

Metaanalyser som grundlag for evidensbaseret misbrugsbehandling

En diskussion af fem metaanalyser af behandlingseffekter

Indledning

Metaanalyser er opgørelser, hvori forskellige undersøgelses resultater lægges sammen, med det formål at øge datamængden og dermed øge sikkerheden for validiteten af en bestemt konklusion. Ideelt set kan man se metaanalysen som et forsøg på i praksis at tage konsekvensen af idealet om, at videnskabelige studier skal være replicerbare. I metaanalysen forsøger man at kontrollere, om replikationerne har vist, at konklusioner fra tidlige lovende studier om en bestemt sammenhæng var holdbare.

I denne artikel gennemgås en række nyere metaanalyser med det formål at identificere faktorer, der bidrager til at øge eller mindske værdien af metaanalyser.

En serie af metaanalyser udkommer i denne tid med fokus på afhængighed af rusmidler. Wilson gennemgik for et par år siden litteraturen og præsenterede metodologiske styrker og svagheder i en række metaanalyser (2000).

Wilson fremhævede blandt andet nogle meget vigtige aspekter af de statistiske metoder i metaanalyser, herunder valg af regressionsanalytisk metode, analyse af varians og så videre.

Siden Wilsons gennemgang er nye metaanalyser af effekter af misbrugsbehandling blandt andet blevet præsenteret af Prendergast og kolleger (2002), SBU-rapporten (2001), Griffith og kolleger (2000) og Irvin og kolleger (1999). Endvidere har Pearson og Lipton præsenteret en metaanalyse af behandling i forbindelse med straf (1999).

Metaanalyser kan bidrage til at identificere dokumenteret effektive behandlinger, og de kan i den sidste ende bidrage til, at behandlingen som sådan bliver evidensbaseret.

Cochrane Collaboration, et internationalt netværk, der arbejder for indførelse af evidensbaseret medicin, anvender metaanalyser som standard for reviews over behandlinger med følgende begrundelse: "*Systematic reviews [her: metaanalyser, M.H.] establish where the effects of healthcare are consistent and research*

Denne artikel er en del af forskningsaktiviteterne ved Center for Rusmiddelforskning, Aarhus Universitet. Jeg takker forskningsassistent Helle Balsby for nyttige kommentarer til manuskriptet.

results can be applied across populations, settings, and differences in treatment (e.g. dose); and where effects may vary significantly. The use of explicit, systematic methods in reviews limits bias (systematic errors) and reduces chance effects, thus providing more reliable results upon which to draw conclusions and make decisions." (Clarke & Oxman 2002, sektion 1).

Den enkelte undersøgelse kan altså give en foreløbig støtte til en bestemt konklusion, men med metaanalyser inddrages replikationsstudier, sådan at man får et samlet overblik over den eksisterende litteratur i forhold til et bestemt spørgsmål.

Cochrane Collaboration anbefaler en meget systematisk og forholdsvis snæver definition af spørgsmålet for undersøgelsen. *"As with any research, the first and most important decision in preparing a review is to determine its focus This is best done by asking clearly framed questions. Such questions are essential for determining the structure of a review Specifically, they will guide much of the review process including strategies for locating and selecting studies or data, for critically appraising their relevance and validity, and for analysing variation among their results."* (Clarke & Oxman 2002, afsnit 4.1).

Laver man en metaanalyse med visse systematiske forskelle i undersøgelsesdesign eller målgrupper, har man nogle statistiske redskaber til at analysere eventuelle modstridende resultater, som ikke umiddelbart er tilgængelige ved et narrativt review, hvor litteraturen læses igennem. Der findes veldokumenterede og vidt udbredte metoder til at udføre statistiske kontroller, der tager udgangspunkt i en vægtning af den enkelte undersøgelse ud fra usikkerheden i undersøgelsen, og man kan lave beregninger, som eksempelvis den såkaldte Q-statistik, der undersøger, om de enkelte undersøgelser varierer mere i effekt, end man ville forvente at finde, hvis de var tilfældige udpluk af den samme undersøgelse. Endvidere kan man statistisk inddrage sådan-

ne faktorer som for eksempel kvalitetskontrol og diagnostisk fremgangsmåde i en regressionsanalyse med henblik på at finde de vægtede sammenhænge mellem metodologiske variable og effekten af behandlingen.

En stor trussel mod kvaliteten af metaanalyser, kendt under navnet "garbage-in, garbage-out", handler om, at metaanalyser aldrig bliver bedre end de enkeltundersøgelser, som de består af. Hvis man analyserer effektstørrelsen på en stor mængde godt og dårligt gennemførte undersøgelser, så risikerer man at få en stor, dårlig analyse. Den eneste rimelige måde at håndtere dette på er at lave en relevant kvalitetskontrol af undersøgelseerne, inden man gennemfører metaanalysen. Herved kan man ekskludere ubrugelige eller dårligt udførte undersøgelser på forhånd, og derved undgår man at komme frem til de forkerte konklusioner. Typisk vil en sådan kvalitetskontrol blive udført ved, at uafhængige ratere gennemlæser de samme artikler og giver artiklerne point ud fra sammenlignelige kriterier. Et eksempel på sådanne kriterier findes i SBU-rapporten (2001, Appendix).

Endnu en trussel mod kvaliteten af metaanalyser er bias. Mest omtalt er publikationsbias eller "skrivebordsskuffefænomenet", hvor negative resultater bliver liggende i skrivebordsskuffen, mens positive resultater bliver publiceret. Men denne bias kan også tage form af citationsbias, hvor flere databaser, artikler og bøger refererer til den positive undersøgelse, der derfor bliver citeret flere steder og dermed bliver lettere at identificere for den arbejdsgruppe, der laver metaanalysen. Endvidere er der rapporteringsbias, hvor man risikerer at kun data, der er "interessante", altså statistisk signifikante, rapporteres i artiklen (Clarke & Oxman 2002, sektion 8.11.1). Bias kan undersøges med visse statistiske metoder, blandt andet "funnel plot" eller skorstensgraf, hvor man undersøger, om der er tendens til, at de mindste og mest usikre undersøgelser viser de højeste effekter, mens de

største og mest sikre undersøgelser viser de mindste effekter. En anden metode er at beregne antallet af negative undersøgelser, der kunne dukke op af skrivebordsskufferne og vælte konklusionen på metaanalysen. Denne metode kaldes "the failsafe N". Men før man kan vælge eller have nytte af en statistisk metode, er det her afgørende at forstå, hvad bias er. Bias er altså alle de faktorer, der kunne forvrænge det billede, som i den sidste ende kommer ud af metaanalysen.

Metode

Dette er en gennemgang af fem forskellige nyere metaanalyser af psykosocial behandling for misbrug, som hver for sig repræsenterer metodiske tilgange til metaanalysen. De adskiller sig især ved at have taget udgangspunkt i forskellige forskningsspørgsmål, men har dog det til fælles, at de blandt andet søger at lave en syntese af litteraturen ud fra et bestemt spørgsmål. Derved adskiller de sig fra mere analytiske reviews, der eksempelvis forsøger at besvare spørgsmål af typen "hvilke faktorer bidrager til X". Et eksempel på de mere analytiske reviews er Brewer og kollegers metaanalyse af faktorer, der forudsiger fortsat misbrug ved opiatmisbrug (1998).

Til forskel fra Wilsons gennemgang fokuseres her mere på de spørgsmål, som metaanalyserne synes at kunne belyse, frem for på de statistiske metoder i analysen. De statistiske metoder vil kun blive berørt i den udstrækning, hvor de direkte har betydning for vurderingen af metaanalysernes evne til at producere klinisk og forskningsmæssigt betydningsfulde konklusioner.

Reviewet tager udgangspunkt i to sider af metaanalysens mange mulige formål:

Den kliniske relevans af den type af fund, som kan gøres igennem de pågældende metaanalyser. Altså: hvilke implikationer har den pågældende metaanalyse for kliniske beslut-

ninger, herunder beslutninger på institutionsniveau, eller med hensyn til tilrettelæggelsen af behandlingssystemer.

Den teoretiske relevans af den type af fund, som gøres igennem disse metaanalyser. Det vil sige: på hvilken måde kan metaanalyser bidrage til at falsificere hypoteser, som følger af specifikke videnskabelige teorier, eller til at nye hypoteser kan rejses, eller til at nye teoretiske modeller lader sig formulere i lyset af metaanalysernes resultater.

Prendergast et al. 2002

Prendergast og kollegers metaanalyse er en gennemgang af alle de evalueringer, kliniske undersøgelser og kvasiekperimentelle studier, som forskergruppen har kunnet identificere, og som er foretaget imellem 1965 og 1996 i USA eller Canada. Studierne skulle være sammenligninger af behandling med ingen behandling eller en eller anden form for opmærksomhedsplacebo. Analysen er særdeles grundig på den statistiske side med en omfattende scoringsnøgle til kodning af undersøgelser tilgængelig fra udgiverens hjemmeside. Desuden er der en grundigt beskrevet fremgangsmåde for udregning af effektstørrelser, vægtning, scoring af metodologiske variable, populationsbeskrivelser og den regressionsanalyse, der skal besvare undersøgelsens spørgsmål. Wilson nævner da også den tidligere metaanalyse fra denne gruppe som en af de stærkeste og mest vel gennemførte på misbrugsområdet i sit review (2000). Endvidere forsøger gruppen gennem statistiske kontroller at undersøge publikationsbias. Der er tale om kontroller, der måske, som jeg skal komme ind på nedenfor, ikke kan besvare spørgsmålet i dette tilfælde. Endelig er der gjort meget for at opspore eventuelle skjulte undersøgelser.

De metodologiske variable, der søges taget højde for, ligger i forlængelse af den eksiste-

rende litteratur på det psykosociale område. Så vidt undersøgelsens stærke sider.

Det lykkedes for gruppen at identificere 78 undersøgelser af forskellig kvalitet.

Prendergast og kolleger stiller det altomfattende spørgsmål, om behandling virker. De konkluderer, at svaret er ja. At behandling virker vil sige: *"With these caveats in mind, the findings from this meta-analysis indicate that drug abuse treatment, as it is practiced in the United States, is effective in reducing drug use and crime. Overall, people with drug abuse problems are better off being in treatment than not."* (2002).

Men ligger denne konklusion inden for, hvad Prendergast og kolleger kan konkludere med deres data i hånden? Ja og nej. Analysen lider af mangler, der gør konklusionen tvivlsom. Disse mangler kommer til udtryk på forskellige måder i rapporten.

Den største og alvorligste mangel er det totale fravær af en definition af behandling. Det er på sin vis en logisk følge af, at undersøgelsens spørgsmål er, om behandling virker. Det underliggende begreb er så bredt, at det bliver umuligt at tænke sig egentlige operationaliseringer.

Den manglende definition af behandling gør sammentællingen ganske svær at forstå, endside at sætte den i relation til en klinisk eller en politisk virkelighed. For nu at bruge en aldeles uvidenskabelig metode kan vi opgøre de første fem behandlinger i litteraturlisten i den rækkefølge, hvori de optræder (de fem der udmærker sig ved at have en første forfatter, hvis efternavn starter med bogstavet A): kortvarig metadonafgiftning, psykoterapi ved siden af metadonbehandling, akupunktur for kokainafhængighed og adfærdsterapi. Længere nede finder vi urinprøvekontrol, softball, døgnbehandling, contingency management, strukturel familierapi, transcendent meditation og så videre. Sammenlægger vi effekten af disse, hvilke konklusioner kan en positiv og signifikant effektstørrelse så give os om be-

handling af afhængighed af rusmidler? At vi kan gøre fuldstændig hvad vi vil, bare vi gør noget? Det er jo drontens berømte kendelse fra Alice in Wonderland, at "alle har vundet, og alle skal have præmier" (Luborsky et al. 2002). Og det er det, dronten siger, efter at alle har løbet rundt imellem hinanden i en stor forvirring (Carrol, 2000). Men der var altså også undersøgelser, der viste effektstørrelser på 0 eller lavere, hvor den behandlede gruppe altså klarede sig dårligere end den ikke-behandlede. De skal vel ikke have præmier?

Denne mangfoldighed af behandlingsformer svækker mange af konklusionerne (jf Beutler, 2002).

Problemet er, at den uafhængige variabel ikke engang er minimalt defineret. Det underliggende begreb, behandling, er alt for heterogent.

Gruppen har truffet det valg at prioritere sammenligninger med minimal behandling eller behandling, hvis formål er at være opmærksomhedsplacebo. Alligevel indgår der undersøgelser med "routine treatment" og 11 undersøgelser med "alternative treatment". Artiklen forklarer ikke, hvorfor disse undersøgelser er blevet inkluderet, mens andre undersøgelser med andre typer af "alternative treatment" ikke er det.

Et yderligere eksempel på en alvorlig metodologisk svaghed er, at frafald i behandlingen regnes som en metodologisk variabel på linie med ikke at have sammenlignelige grupper i kontrol- og eksperimentalgruppen, eller på linie med at have dårlig implementering af behandlingen. Men frafald i behandlingen er absolut ikke en metodologisk variabel, ligesom terapeutisk alliance eller "compliance" med behandlingen heller ikke er det. Frafald i behandlingen er en medierende variabel mellem behandlingen og outcome, fordi den behandling, der fastholder klienten i behandling, også har en større chance for at påvirke ham, end den behandling, som klienten tid-

ligt forlader. Og blandt de faktorer, som kan få en klient til at forlade en behandling er, at han ikke føler, at han får noget ud af den.

Et andet eksempel fra de metodologiske variable som nævnes i rapporten er, at sammenligninger mellem grupper, der ikke var sammenlignelige som udgangspunkt, indgår i metaanalysen. Selv om sådanne observationsstudier kan have en vis hypotesedannende funktion, er de næppe egnede til en metaanalytisk tilgang. De indfører en usikkerhed i analysen, som ikke kan fjernes blot ved at kontrollere statistisk for denne variabel i en regressionsanalyse. Som mindstemål burde man kunne forvente, at der blev givet separate opgørelser af effektstørrelserne for de forskellige typer af sammenligninger. Metaanalysens styrke kommer fra dens store statistiske power, men dens svaghed fra dens sammensatte materiale. Her forsøger Prendergast og kolleger som nævnt at drage en konklusion om kausalitet og klinisk relevans ("*clients are better off...*") ud fra en blanding af materiale, der næppe tillader dette, og materiale der formodentlig tillader en sådan konklusion (eksempelvis kliniske studier). Her havde det nok været mere relevant at skærpe eksklusionskriterierne en smule.

Endvidere rapporteres den overordnede effektstørrelse uden nogen smålig skelen til, hvad den måtte betyde i forskellige undersøgelser. En vægtet effektstørrelse på 0.33 siger, som forfatterne meget rigtigt påpeger, ikke meget om, hvad det er for en forskel, effektstørrelsen udgøres af. Men forsøgene på at omtolke effektstørrelsen til procenter i behandlingsgrupper og kontrolgrupper, der profiterer af behandlingen, hjælper kun lidt på forståelsen. For der er afgørende forskel på at mindske et sidemisbrug hos tunge stofmisbrugere i metadonbehandling med sideløbende kokainafhængighed, at forebygge tilbagefald efter afgiftning og døgnbehandling og at lave familiebehandling med et ungt menneske med et problematisk hashforbrug.

Rapportens forsøg på at tage højde for publikationsbias undergraves af den store sammensathed i materialet. Publikations- og rapporteringsbias kan udgøre et seriøst problem for tiltag i den mere kuriøse ende (såsom softball-behandling), men de udgør et mindre problem i forhold til behandlingsformer, der har en større opbakning i det videnskabelige miljø, hvor det er sandsynligt, at også negative resultater vil blive rapporteret, og at undersøgelserne vil blive publiceret i tidsskrifter. Derfor siger effektstørrelsernes normalfordeling heller ikke meget om publikationsbias eller rapporteringsbias.

Endelig siger en metaanalyse af kliniske undersøgelser, selv om den måtte være langt bedre udført end denne, faktisk ikke noget om, hvorvidt det er bedre for mennesker med misbrugsproblemer at være i behandling end ikke at være det. Den siger noget om, at det er bedre at være i behandling end at være på en venteliste, i en gruppe, der får minimal opmærksomhed, eller i en gruppe, der får opmærksomhedsplacebo. Selve det nederlag det er at overgive sig til behandling frem for at forsøge at løse sine problemer selv, kan tænkes at være grundlag for, at man mindsker troen på sig selv og er dårligere stillet, end man ville være uden behandling. Eller omvendt kan selve det at beslutte sig for at søge behandling i virkeligheden være et afgørende skridt i retning af at få det bedre. Når man først har kastet håndklædet i ringen og er gået i behandling, og man så bliver henvist til en "minimal kontakt"-gruppe, er det muligt, at nogle typer af klienter bliver skuffede og opgiver håbet, i det mindste for en tid. Dette problem er helt parallelt til problemet med patienter i kliniske undersøgelser af eksempelvis antidepressiv medicin, der får en kalktablett og som opdager, at de er på placebopillen, fordi de ikke mærker noget til bivirkninger. Også her kan det være svært at sige, om manglende bedring skyldes skuffelsen over ikke at have fået hjælp til sin depression, eller om den manglende

bedring er udtryk for sygdommens naturlige, ubehandlede forløb.

Prendergast har valgt at ekskludere dosisstudier og studier, der sammenligner typer af behandling med mere eller mindre bona fide. Det kan man selvfølgelig have sine grunde til: en er, at disse undersøgelser kan tænkes ikke at have så stor effekt, hvis drontens kendelse står til troende.

Men ikke desto mindre er det langt mere interessant, både klinisk og politisk, hvilken behandling der hjælper, og hvilken intensitet behandling bør have for bedst at kunne hjælpe mennesker med misbrugsproblemer, end om der er en lille, statistisk signifikant forskel på den ene og den anden type af undersøgelser.

Der er også et andet sigte med Prendergast og kollegers metaanalyse, nemlig at lave en analyse af variable, der er knyttet til høj effekt. Til dette formål laver de en regressionsanalyse.

Herudfra konkluderer Prendergast to ting, der kunne være relevante, hvis konklusionen var troværdig ud fra analysen: den ene er, at den bedst implementerede behandling giver den største effekt. Det lyder umiddelbart overbevisende, men med analysens mange metodologiske svagheder in mente bliver det svært at vurdere, om denne konklusion står til troende. Den anden er, at et højere niveau af teoretisk refleksion er relateret til en mindre effektstørrelse. Denne mere provokerende konklusion ville ligeledes have vundet ved at bygge på nogle lidt mere konsistente data. Hvis den står til troende, kan den betyde flere ting: Prendergast nævner nogle, nemlig at de undersøgelser, der bygger på en kompleks eller gennemtænkt teori, ligger længere fra den kliniske virkelighed, og at de mindre interessante empiriske fund stimulerer forskerne til at elaborere fantasifuldt på den teoretiske baggrund. Men en anden mulighed, som gør at denne konklusion måske er tvivlsom, kunne være, at de mere teoretisk funderede studier er mere samvittighedsfuldt udført, lider mindre

af publikations- eller rapporteringsbias, oftere er randomiseret og derfor lettere giver anledning til, at negative sammenhænge bliver rapporteret. Eksempelvis er sandsynligheden for, at man overhovedet vil opfatte noget som en "undersøgelse", hvis man har sat en gruppe klienter til at spille softball, og en anden gruppe til at lave noget andet, nok højere, hvis man har et positivt resultat. Endvidere kan de teoretisk gennemtænkte behandlinger være givet til tungere misbrugere, med flere hjemløse, arbejdsløse, psykiatrisk dårligt fungerende misbrugere – alt sammen variable, som man ikke har målt i undersøgelsen. Ikke-behandlingen kan også have været mere behandling i de teoretisk veldefinerede behandlinger, hvorved forskellen på behandling og ikke-behandling kan være mindsket. Som nævnt er ikke-behandling meget forskelligt i Prendergast's analyse, og han opstiller ingen undergruppedata, der kunne belyse dette spørgsmål.

Alt i alt kan undersøgelsen ikke bidrage til udvikling af en evidensbaseret praksis i forhold til behandling af misbrug og afhængighed, og selve dens hovedkonklusion, at behandling er bedre end ingen behandling, er, bortset fra at være ret intetsigende, ikke understøttet i data. De teoretiske implikationer for forståelsen af såvel misbrug og afhængighed som virkningsmekanismerne i behandling af misbrug og afhængighed må siges at være helt uklare.

SBU-rapporten

SBU-rapporten (SBU 2001, bind 2) lider som Prendergast og kollegers rapport af svagheder, om end den dog har langt flere styrker. SBU-rapporten er tidligere blevet omtalt og anmeldt i NAT. I Poikolainens anmeldelse blev påpeget, at rapporten blandt andet savner analyser af heterogeniteten i effektstørrelserne, og at der de fleste steder mangler en rele-

vant analyse af publikationsbias (Poikolainen, 2002). Poikolainen påpeger endvidere, at der savnes visse dele af procedurebeskrivelsen i analysen, herunder beskrivelser af interrater-pålidelighed. Rapporten består af flere, meget forskellige dele. Bergmark kritiserer rapporten på en række punkter, blandt andet for at have uheldige definitioner af behandling (2001).

Visse dele af SBU-rapporten indeholder slet ikke metaanalyser, og en del af konklusionerne drages på basis af narrative reviews. Eksempelvis har man i hele kapitlet om medicinsk behandling af narkotikamisbrug ikke ment, at metaanalyser var meningsfulde, og mange andre spørgsmål har man afstået fra at analysere gennem metaanalyser, selv om visse af de metaanalyser, der indgår i rapporten, for et overfladisk blik har en helt tilsvarende heterogenitet i materialet (om den har den statistiske heterogenitet, er det desværre ikke så let at sige noget om).

Den største svaghed i SBU-rapporten er imidlertid det samme som hos Prendergast og kolleger: fraværet af det klinisk relevante spørgsmål, som kunne blive omdrejningspunkt for den enkelte analyse. I stedet for at kondensere og fokusere på klinisk relevante problemer og klinisk relevante løsninger forsøger SBU-rapporten at besvare generelle spørgsmål.

Dermed lever den desværre ikke op til sit opdrag om at udgøre en medicinsk teknologivurdering.

I kapitlet om behandling af alkoholmisbrug opdeles a priori i forskellige typer af behandling: 1) behandling rettet mod at motivere til ændring, 2) metoder til forandring af selve misbrugsadfærden, 3) metoder der søger at fjerne årsagerne bag misbruget 4) alment støttende rådgivning og 5) behandling rettet mod partner eller pårørende.

Denne inddeling udbygges videre, så der laves specifikke analyser af forskel på behandlingseffekt hos mænd og kvinder, hjemløse misbrugere og psykisk syge misbrugere.

I opsummeringen af undersøgelserne rapporteres signifikansen af de enkelte effekter, som den tolkes af gruppen. Men hvad vil det eksempelvis sige, at effekten for hjemløse misbrugere eller kvinder er sådan og sådan? Eller at effekten af metoder til at ændre selve misbrugsadfærden er sådan og sådan? Som i Prendergast og kollegers analyse kan man her ikke se, hvilke metoder der kan betegnes som dokumenteret effektive, og hvilke der må betegnes som ikke dokumenteret effektive. Dermed bliver hele spørgsmålet om behandlingseffekt mere et spørgsmål om diskussionen om, hvorvidt behandling virker, end et spørgsmål om, hvilken behandling der er evidensbaseret. Sammenligner man dette med en tilsvarende metaanalyse af behandlinger for cancer, så kunne man tænke sig en analyse, hvor man spørger: er cancerbehandling effektiv. Derefter inddeler man, som SBU-rapportens analyse af alkoholbehandlinger, i en række typer af målsætninger: motivation og adfærdsændring, og var det cancer kunne det eksempelvis være vækst i cancerceller og levetid. Endelig inddeler man så efter forskellige typer af cancer, såsom cancer i lever, hjerne, lymfesystem, mave, knogler og så videre. Til slut sammenlægger man effekten af alle forskellige typer af behandling i hver gruppe, kirurgiske indgreb, kemoterapi, strålebehandling, kostbehandling og så videre. Lad os antage, at man så kan konkludere, at eksempelvis behandling af testikelkræft har en højere effekt end behandling af kræft i leveren. Dette er vigtigt nok at vide, men det giver ikke noget grundlag for evidensbaseret behandling. At en slags kræft har vist sig at kunne behandles bedre end en anden slags kræft siger jo intet om, hvilken behandling der har vist sig dokumenteret effektiv, og hvilken behandling der aldrig har vist sig effektiv i kliniske undersøgelser. Er det alt hvad man ved, kan man risikere, at flere testikelkræftpatienter må lide en unødvendig død, alene fordi man glemte at stille det mere praktiske spørgsmål: hvad er

det så, der virker på testikelkræft.

I kapitlet om psykosocial behandling af stofmisbrug udgør inddelingen i støttende metoder, mere adfærdsorienterede metoder og mere psykoterapeutiske metoder et stort fremskridt i forhold til Prendergast og kollegers analyse. Her nærmer vi os en beskrivelse af behandlingsformer, der kan defineres som forskellige, og som dermed kan sige os noget om, hvilken type af behandling, der er relevant for hvilke grupper af klienter. Alligevel svækkes konklusionerne på visse punkter. Et problem er, at opdelingen efter hovedstof ikke nødvendigvis er den afgørende. Kokainafhængige kan være mennesker i arbejde, med familie, bolig og en kassekredit, der ikke kan mere, eller med en arbejdsgiver, der har slået bremsen i over for sygefraværet; eller der kan være tale om crackrygere, der har tilbragt 20 år dels i kondemnerede bygninger med rygerøret i hånden, dels med at lave indbrud og sælge stoffer og dels i fængsel, siden de forlod deres enlige, arbejdsløse mor som 13-årige; og det samme gælder opiatafhængige: de kan ligeledes være lønmodtagere, der har familieliv og en stærk social integration, eller de kan være hjemløse misbrugere med store sociale problemer. Her er der nok tale om, at "different strokes for different folks" er på sin plads. En anden sag er så, at data for psykosociale belastninger på misbrugere inden behandling er noget, der ikke altid er lige godt beskrevet i kliniske undersøgelser.

Opdelingen i tre typer af behandling, støttende, adfærdsorienteret/omlærende og psykoterapeutisk, er forholdsvis kompleks og kræver i det mindste en kvalificering i form af klare kriterier og rapport af inter-rater pålidelighed. Endvidere savnes her Q-statistikken for heterogenitet i effektstørrelser, netop fordi alle tre grupper af behandling rummer en ganske betydelig grad af heterogenitet med hensyn til både terapeutisk metode og dermed mulige effekter.

Alt andet lige må man dog sige, at kapitler-

ne om psykosociale behandlinger går et skridt i retning af at lave klinisk relevante metaanalyser, når forfatterne i det mindste laver visse inddelinger og skelner mellem forskellige typer af målsætninger for behandlingen. Den når ikke i mål, fordi opdelingen forbliver for uklar, og fordi målgrupperne måske ikke repræsenterer den klinisk relevante opdeling, ligesom forfatterne desværre ikke er tilstrækkeligt stringente i gennemførelsen af deres opdelinger (se Bergmark 2001).

At indsnævre målgrupperne vil sige at tabe statistisk power, og det er altid et spørgsmål om at bevare metaanalysernes styrke i form af statistisk power, samtidig med at man får defineret et problem, der er konkret nok til, at det er muligt at besvare meningsfyldt. Selv om SBU-rapporten på nogle måder gør det bedre end Prendergast og kolleger, må man alligevel konkludere, at man efter læsning af rapporten ikke står tilbage med meget, der kan danne grundlag for meningsfulde kliniske eller politiske beslutninger. Rapporten vil mere end den kan, men vi må håbe, at den kan udgøre en forberedende fase til senere, mere brugbare opgørelser af litteraturen med brug af metaanalytiske metoder.

Irvin m. fl. (1999). Relapse Prevention

Irvin og hendes kolleger præsenterede en metaanalyse af effekten af tilbagefaldsforebyggelse. I modsætning til SBU-rapporten var de brede i valg af hovedstof til den enkelte analyse, men til gengæld fokuserede de kun på en enkelt metode, nemlig tilbagefaldsforebyggelse.

Ud over stof- og alkoholmisbrug behandler Irvin og hendes kollegers analyse også tobaksafhængighed i deres analyse.

De konkluderer, at relapse prevention er mest effektiv for alkoholmisbrug, fulgt af blandingsmisbrug og kokainmisbrug, og mindst effektiv for tobaksafhængighed. Frem

for at kontrollere for forskellige intervenserende variable gennem en regressionsanalyse har de valgt at lave parrede sammenligninger, sådan at de opdeler materialet i to grupper og derefter deler igen, og i hvert tilfælde analyser de heterogenitet og effektstørrelser. Denne metode er muligvis mere følsom for forskerbias, fordi den indebærer en del skøn undervejs i analysen (og det er derfor Clarke og Oxman starter kapitlet om analyse i Reviewers handbook med ordene: "Don't start here").

Til gengæld giver deres konklusioner svar på nogle spørgsmål af typen: "Er denne behandling virksom? Kan den tænkes at være virksom i min behandling? Vil den være virksom for den type af patienter, som jeg har i behandling?"

Eksempelvis kan de konkludere, at relapse prevention er lige effektiv i gruppebehandling og individuel behandling. Den er også lige effektiv i døgnregi og ambulant regi, men giver altså vidt forskellig effekt afhængig af målgrupper.

Disse pragmatiske spørgsmål er af stor betydning for behandlere, og Irvin m. fl. repræsenterer et skridt frem imod at kunne besvare disse spørgsmål, for så vidt angår relapse prevention.

En yderligere styrkelse af Irvin og kollegers metaanalyse kunne ligge i at lave en vurdering af de forskelle i udførelsen af behandlingen, der måtte ligge i de forskellige undersøgelser. Eksempler på dette kunne være vurderinger af kontrollen af gennemførelse af den terapeutiske metode, beskrivelsen af manualen for behandlingen, eller hyppigheden af samtaler eller gruppemøder.

Med hensyn til specifik effekt viser undersøgelsen, at relapse prevention kun har sin effekt i sammenligning med minimal behandling eller opmærksomhedsplacebo. Dermed kan man sige, at metaanalysen ikke udpeger relapse prevention som en generelt bedre underbygget metode end andre metoder, men blot at den eksisterende forskning støtter den-

ne behandling.

Teoretisk svækker analysen Marlatts antagelse om, at alle former for afhængighedsadfærd ("addictive behaviours") basalt set har den samme indre logik og derfor skal behandles med den samme metode (præsenteret i Marlatt og Gordon, 1985). Denne konklusion svækkes dog igen af mangelfulde data med hensyn til gennemførelsen af behandlingen, eksempelvis i form af graden af individualisering af behandlingsplaner.

Griffith m.fl. (2000): Contingency management

Griffith m. fl. er den metaanalyse i denne gennemgang, der har haft det snævraste fokus, nemlig belønning af rene urinprøver for klienter i metadonbehandling. Inklusionskriterierne her var langt mere elaborerede end i nogen af de andre analyser. I analysen indgår kun behandling med metadon, løbende urinprøvekontrol og sammenligning af klienter, der modtog belønning for urinprøver vs. klienter, der ikke modtog en sådan belønning.

Deres metaanalyse producerede data, der kan besvare klinisk og teoretisk relevante spørgsmål. Her er nogle eksempler:

- Øgning af metadondosis ved rene urinprøver gav den største effekt som belønning, fulgt af øgede taghjemdoser ved rene urinprøver.
- En kombination af positiv og negativ reinforcement gav en større og mere konsistent effekt end negativ reinforcement eller positiv reinforcement alene.
- Hurtig reinforcement gav større effekt end forsinket reinforcement.
- Hyppige urinprøver (3 gange om ugen eller oftere) gav en større effekt end sjældnere urinprøver.
- Fokus på et stof, frem for på flere stoffer, gav en større effekt.

Disse fund er alle konsistente med klassisk

adfærdsteori. De bidrager dermed til at gøre contingency management til en potentiel "guldstandard" for adfærdsbaseret behandling, hvorfor nye og bedre behandlinger for den nævnte målgruppe kan tage udgangspunkt i en sammenligning med contingency management og en opmærksomhedsplacebo.

Muligheden for at udfordre klassisk adfærdsteori, for at se den i nyt lys og undersøge de underliggende antagelser, er dog begrænset. I forhold til adfærdsteori kan CM-behandling ses som et replikationsstudie, der generaliserer fundene til en ny population.

I deres analyse har de valgt at indtage såvel randomiserede som ikke-randomiserede undersøgelser, hvilket dog svækker deres konklusioner en smule. De randomiserede undersøgelser gav en lavere effekt end de ikke-randomiserede, hvilket kan give mistanke om, at de ikke-randomiserede undersøgelseres konklusioner er for optimistiske. Men i og med at Griffith og kolleger rapporterer dette, giver de os også mulighed for at tage stilling til spørgsmålet, i modsætning til Prendergast og hans medarbejdere.

Pearson og Lipton 1999. Behandling i forbindelse med straf

Pearson og Lipton opbygger deres metaanalyse som et sæt af analyser (1999). Overordnet set behandler de undersøgelser af behandlingsinterventioner, som er foretaget i forbindelse med afsoning af domme for straffelovsovertrædelser. Imidlertid opstiller de herunder en række spørgsmål. De underopdeler efter kvalitet og herefter forskellige typer af behandling: 12-trinsbehandling, "boot camp", terapeutisk samfund, metadonvedligeholdelse, og kognitiv adfærdsterapi.

Det fælles er, at behandlingen skal foregå inden for rammerne af eksempelvis et fængsel eller en institution.

Deres rapport giver svar på en række klinisk

relevante spørgsmål: findes der overhovedet dokumenteret effektive behandlingsformer for misbrug inden for rammerne af straf, er "boot camp" dokumenteret effektiv, er terapeutiske samfund dokumenteret effektive og så videre. Inden for hver type af behandling giver de effektstørrelser underopdelt efter kvalitetsvurdering og en analyse af effektstørrelsernes heterogenitet.

Den politiske og kliniske relevans er her helt tydelig, mens det er sværere at drage teoretiske konklusioner af analysen. Ud fra en pragmatisk synsvinkel er dette oplagt en relevant metaanalyse.

Konklusion. De tre store problemer: bias, heterogenitet og garbage-in garbage-out

Biasproblemet i metaanalyser er nok det mindst afgørende. Bias kan ikke elimineres, men der findes dog relevante metoder til at kontrollere for bias, og når et givet spørgsmål har nået en vis opmærksomhed i det videnskabelige miljø, må man forvente, at også negative forskningsresultater betragtes som relevante og interessante og dermed bliver rapporteret og publiceret. Som nævnt vil dette problem dog vedblive at være alvorligt for helt nye og eksperimentelle behandlinger. Metaanalyser af behandlingseffekter hører derfor en tid til, hvor en behandlingsform har nået en vis modenhed, og der er lavet en serie af replikationsstudier, der tilsammen udgør et korpus af litteratur. Efter min bedste vurdering har softball-behandling endnu ikke nået dette stadie.

Heterogenitet med hensyn til de oprindelige undersøgelser i en metaanalyse er et uomgængeligt vilkår i metaanalyser. Det er igennem heterogenitet, at konklusionerne opnår en ekstern validitet, hvilket skal forstås sådan, at hvis en behandlingseffekt er vist på en gruppe, så kan metaanalyser bruges til at opføre, om det er vist, at denne behandlingseff-

fekt kan generaliseres til andre grupper.

Hvis materialet, som i Irvin og kollegers undersøgelse, er heterogent på den måde, at det består af patienter, der er afhængige af forskellige stoffer, så kan metaanalysen bidrage til at vurdere, hvorvidt behandlingsmetodens effekt kan generaliseres til andre misbrugsstoffer. Og når den indeholder undersøgelser, der er foretaget i forbindelse med døgnbehandling og undersøgelser, der er foretaget i ambulant regi, så kan den belyse, om metoden kan generaliseres mellem disse forskellige behandlingsmodaliteter. Dermed er heterogenitet ikke en trussel mod metaanalyseres kvalitet, men i sig selv en kvalitet ved metaanalyser. Griffith og kollegers metaanalyse kan med sin noget større homogenitet i målgrupper være svær at anvende, hvis man ikke har metadonbehandling i sit behandlingsprogram, eller hvis budgetterne ikke kan bære mere end en enkelt urinprøve per uge, eller måske endda kun stikprøvekontrol eller slet ingen urinprøvekontrol.

Garbage-in garbage-out problemet kræver ligeledes kompromisser – men de rette kompromisser. Det diskuteres i Cochrane Collaboration's håndbog (Clarke & Oxman 2002), om forskere, der kender et område godt, kan blive fristet til at ekskludere studier, hvis resultater ikke støtter en hypotese, som de har tilbragt et halvt liv med at undersøge. Man kan på den måde risikere, at man ender med at stå tilbage med, at der simpelt hen ikke er undersøgelser, der belyser det spørgsmål, man ville have analyseret, når man har opstillet nogle meget stramme eksklusionskriterier. I denne situation må man vurdere, om udviklingen står der, hvor man bør lave en klinisk undersøgelse, der kunne leve op til sådanne kriterier, frem for at lave en metaanalyse, der bliver "garbage", eller om man vil løse op på visse af eksklusionskriterierne.

Som det ses af ovenstående gennemgang, er de mest klinisk relevante metaanalyser i øjeblikket de, der har opstillet et spørgsmål til en

enkelt psykosocial behandlingsform. Med hensyn til teoretisk relevans må man også sige, at en enkelt behandling giver større teoretisk relevans. Eksempelvis kan en række af adfærdsteoriens påstande testes i Griffith og kollegers metaanalyse.

Diskussion: Er tiden moden til evidensbaseret misbrugsbehandling?

De ovenstående metaanalyser kan betegnes som forsøg på at bidrage til evidensbaseret misbrugsbehandling. Evidensbaseret behandling bygger på den antagelse, at en given behandling kan have en dokumenteret effekt. En dokumenteret effekt ses i randomiserede kliniske undersøgelser samt i metaanalytiske opgørelser af flere sådanne undersøgelser.

At en behandling er evidensbaseret betyder på den anden side ikke, at behandlingen er alle andre behandlinger overlegen. Eksempelvis findes der flere forskellige behandlinger, der er dokumenteret effektive for behandling af depressioner: SSRI-præparater, tricykliske antidepressiva, ECT, kognitiv terapi og interpersonel terapi kan alle betegnes som dokumenteret effektive behandlinger for depressioner, og man kan kun i begrænset omfang lave et evidensbaseret valg imellem disse forskellige behandlinger, hvis man har en deprimeret patient (Beutler 2000).

Derimod er benzodiazepiner eller lithiumsalt ikke evidensbaseret behandling med indikationen depression, fordi de har produceret inkonsistente resultater i kliniske undersøgelser, og samlet set ikke kan siges at have vist sig effektive i behandling af depression.

Men spørgsmålet er så, om man på samme måde kan sige, at der eksisterer evidensbaseret psykosocial behandling for afhængighed af rusmidler. På basis af Prendergast' og SBUs gennemgange er det umiddelbart svært at sige, at der eksisterer evidensbaseret behandling, mens man på basis af Griffith m.fl. kan

sige, at Contingency Management er en evidensbaseret behandling for sidemisbrug, når man er i behandling med metadon. Ligeledes kan tilbagefaldsforebyggelse betegnes som en evidensbaseret behandling for alkohol- og blandingsafhængighed efter Irvin og kollegers gennemgang, men ikke for kokain- eller tobaksafhængighed. Vi kan også sige, at inddragelse af familie i behandlingen, hvor dette er muligt, er en evidensbaseret behandling (se Wilson 2000).

I begge tilfælde er det i sammenligning med en mindre intensiv intervention, at behandlingerne har vist sig effektive, på samme måde som farmakologisk behandling viser sig mere effektiv i sammenligning med placebobehandling.

På den anden side er det muligt, måske endda troligt, at de organisatoriske forhold omkring gennemførelse af behandling spiller en mindst lige så stor rolle som den konkrete behandlingsideologi (Prendergast et al. 2000; McCaughrin og Price 1992). Spørgsmålet er så, om misbrugsbehandling derved adskiller sig fra mange andre former for interventioner i forhold til mennesker, fra hospitalsbehandling, over jordemoderstøtte under fødsler til grundskoleundervisning; at de specifikke metoder kun bidrager med en del af den samlede betydning, som interventionen får for modtagerne. Måske er det i virkeligheden ofte kun en mindre del, der kan forklares ved specifikke metoder. Dette er ikke anderledes, end at en velfungerende skole yder en langt bedre undervisning end en kaotisk skole, mens en skole med blot et minimum af organisation kan yde en noget bedre undervisning med gode lærebøger end med dårlige.

Det bør på den anden side ikke betyde, at vi beslutter os for, at alt er lige godt. For ligesom lithiumsalt og tricykliske antidepressiva ikke er lige godt som behandling af depression, er der også behandlinger, der er mere og mindre effektive over for misbrug og afhængighed. Vi ved i dag noget om behandlingsformer, der

næppe har effekt, og noget om behandlinger der har effekt.

Et par enkelte ting kan vi sige på basis af helt stringent gennemførte undersøgelser: Relapse prevention har ikke dokumenteret effekt for kokainafhængighed og tobaksafhængighed (Irvin et al. 1999). Vi ved også, at "boot camp" ikke er behandling for hverken stof- eller alkoholmisbrug eller kriminel adfærd, mens terapeutiske samfund og kognitiv terapi kan være det (Pearson & Lipton 1999).

Metaanalyser bør absolut spille en rolle i fremtidige opsummeringer af den eksisterende forskning på misbrugsbehandlingsområdet. Men ligesom en enkelt undersøgelse ikke kan tilrettelægges uden et relevant forskningsspørgsmål, sådan kan metaanalyser heller ikke tjene noget formål, uden at der ligger et videnskabeligt spørgsmål bag dem.

REFERENCER

- Bergmark, A. (2001): Psykosocial behandling av alkohol- och narkotikaproblem. Kan socialtjänsten dra några slutsatser? Nordisk alkohol- & narkotikatidskrift 18 (5-6): 510-513
- Beutler, L. E. (2002): The Dodo Bird Is Extinct. Clinical Psychology: Science and Research 9: 30-34
- Brewer, D. D. & Catalano, R. F. & Haggerty, K. & Gainey, R. & Flemming, C. (1998): A meta-analysis of predictors of continued drug use during and after treatment for opiate addiction. Addiction 93 (1): 73-92
- Carroll, L. (2000): Alice i Undreland. Nyoversat og gendigtet af Ejgil Søholm. Apostrof
- Clarke M, & Oxman A. D. (red) (2002): Cochrane Reviewers Handbook 4.1.5 [updated April 2002]. I: The Cochrane Library, Issue 2, 2002. Oxford: Update Software. Updated quarterly
- Gorman, D.M.; Derzon, J. H. (2002): Behavioral traits and marijuana use and abuse: A meta-analysis of longitudinal studies. Addictive Behaviors 27: 193-206
- Griffith, J. D. & Rowan-Szal, G. A. & Roark, R. R. & Simpson, D. D. (2000): Contingency management in outpatient methadone treatment: a

meta-analysis. *Drug and Alcohol Dependence* 58: 55–66

Irvin, J. E. & Bowers, C. A. & Dunn, M. E. & Wang, M. C. (1999): Efficacy of Relapse Prevention: A Meta-Analytic Review. *Journal of Consulting and Clinical Psychology* 67 (4): 563–570

Luborsky, L. & Rosenthal, R. & Diguier, L. & Andrusyna, T. P. & Berman, J. S. & Levitt, J. T. & Seligman, D. A. & Krause, E. D. (2002): The dodo bird verdict is alive and well -mostly. *Clinical Psychology: Science and Practice* 9: 2–12

Marlatt, G. A. & Gordon, J. R. (red.) (1985): *Relapse Prevention. Maintenance strategies in the treatment of addictive behaviors.* Guilford

McCaughrin W. C. & Price R. H. (1992): Effective outpatient drug treatment organizations: program features and selection effects. *International Journal of the Addictions* 27 (11): 1335–58

Pearson, F. S. & Lipton, D. S. (1999): The effectiveness of educational and vocational programs: CDATE metaanalyses. Paper presented at the annual meeting of the American Society of Criminology.

Toronto. ON.

Poikolainen, K. (2002): SBU lyckades inte bringa reda i missbrukarvården. *Nordisk alkohol- & narkotikatidskrift* 19 (3): 224–225

Prendergast M. L. & Podus D. & Chang, E. (2000): Program factors and treatment outcomes in drug dependence treatment: an examination using meta-analysis. *Substance use and Misuse* 35 (12–14): 1931–65

Prendergast M. L & Podus, D & Chang, E & Urada, D. (2002): The effectiveness of drug abuse treatment: a meta-analysis of comparison group studies. *Drug and Alcohol Dependence* 67: 53–72,

SBU (2001): *Behandling av alkohol- och narkotikaproblem. En evidensbaserad kunskapssammanställning.* SBU. Statens beredning för medicinsk utvärdering. Rapport nr. 156. Volym I–II

Wilson D. B. (2000): Metaanalyses in alcohol and other drug abuse treatment research. *Addiction* 95 (Suppl. 3): 419–38.

Summary

Morten Hesse: *Meta-analysis as basis for evidence-based substance abuse treatment. A discussion of five meta-analyses of treatment outcome.*

Evidence-based medicine rests on the assumption of a hierarchy of evidence. The highest level of evidence is assumed to come from meta-analyses of randomised clinical trials.

Five illustrative recent meta-analyses of psychosocial treatment for substance dependence are discussed (Prendergast et al. 2002; Irvin et al. 1999; SBU 2001; Griffith et al. 2000; Pearson & Lipton 1999). In all cases the methodological strengths and weaknesses of the study are discussed, as well as the clinical and theoretical implications of the analysis. Studies that attempt to address very broad questions such as "does treatment work" produce conclusions that appear inconsistent and may be flawed, and the clinical and theoretical implications of the findings of such meta-analyses are unclear. In contrast, analyses of specific treatments produce findings that are clinically and theoretically meaningful. In cases where the populations or settings studied are diverse, but the type of treatment is similar, well-conducted meta-analyses may produce significant data on the degree to which the apparent effectiveness of a specific treatment can be generalised to new populations or settings.

It is concluded that the efficacy of specific classes of interventions for drug or alcohol use disorders may be a relevant target for meta-analyses, whereas meta-analyses that attempt to do too much by asking very global questions are unproductive, and that although meta-analyses are generally effective at addressing the question of whether a given treatment has support in the existing research, they are not able to answer questions about what should be the means and ends of drug policy and drug treatment at a more general level.

Keywords: Metaanalysis, evidence-based practice, drug abuse, review.

Yhteenvedo

Morten Hesse: *Meta-analyytit näyttöön perustuvan päihderiippuvuushoidon perustana. Käsitteilyssä viisi hoidon tehokkuuden meta-analyysejä*

Näyttöön perustuva lääketiede nojautuu oletukseen näyttöjen hierarkisuudesta. Korkeimman tason näyttönä pidetään tietoa, jota saadaan satunnaistettujen kliinisten kokeiden meta-analyyseistä.

Tutkimuksessa käsitellään viittä havainnollista viime aikoina tehtyä päihderiippuvuuden psykososiaalisen hoidon meta-analyysejä (Prendergast et al. 2002; Irvin et al. 1999; SBU 2001; Griffith et al. 2000; Pearson & Lipton 1999). Kussakin tapauksessa käsitellään tutkimuksen metodologisia vahvuuksia ja heikkouksia, sekä analyyisin kliinisiä ja teoreettisia seuraamuksia. Tutkimukset, joissa pyritään vastaamaan hyvin laajoihin kysymyksiin, kuten "tehoaako hoito", tuottavat epäjohtomukaisilta vaikuttavia ja usein virheellisiä johtopäätöksiä. Kyseisten meta-analyyseiden tulosten kliiniset ja teoreettiset seuraamukset ovat epäselviä. Sitä vastoin yksittäisten hoitomuotojen analyysejä tuottavat tuloksia, jotka ovat kliinisesti ja teoreettisesti merkittäviä. Tapauksissa, joissa tutkitut populaatiot tai ympäristöt vaihtelevat mutta hoitomuoto on sama, huolellisesti tehdyt meta-analyysejä voivat tuottaa merkittävää tietoa siitä, missä määrin tietyn hoitomuodon todettu tehokkuus on yleistettävissä uusiin populaatioihin tai ympäristöihin.

Johtopäätöksenä todetaan, että tiettyjen huume- ja alkoholiriippuvuuden hoidossa käytettyjen interventiotyyppien tehokkuus voi soveltua hyvin meta-analyyseihin kohteeksi, kun taas meta-analyysejä, joissa etsitään vastauksia liian kaikenkattaviin kysymyksiin, ovat tehottomia. Vaikka meta-analyyseistä selviää helposti onko tietyllä hoitomuodolla tutkimuksen tuki takanaan, ne eivät tuo vastauksia kysymyksiin siitä, minkälaisia huumeepoliittikan ja hoitomuotojen keinojen ja lopputulosten tulisi olla yleisemmällä tasolla.