoxy? *Children and youth services review*. www.elsevier.com/locate/childyouth
- Littell, Julia H., Melania Popa, & Burneé Forsythe (2005) *Multisystemic therapy for social, emotional and behavioural problems in youth aged 10-17*. Rapport 05:01, Nordic Campbell Center www.nc2.net
- Ogden, Terje & Amlund-Hagen, Kristine (2006) Multisystemic Treatment of Serious Behaviour Problems in Youth: Sustainability of Effectiveness Two Years after Intake, *Child and adolescent mental health*, in print.
- Ogden, Terje & Halliday-Boykins, Colleen, A. (2004) Multisystemic treatment of antisocial adolescents in Norway: Replication of clinical outcomes outside of the US. *Child and adolescent mental health*, 9, 76-82.
- Petrosino, Antonio & Haluk Soydan (2005) The impact of program developers as evaluators of criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of experimental criminology*, 1, 435-450.
- Storvoll, Elisabeth (1997) *Barn og unge med alvorlige atferdsvansker: Hvem er de, og hvilken hjelp blir de tilbudt?* Rapport 21. Oslo: NOVA.
- Sundell, Knut (2005) *Fackgranskning av Systematic and nonsystematic reviews of effects of Multisystemic Therapy (MST)* av Julia Littell. Rapport til Institutt for Metoder i Sosialt arbeid (IMS), Sosialstyrelsen, Stockholm.
- The Cochrane Collaboration (2002) The Cochrane Collaboration Open Learning Material. Retrieved from <http://www.cochrane-net>.

org/openlearning/HTML
 Wilson, Sandra. J. (2004) *Do program developers influence program outcomes? An example from meta-analysis*. Social methods seminar, 9 december, 2004, Centre for evaluation of social services, National Board of health and welfare, Stockholm.

Summaries

We comment on a systematic review and meta-analysis of Multisystemic Therapy (MST) presented by Littell, Popa, and Forsythe (2005) for the Campbell Collaboration. The authors concluded, based on 8 evaluations of MST, that it is premature to determine the effectiveness of MST as compared to other services. In this article, we argue that the foundation on which this review is based is insufficient and assert that it is the meta-analysis that is premature. In consequence, the review seems unfit for evaluating the effectiveness of MST on adolescent behavioral problems, particularly for the Nordic countries.

Virkar MST? Athugasemdir við kerfisbundið rannsóknaryfirlit og meta-analysis af MST
 Í greininni gerum við athugasemdir við hið kerfisbundna rannsóknaryfirlit og meta-analysis (eftirgreiningu) af Multisystemic Therapy (MST) kynnt af Littell, Popa og Forsythe (2005) fyrir Campell Collaboration (Nordic Campell Center).

Rannsóknin hefur vakið athygli og þar með lagt sitt af mörkum til endurvakinnar umræðu um notkun MST og aðferða sem byggja á sannprófun. Á grundvelli mats á átta MST prófunum komust höfundar að þeirri niðurstöðu, að það sé ekki tímabært að ákvarða hversu árangursríkt MST sé samanborið við aðra meðferð/þjónustu. Þó eru ekki merki um að MST sé verra en önnur meðferð. Í þessari grein færur við rök fyrir því að rannsóknaryfirlitið sé ekki reist á nægilega traustum grundvelli og staðhæfum að það sé eftirgreiningin (meta-analysis) sem sé ótímabær. Þar af leiðandi virðist rannsóknaryfirlitið ekki hæft sem gildismat á árangri MST í starfi við alvarlegar hegðunarraskanir unglinga, sérstaklega á Norðurlöndunum.

Vaikuttaako MST? Kommentteja MST:tä koskevalle systemaattiselle tutkimuskatsaukselle ja meta-analysille

Tässä artikkelissa kommentoimme multisysteemisen terapian (MST) systemaattista tutkimuskatsausta ja meta-analyysiä, jonka Littell, Popa ja Forsythe (2005) toteuttivat Cambell Collaboratorin –yhteistyöverkostolle. Katsauksen kahdeksaan MST-tutkimukseen pohjautuva johtopäätös on, että on ennen aikaista määritellä MST:n vaikuttavuus suhteessa muihin palveluihin. Tässä artikkelissa esitämme, että katsauksen johtopäätösten perusta on puutteellinen ja väitämme, että se on meta-analyysi, joka on ennen aikainen. Tästä syystä katsaus ei tunnu soveltuvan arvioimaan MST:n vaikuttavuutta nuorison käyttäytymisongelmiin, etenkin Pohjoismaissa.