

EVIDENSKRAVET I PSYKOTERAPI:
Kan der tilvejebringes metodologisk stringent evidens
for psykoterapiens virkning?

Esben Hougaard

Interessen for evidensbaseret klinisk praksis er vokset voldsomt inden for de sidste ti år (Hougaard, 2007b). Evidensbaseret praksis handler om, at oplysning, undersøgelse, forebyggelse og behandling skal bygge på forskningsresultater, hvilket har fået tiltagende betydning i takt med fremkomsten af (og tiltagende brugerefterspørgsel efter) nye behandlingsformer og de stadigt øgede sundhedsudgifter. Sundhedsmyndigheder, politikere, offentlighed og brugere ønsker informationer om dokumentation for særlige behandlings virkning, sikkerhed og omkostningseffektivitet.

Betegnelsen »evidensbaseret medicin« blev første gang anvendt af forskere ved McMaster Universitet i Canada i 1992 om et »nyt paradigme, der afsvækker betydningen af intuition, usystematiske kliniske erfaringer og patofysiologisk rationale som tilstrækkelige grunde til klinisk beslutningstagen og vægter undersøgelse af dokumentation fra klinisk forskning.« (Guyat et al., 1992). Teori og kliniske erfaringer nedprioriteres i denne definition. Den nu etablerede definition, som også blev foreslået af forskere ved McMaster Universitet, lyder imidlertid således: Evidensbaseret medicin består i »integrationen af den bedste forskning med klinisk erfaring og patientens værdier« (Sackett et al., 1996; 2001, p. 1). Det er denne, væsentligt mere fleksible definition, der i dag er den officielle inden for evidensbaseret medicin.

Evidensbaseret medicin opstod omtrent samtidigt med Cochrane-samarbejdet, kliniske praksisvejledninger eller referenceprogrammer og bevægelsen for empirisk valideret psykoterapi i Selskabet for Kliniske Psykologer i den Amerikanske Psykolog Forening (Hougaard, 2007b). Disse bevægelser var med til at sætte øget fokus på empirisk dokumentation for psykoterapiens effekt. Betegnelsen »evidensbaseret« anvendes nu inden for en række forskellige discipliner, hvorfor man mere bredt taler om evidensbaseret

praksis. Den Amerikanske Psykolog Forening vedtog i 2005 principper for evidensbaseret psykologisk praksis (APA, 2006), og i 2007 fulgte Den Norske Psykologforening trop (Aanonsen, 2007).

I artiklen her vil jeg redegøre for og diskutere evidenskravene i psykoterapeutisk forskning med fokus på dokumentation af effekt. Det synspunkt vil blive fremført, at det er vigtigt at kende til metodologien i de såkaldte randomiserede kontrollerede undersøgelser og at forbedre metodologien i psykoterapeutisk effektforskning, men at der er forskelle i mulighederne for at opnå evidens i psykoterapi og i medicin, og at de fleste praktiske spørgsmål i forbindelse med psykoterapi må besvares ud fra plausibilitet snarere end stringent evidens.

Individualiseret evidensbaseret praksis

Megen kritik af evidensbaseret psykologisk praksis har rettet sig mod en forestilling om, at den degraderer psykoterapeuten fra en »disciplineret, improviserende kunstner« til en »manualdrevet tekniker« (Bohart, O'Hara & Leitner, 1998, p. 145). Denne opfattelse knytter sig imidlertid til bevægelsen for empirisk valideret psykoterapi snarere end til evidensbaseret praksis, hvor man har taget afstand fra »køgebogsmedicin« (Sackett et al., 1996). Og selv om bevægelsen for empirisk valideret psykoterapi fordrede brug af manualer i forskningen for at sikre psykoterapiens teoretiske integritet, har man aldrig foreslået, at al psykoterapi i praksis skal foregå efter manual (Hougaard, 2007a). Som fremhævet af Bonnie Spring (2007) er den etablerede definition af evidensbaseret praksis (Sackett et al., 2001) udtryk for idiografiske principper for beslutningstagen med tre selvstændige led: 1) brugen af den bedste, tilgængelige forskning, 2) klinisk erfaring og kompetence og 3) patientens karakteristika og ønsker.

Den Amerikanske Psykolog Forenings retningslinjer

I 2005 nedsatte præsidenten for den Amerikanske Psykolog Forening (APA), Ronald F. Levant, en arbejdsgruppe, der fik til opgave at udarbejde en rapport om evidensbaseret praksis for psykologer. Arbejdsgruppen bestod af et bredt udvalg af forskere og klinikere, herunder også personer med skarpe kritikpunkter over for empirisk valideret psykoterapi som Bruce Wampold (2001) og Drew Westen (Westen, Novotny & Thompson-Brenner, 2004). Arbejdsgruppen fuldførte arbejdet i enighed med en rapport, der blev vedtaget som foreningens officielle politik (APA, 2006).

Evidensbaseret psykologisk praksis defineres af APA som »integrationen af den bedste, tilgængelige forskning med klinisk ekspertise sammenholdt med patientens karakteristika, kultur og præferencer« (APA, 2006, p. 273).

Definitionen lægger sig tæt op ad Sackett et al.'s (1996) definition af evidensbaseret medicin, men rapporten indeholder en uddybet indholdsbestemmelse af klinisk ekspertise og af patientkarakteristika. APA anvender betegnelsen *patient*, men det vedkendes, at betegnelser som klient, person eller bruger i nogle sammenhænge kan være mere hensigtsmæssig.

Det fremhæves, at der er forskelle mellem evidensbaseret psykologisk praksis og empirisk valideret psykoterapi ved at førstnævnte tager udgangspunkt i patienten (og ikke i behandlingsmetoden), ved at der fokuseres på andre områder end behandling, samt ved at vurderingen sker ud fra mange forskellige forskningsformer.

Forskellige former for forskningsevidens

APA (2006) opstiller en liste over forskellige forskningsdesigns og deres hovedanvendelse (se Tabel 1).

Tabel 1. *Forskellige forskningsformer anbefalet af APA (2006)*

-
- a) Klinisk observation, individuelle case-studier og almen psykologisk videnskab er værdifulde i forbindelse med innovation og hypotesedannelse
 - b) Kvalitativ forskning kan anvendes til at beskrive patienters subjektive oplevelser
 - c) Aggregerede case-studier kan anvendes som en form for praksisforskning
 - d) Enkelt-case eksperimentel forskning er særligt egnet til at etablere kausale sammenhænge på individniveau
 - e) Epidemiologisk og etnografisk forskning er særligt egnet til at undersøge tilgængelighed, brug og accept af særlige behandlingsmetoder
 - f) Proces-effekt-studier er særligt egnede til at identificere forandringsmekanismer
 - g) Naturalistisk forskning under almindelige kliniske vilkår (forskning i klinisk brugbarhed eller »effectiveness«) er egnet til at undersøge behandlings økologiske eller eksterne validitet
 - h) Randomiserede kontrollerede forsøg er standarden for at drage konklusioner om behandlings effekt (»efficacy«) med kontrol af interne validitets-trusler
 - i) Metaanalyser er et systematisk middel til at syntetisere forskning, teste hypoteser og estimere effektstørrelser
-

Der er altså tale om en bred vifte af forskningsstrategier afpasset til forskellige formål. Som inden for evidensbaseret medicin anses randomiserede kontrollerede kliniske forsøg (RCT, »Randomized Controlled Trial« eller »Randomized Clinical Trial«) for standarden for at drage konklusioner om

behandlingseffekt (pkt. h i Tabel 1), særligt i forbindelse med metaanalyser af flere RCTs (pkt. i). Det anføres dog også – i lighed med foreningens tidligere anbefalinger (APA, 1995; 2002) – at der bør lægges vægt på den tostrengede evaluering af hhv. behandlingseffekt i kontrollerede forsøg («efficacy») og af klinisk brugbarhed i daglig praksis («effectiveness») (jf. pkt. g).

Randomiserede kontrollerede forsøg

Randomisering betyder statistisk tilfældig fordeling, og RCT er guldstandard i behandlingsforskning inden for evidensbaseret medicin. Ifølge Sackett et al.s (2001) håndbog om evidensbaseret medicin er randomisering den vigtigste metodiske forholdsregel i forbindelse med behandlingsforsøg. Hvis der findes randomiserede undersøgelser på området, og en undersøgelse ikke er randomiseret, så spring den over, som det udtrykkes (p. 108). Randomisering sikrer gennem tilfældig gruppeallokering, at forskelle mellem grupperne, der kunne have indflydelse på resultatet, statistisk udraderes, da variablerne i gennemsnit fordeler sig ligeligt i grupperne. På denne baggrund anses RCT for den optimale strategi til at kontrollere interne validitetstrusler i behandlingsforskning, således at fundne forskelle mellem grupperne pålideligt kan tilskrives behandlingen (Shadish, Cook & Campbell, 2002).

Der er kommet stærkt øget opmærksomhed på metodisk stringens inden for medicin i kølvandet på en række eksempler på fejlagtige konklusioner baseret på forskning med metodiske fejl – ikke mindst i forbindelse med sager, hvor medicinalfirmaer har haft økonomiske interesser. Således er de metodiske krav til medicinsk forskning blevet stærkt forøgede inden for de sidste ti år. Psykologer er traditionelt rimeligt velbevandrede i eksperimentel logik, men den praktiske afholdelse af RCTs giver særlige problemer, som har fået mindre bevågenhed blandt psykologer end blandt medicinske forskere (Spring, 2007). Det er da også påpeget, at færre psykologiske end medicinske publikationer lever op til almindeligt anerkendte metodekrav for RCTs (Spring et al., 2007). Det gælder imidlertid også for psykoterapeutisk forskning, at kravene til metodisk kvalitet er skruet i vejret inden for de senere år.

Hovedkriterier for et velgennemført RCT er anført i Tabel 2. Som det fremgår af kravene i tabellen, er det en stor og kompliceret opgave at gennemføre et velkontrolleret RCT efter metodisk stringente kriterier.

Det er vigtigt at være opmærksom på, at meget få psykoterapiundersøgelser opnår høje scores på skalaer til bedømmelse af metodisk kvalitet efter sådanne kriterier, og jeg kender ingen, som tilfredsstillende opfylder alle krav. Der er imidlertid tale om en nyttig tjekliste, man kan konsultere for at vurdere, hvilke krav der er mulighed for at opfylde inden for en undersøgelses praktiske muligheder. Alle undersøgelser har metodiske svagheder, og selv

mindre gode undersøgelser indeholder som regel mere pålidelig information end den, som tilvejebringes ud fra kliniske indtryk eller vurderinger. Undersøgelser skal primært vurderes i forhold til vidensniveauet på det pågældende område og ikke ud fra absolutte standarder. Ved den indledende afklaring af særlige metoders virkning er det ikke hensigtsmæssigt at starte med et stort RCT, men derimod med mindre krævende forskningsmetoder, fx systematisk enkelt-caseforskning eller case-serier (jf. Tabel 1, pkt. c og d).

Kravene til et velgennemført RCT er forståelige ud fra almen metodologi som strategier til at imødegå potentielle validitetstrusler i forbindelse med kausalslutninger (Shadish et al., 2002). Principielt set er de lige så relevante for psykologisk som for medicinsk forskning. Det er veldokumenteret, at fravigelse af kravene vil kunne føre til fejlkilder; ofte med den konsekvens at behandlingsresultatet overvurderes (Wulf & Gøtzsche, 2000; Sackett et al., 2001). Nogle af kravene er imidlertid umulige at implementere i psykoterapeutisk forskning (fx dobbelt eller tredobbelt blindhed; Tabel 2, pkt. 6), og nogle, såsom opnåelse af tilstrækkelig power (pkt. 5), er særdeles vanskelige at tilgodese (jf. senere). Utilstrækkelig power i enkeltundersøgelser kan kompenseres ved at opgøre resultaterne på tværs af studier gennem metaanalyser.

Tabel 2. *Optimale krav til kontrolleret effektforskning*

-
1. Der skal være tale om en velafgrænset og relevant klinisk problemstilling^{abc} med specifikke formål og hypoteser.^a
 2. Den videnskabelige baggrund for problemstillingen skal være klart beskrevet.^a
 3. Fordeling af forsøgspersonerne skal foregå efter eksplicit beskrevet randomisering.^{abcde} (Fx tilfældige talserier fra statistisk værk eller frembragt pr. computer).
 4. Randomiseringens resultat skal holdes skjult for forskere og behandlere.^{abce} (Helst skal den foregå et andet sted end behandlingen).
 5. Stikprøven skal være tilstrækkelig stor til at sikre rimelig chance for signifikant resultat.^{ac} Der skal være forhåndsudført power-analyse.^a
 6. Forsøgspersoner, behandlere og forskere skal så vidt muligt være »blinde«
mht. behandlingen.^{abce} (Normalt kan kun forskerne/bedømmerne være »blinde« i psykoterapeutisk forskning. Undertiden tjekkes her behandlingernes »troværdighed« for patienterne som en kompensation for deres manglende blindhed).
 7. Alle relevante slutresultater (outcome) skal være målt standardiseret, gyldigt og pålideligt.^{abcd}
 8. Hovedeffekt mål skal være specificeret på forhånd.^a (For at undgå »signifikansfiskeri«).
 9. Alle randomiserede patienter skal inddrages i analyserne (»intention to treat-analyse«).^{abce} (Ved frafald kan sidste observation fremskrives som slutresultatet [»last observation carried forward«] eller frafaldne tilfælde

- kan betragtes som ikkebedrede ved dikotome effektmål [bedrede/ikke bedrede]).
10. Tilfredsstillende statistisk analyse, inklusive korrektion for multiple tests^{ac}, fx ved Bonferroni-korrektion. (Her dividerer man den forhåndsfastsatte p-værdi med antallet af sammenligninger; fx fastsættes p-værdien til 0,01 i st. f. 0,05 ved fem sammenligninger til belysning af samme spørgsmål).
 11. Der skal være angivet effektstørrelser, og ikke kun signifikansniveau, for centrale resultater.^{ae}
 12. Det skal tjekkes, om resultaterne er homogene mellem de forskellige undersøgelsessteder (ved multicenterundersøgelser).^b
 13. Der skal være followup af rimelig længde (mindst 3 måneder).^{ce}
 14. Der skal være en klar beskrivelse af behandlingen, inklusive supplerende behandling.^{ac} (Herunder samtidig medicin i psykoterapiundersøgelser, og medicinen skal holdes konstant under og helst i en periode før behandlingen).
 15. Behandlingen skal udføres efter terapimanual^d (Særligt ved psykoterapi for at sikre behandlingens teoretiske integritet).
 16. Der skal være eftertjek af, om behandlingen er udført forskriftsmæssigt (adhærent) og kompetent. (Særligt ved psykoterapi).^d
 17. Grupperne skal behandles ens, bortset fra den undersøgte behandling^{cc} (Fx mht. omfanget af behandlerkontakt).
 18. Der skal være redegjort for behandlingskomplians, inklusive fremmøde til behandlingen.^c
 19. Bivirkninger eller andre utilsigtede virkninger af behandlingen skal registreres.^{ace}
 20. Der skal være en klar beskrivelse af stikprøvens rekruttering med mulighed for vurdering af dens repræsentativitet.^{ac} (Særligt ved annoncerrekruttering kan der være problemer med klinisk repræsentativitet).
 21. Der skal være specifikation af diagnosegrundlag og af inklusionskriterier.^{ac}
 22. Det skal tjekkes, om grupperne er sammenlignelige på centrale variabler inden behandlingen.^{abce} (Evt. kan der tages højde for forskelle i den statistiske test, fx ved kovariansanalyse).
 23. Frafald inden og under behandlingen skal beskrives og årsagerne angives.^{ac} (Det er almindeligt, og i nogle tidsskrifter påkrævet, at angive alle patienters skæbne fra den første screening og frem til followup i et »flowdiagram«)^a.
 24. Stikprøvens demografiske variabler (fx alder, køn, ægteskabelig status, arbejdsfunktion, socioøkonomisk status og klinisk historie) skal være anført.^c
 25. Der bør være angivelse af interesser, herunder økonomisk støtte.^c (Kræves af nogle medicinske og psykiatriske tidsskrifter).

Tabelfooter: ^a CONSORT Group (2001); ^b Sundhedsstyrelsens Tjekliste 2: Randomiserede kontrollerede undersøgelser (www.sst.dk); ^c The Cochrane Collaboration Depression, Anxiety and Neurosis Quality Rating System (Moncrieff et al., 2001); ^d Chambless & Hollon (1998); ^eSackett et al. (2001).

Systematiske forskningsoversigter

Selv store, velgennemførte undersøgelser er behæftet med fejkilder, og derfor er en enkelt undersøgelse aldrig et sikkert vurderingsgrundlag. Forskningsresultater må gentages og vurderes samlet i forskningsopgørelser.

Metodologien i forskningsoversigtsopgørelser eller reviews har været genstand for en revolution de sidste tyve år. Med de kliniske praksisvejledninger eller »guidelines« (på dansk referenceprogrammer), som vandt indpas inden for medicin og psykologi omkring 1990, blev der institueret en række rutiner for en tværfaglig vurdering af evidens og opstilling af vejledende anbefalinger inden for særlige områder, fx diagnosticering og behandling af en særlig lidelse. Omtrent samtidig blev metaanalyser etableret som standardmetoden i forbindelse med forskningssyntese. Vurderingskriterier for systematiske oversigtsartikler eller metaanalyser opstillet af Sundhedsstyrelsen ses i tabel 3 (www.sst.dk).

Tabel 3. *Sundhedsstyrelsens Tjekliste 1: Systematiske oversigtsartikler og metaanalyser*

-
1. Er der en velafgrænset og relevant klinisk problemstilling?
 2. Er der beskrivelse af den anvendte metodologi? (Der kræves en systematisk metode til identificering og vurdering af litteraturen med eksplicit beskrivelse af søgestrategier).
 3. Var litteratursøgningen tilstrækkelig grundig til at identificere alle relevante undersøgelser? (Alle relevante databaser skal afsøges. Som et minimum anføres Embase, PubMed, Cochrane-biblioteket [samt ved psykoterapi PscINFO]).
 4. Var undersøgelseernes kvalitet vurderet og taget i betragtning?
 5. Berører analysen alle potentielle positive og negative effekter af interventionen?
 6. Var det rimeligt at kombinere undersøgelseerne? (Undersøgelseerne må være rimeligt sammenlignelige mht. populationer, metoder og effektmål, ligesom resultaternes ensartethed må vurderes).
 7. Kan analysens konklusion udledes af den præsenterede evidens?
-

Kliniske praksisvejledninger

Kliniske praksisvejledninger eller referenceprogrammer udarbejdes af sundhedsmyndigheder overalt i verden efter omtrent samme skabelon (se Tabel 4). Der er tale om et omfattende og grundigt arbejde med systematisk litteraturindsamling og vurdering, hvor hver enkelt videnskabelig publikation vurderes for metodisk kvalitet på særlige skemaer (trin 4 i processen), jf. kriterier i Tabel 2 og 3 ovenfor. Udfærdigelsen af de fokuserede spørgsmål (trin 3) skal ske i det såkaldte PICO-format, hvor der specifikt spørges til,

hvilket særligt problem (P) der opnår hvilket præcist defineret »outcome« (O), når det udsættes for en velafgrænset intervention (I) set i forhold til en særlig kontrolforanstaltning (C).

Tabel 4. *Processen i udviklingen af kliniske praksisvejledninger*

-
1. Udvælgelse af emne for referenceprogram
 2. Etablering af tværfaglig arbejdsgruppe
 3. Udfærdigelse af fokuserede spørgsmål
 4. Systematisk litteratursøgning og -vurdering
 5. Udarbejdelse af udkast til referenceprogram med graduerede evidensudsagn og anbefalinger
 6. Offentlig høring og peer-review
 7. Offentliggørelse og implementering af anbefalinger
 8. Revision og opdatering
-

Efter litteraturgennemgangen udarbejdes der et udkast til referenceprogram med graduerede evidensudsagn og praktiske anbefalinger (trin 5 i Tabel 4) efter skabelonen i Tabel 5. Udkastet sendes til høring samt til peer-review blandt udenlandske forskere (trin 6), inden referenceprogrammet offentliggøres, og implementering af anbefalingerne påbegyndes (trin 7). Det anføres i programmet, inden for hvilken tidsramme (fx to år) det skal opdateres (pkt. 8).

Referenceprogrammets centrale funktion er at udforme praktiske anbefalinger. Svarende til Sackett et al.'s (2001) definition af evidensbaseret medicin er anbefalingerne alene vejledende: »Det er i sidste instans altid den enkelte behandlers eget ansvar at skønne, hvad der er rigtigt at gøre i en bestemt klinisk situation – ud fra videnskabelig evidens, erfaring, klinisk skøn og patientens ønsker.« (Sundhedsstyrelsen, 2007a, p. 11). Bemærk, at anbefalingernes graduering (fra A til D i Tabel 2) gælder evidensgrundlaget og ikke deres vigtighed ud fra en klinisk betragtning.

Også kliniske praksisvejledninger kan vurderes efter en almen skabelon i *Appraisal of Guidelines for Research and Evaluation* (AGREE, 2001 [www.sst.dk]). Vurderingerne i AGREE bygger på, at spillereglerne i Tabel 4 er overholdt, og på grundigheden i fremskaffelse og vurdering af videnskabelig evidens (jf. kriterier i Tabel 3). Der lægges imidlertid også vægt på økonomiske analyser og organisatoriske forslag til implementering af anbefalingerne. Dertil gives der point for, at arbejdsgruppen, der udarbejder praksisvejledningen, repræsenterer alle relevante faggrupper, samt at patienternes synspunkter og ønsker er søgt inddraget, fx ved at repræsentanter for patientorganisationer indgår.

Tabel 5. Sundhedsstyrelsens graduering af litteratur (www.sst.dk)

Publikationstype	Evidens	Styrke
Metaanalyse, systematisk oversigt over RCT (Randomized Controlled Trial)	Ia	A
Randomiseret, kontrolleret studie	Ib	
Kontrolleret, ikke-randomiseret studie	IIa	B
Kohorteundersøgelse	IIb	
Casekontrolundersøgelse	III	C
Deskriptiv undersøgelse		
Mindre serier, oversigtsartikel	IV	D
Ekspertvurdering, ledende artikel		

Den tværfagligt sammensatte gruppe åbner også mulighed for anbefalinger baseret på klinisk konsensus i forbindelse med spørgsmål, hvor der ikke foreligger evidens. I praksisvejledningerne udviklet af National Institute for Clinical Excellence (NICE) i England markeres sådanne anbefalinger med »GPP« (Good Practice Point). I de danske referenceprogrammer markeres de med »D✓« (Sundhedsstyrelsen 2007a) eller blot »✓« (Sundhedsstyrelsen 2007b). Således er følgende anbefalinger fra referenceprogrammet for angstlidelser baseret på klinisk konsensusvurdering og ikke på evidens: »Patienter bør så vidt muligt have et informeret valg mellem dokumenteret virksomme behandlingsmetoder (D✓)« og »Nogle patienter med svære angsttilstande bør have tilbud om fortsat, intermitterende behandlingskontakt (D✓)« (Sundhedsstyrelsen, 2007a, pp. 15-16). Referenceprogrammer bygger således både på evidens og på konsensus i ekspertpanelet.

Metaanalyser

Traditionelle, kvalitative eller narrative forskningsoversigter beskrev hver enkelt undersøgelse og opsummerede almene resultater fra undersøgelserne, eller de optalte signifikante sammenhænge ud fra en oversigtstabel (den såkaldte »boks-score«-metode). En mere effektiv måde at analysere data på tværs af enkeltstudier består i metaanalyse – dvs. »analyse af ana-

lyse«. Metoden blev først introduceret af Smith & Glass i 1977, og den er nu efterhånden blevet standardmetode i forbindelse med forskningsoversigtsopgørelser. Metaanalyser opgør den gennemsnitlige effektstørrelse (ES) på tværs af undersøgelserne. Det mest anvendte mål for ES er Cohens d , der er et udtryk for den standardiserede forskel mellem gruppernes middelværdier (M); dvs. forskellen divideret med den statistisk gennemsnitlige (»pooled«) standardafvigelse (SD) for grupperne svarende til formelen: $d = (M_{\text{behandling}} - M_{\text{kontrol}}) / SD(\text{within})_{\text{pooled}}$ (Cohen, 1988; Rosenthal, 1984).¹

Cohen (1988) foreslår at benævne d -værdier på hhv. 0,20, 0,50 og 0,80 som hhv. små, moderate og store ESs. Man kan få en intuitiv forestilling om værdien af disse ESs, ved at de omregnes til procentvise forskelle mellem grupperne, hvis resultaterne dikotomiseres i antal bedre/ikke-bedre (Rosenthal, 1984). En d -værdi på 0,80 svarer til en forskel på 37 procentpoint mellem grupperne (fx således at 68,5% har tilfredsstillende udbytte i behandlingsgruppen mod 31,5% i kontrolgruppen); en d -værdi på 0,50 svarer til en forskel på 24 procentpoint; medens en d -værdi på 0,20 svarer til en forskel på 10 procentpoint.

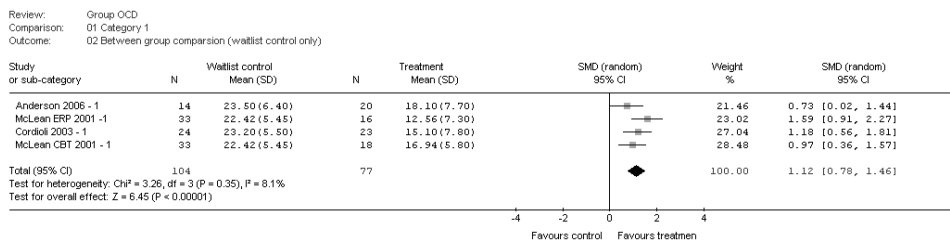
Et raffineret databehandlingssystem til metaanalyser, Cochrane Review Manager (RevMan), er udviklet inden for Cochrane Sammenslutningen (www.cc-ims.net/RevMan). Systemet fodres med oplysninger fra primærundersøgelserne om de enkelte gruppers antal prøvepersoner, middelværdier og standardafvigelser, hvorefter det udregner ESs med sikkerhedsgrænser og indtegner dem i et såkaldt »Forest Plot« (se Fig. 1). Fig. 1 omhandler fire undersøgelser, der sammenligner kognitiv adfærdsterapi (KAT) ved obsessiv-kompulsiv tilstand med ventelistekontrolgruppe (Jonsson & Hougaard, submitted), men her bringes tallene blot som illustration af den almene metode. Forrest Plot-diagrammet viser ES i form af Cohens d (Standard Mean Difference, SMD) for de enkelte studier med tilhørende 95% sikkerhedsgrænser (Confidence Limits, CL) samt den gennemsnitlige ES vægтет i forhold til antal patienter i undersøgelserne. Sikkerhedsgrænserne er også indtegnet som vandrette linjer i Forest Plottet, hvor de, som ikke overskrider 0-linien, er udtryk for signifikante ESs. Sikkerhedsgrænserne i figuren er opstillet efter den såkaldte tilfældige (Random) model, hvor det antages, at studierne er udvalgt fra en større gruppe af relevante studier (modsatningen er faste (»Fixed«) modeller, hvor studierne antages at udgøre hele populationen). Den tilfældige model foretrækkes almindeligvis, fordi den giver bedre mulighed for at generalisere til klinisk praksis.²

1 Smith & Glass (1977) anvendte kontrolgruppens SD , men det anbefales nu almindeligvis at anvende den gennemsnitlige SD for de to grupper.

2 Den almene statistiske regel lyder, af den faste model (der giver snævrere sikkerhedsgrænser og dermed større sandsynlighed for signifikans) kan anvendes, hvis metaanalysens resultater er homogene.

Fig. 1. angiver to andre statistiske udregninger i form af Q -statistik og I^2 -statistik (Higgins et al., 2003). Q -statistik omhandler sandsynligheden for, at resultaterne er heterogene (i eksemplet i Fig. 1 er p -værdien $>0,05$, så resultaterne afviger ikke signifikant fra hinanden). I^2 -statistik udregner en værdi for størrelsen af heterogeniteten. En I^2 -værdi mellem 0,30 og 0,50 opfattes almindeligvis som udtryk for moderat heterogenitet, medens værdier på mindst 0,50 er udtryk for høj grad af heterogenitet (i resultatet i Fig. 1 er I^2 -værdien på 0,08, svarende, til at resultaterne er homogene).

Fig. 1. Eksempel på »Forest Plot« i Cochrane Review Manager



RevMan kan også anvendes til metaanalyser ved dikotome effektmål (antal bedrede/ikkebedrede), hvor resultatet opgøres i form af relativ risiko (RR). RR er risikoen for ikke-bedring (eller bedring) i behandlingsgruppen set i forhold til kontrolgruppen. Hvis fx 40% er ikke-bedrede i behandlingsgruppen mod 60% i kontrolgruppen så er $RR = 0,67$ ($40/60$) svarende til en 33% mindre chance for ikke-bedring i behandlingsgruppen. Dette sidste tal (33%) omtales også som den relative risikoreduktion (RRR). En tredje mulighed for at udtrykke ES ved dikotome effektmål er Number Needed to Treat (NNT), der nu ofte fremhæves som et intuitivt lettere forståeligt mål for effekt (Sackett et al., 2001). NNT angiver, hvor mange patienter der skal behandles for at få en ekstra patient med bedring i forhold til kontrolgruppen oprundet til et helt tal. NNT udregnes som 1 divideret med den absolutte risikoreduktion, der i eksemplet er 0,20 ($0,60 - 0,40$). I eksemplet er $NNT = 5$ ($1/0,20$); fem patienter skal altså behandles for at få en ekstra bedret patient.

I de grundige kliniske praksisvejledninger udarbejdet ved National Institute for Clinical Excellence i England har man anvendt RevMan med forhåndsspecificerede grænser (ud fra klinisk konsensus) for, hvornår en ES kan anses for klinisk betydningsfuld. Fx specificerer NICE-programmet for obsessiv-kompulsiv tilstand (OCD), at der skal være tale om en d -værdi på mindst 0,50 eller en $RR \leq 0,80$ (svarende til en RRR på 20%), for at der er tale om en klinisk signifikant forskel (NICE, 2006). Hvis fx den nedre sikkerhedsgrænse for ES er højere end en d -værdi på 0,5 (hvad den jo klart

er i Fig. 1, hvor den nedre grænse er 0,78), udtrykkes det på den måde, at der er *evidens* for, at behandlingen virker bedre end kontrolforanstaltningen. Hvis der er klinisk signifikant forskel mellem gruppernes middelværdier, men sikkerhedsgrensene ikke udelukker en forskel på mindre end 0,5, udtrykkes det på den måde, at der er *begrænset evidens* for, at behandlingen virker bedre end kontrolforanstaltningen. Ud fra sådanne forhåndsfastsatte konventioner kan der opnås en præcis sprogbrug, som er letforståelig for klinikere.

Medens metaanalyser er den mest nøjagtige form for forskningssyntese, indeholder kliniske praksisvejledninger retningslinjer for, hvordan syntearbejdet i praksis kan foregå efter stringente regler med inddragelse af tværfaglig ekspertise i forslag til praktiske anbefalinger inden for særlige kliniske områder.

Problemer ved fremskaffelse af evidens

Power-problemet

Et af de store problemer i forbindelse med RCTs er antallet af prøvepersoner, som skal til, for at undersøgelsen har tilstrækkelig test-styrke eller power. En undersøgelses power er et udtryk for sandsynligheden for at undgå type 2-fejl, hvor man ikke opnår signifikante forskelle mellem grupperne, selv om der rent faktisk er en forskel. Sandsynligheden for type 2-fejl er størst i små stikprøver. Nærmere bestemt vokser sandsynligheden for type 2-fejl (hvorved power mindskes) med: 1) mindre stikprøve; 2) lavere forhåndsfastsat *p*-værdi (fx 0,01 i stedet for 0,05); 3) mindre forskel mellem gruppernes middelværdier; og 4) større spredning (fx målt som *SD*) i grupperne som et udtryk for resultaternes variabilitet. Pkt. 3 og 4 indgår samlet i Cohens *d* (se ovenfor).

Med fastlagt *p*-værdi kan man ud fra forventet ES beregne, hvor mange personer der skal indgå i undersøgelsen for at opnå en bestemt power-værdi. Konventionelt anbefales det at tilstræbe en power-værdi på 0,80, dvs., at der er 80% chance for at opnå signifikant resultat. Der er power-tabeller, hvor man kan slå resultatet op, i mange statistiske indføringsværker, herunder hos poweranalysens »fader«, Jacob Cohen (1988).

Ved sammenligning af to grupper ved hjælp af *t*-test og en fastsat *p*-værdi på 0,05 kræves følgende antal forsøgsparticipanter for at opnå en power på 0,80 ved forskellige ES:

1. Lille ES, dvs. $d = 0,20$: 786 deltagere (393 i hver gruppe)
2. Moderat ES, dvs. $d = 0,50$: 128 deltagere (64 i hver gruppe)
3. Stor ES, dvs. $d = 0,80$: 52 deltagere (26 i hver gruppe)

Almene metaanalyser af psykoterapi (Lambert & Ogles, 2004) peger i retning af, at forskellen mellem psykoterapi og hhv. ventelistekontrol, placebo-kontrol og i forhold til andre psykoterapiformer gennemsnitligt er hhv. stor, moderat og (højst) lille. Så hvis vi forventer tilsvarende resultater i en ventelistekontrolundersøgelse, vil en tilfredsstillende power kunne opnås med 52 deltagere. For at udkonkurrere en middelgod placebobetingelse kræves 128 deltagere. Hvis der derimod er tale om en undersøgelse, der sammenligner velrenommerede og veludførte psykoterapiformer, og hvor vi forventer en lille forskel ($d = 0,20$) i den ene terapiforms favør, skal vi helt op på 786 deltagere. En lille forskel svarende til 10 procentpoint kan godt være praktisk betydningsfuld: Hver gang vi behandler 100 patienter med den mest effektive metode, får vi 10 flere med tilfredsstillende behandlingsresultat.

Power-problemet findes naturligvis i al forskning, men det er særligt vanskeligt at opnå tilstrækkeligt store stikprøver i psykoterapiforskning, hvor behandlingen strækker sig over en længere periode og kræver betydelige ressourcer. Der er således mig bekendt ingen komparative psykoterapiundersøgelser med tilstrækkelig power til at påvise en lille forskel i effekt.

Selektivt frafald

Frafald er en alvorlig fejlkilde i RCTs, fordi frafaldet i grupperne ikke kan anses for tilfældigt, hvilket ødelægger randomiseringens statistisk sikrede gruppesammenlignelighed. Frarafaldet er generelt set betydeligt i psykoterapi. En metaanalyse ud fra 125 undersøgelser af Wierzbicki og Pekarik (1993) fandt et gennemsnitligt frafald på hele 46,9%. En metaanalyse af kognitiv adfærdsterapi for panikangst og generaliseret angst fandt dog kun et frafald på 11% (Mitte, 2005a;b), hvilket måske bedre afspejler frafald under almindelige kliniske forhold. Nogle oversigtsarbejder tyder på større frafald i undersøgelser, der sammenligner psykoterapi og medicin end i undersøgelser af psykoterapi alene (Sundhedsstyrelsen, 2007a), hvilket måske kan hænge sammen med forskelle i patientpræferencer.

Vi kan illustrere problemerne med en stor flercenterundersøgelse, der sammenlignede adfærdsterapi, medicin (clomipramin), kombinationsbehandling med både adfærdsterapi og medicin, samt placebo over for patienter med OCD (Foa et al., 2005). Undersøgelsen, der påbegyndtes i 1990, screenede 833 patienter, men udelukkede 684; for godt halvdelen (372) vedkommende fordi patienterne nægtede at deltage i randomiseringen, heraf 56 fordi de ikke ville risikere af få medicin, og 54 fordi de ikke ville risikere at få adfærdsterapi. Af de 37 patienter, der blev allokeret til adfærdsterapi, sprang 8 fra inden terapiens start mod 11 ud af 47 i medicingruppen. Under behandlingen sprang yderligere 8 fra i adfærdsterapigruppen og 9 i medicingruppen svarende til et samlet frafald på 43% i de to grupper. Det forekommer sandsynligt, at patienter med præference for psykoterapi hovedsageligt sprang fra i medicingruppen, og patienter med præference for medicin hovedsageligt sprang fra i psykoterapigruppen.

Som omtalt i Tabel 2 (pkt. 9) består den gængse strategi til at imødegå selektivt frafald i at lave »intention-to-treat-analyser«, hvor alle patienterne indgår med fremskrivning af deres senest registrerede effektdata (der evt. kan være målt før terapiens påbegyndelse). Men er det relevant at få oplyst virkningen af en behandling over for de personer, som ikke modtager den? eller ved sådanne, der ikke bryder sig om at modtage den? Man kan og bør naturligvis lave såvel intention-to-treat-analyser som completeranalyser, men det løser ikke problemet med den ødelagte randomisering ved stort frafald. Foa et al.'s (2005) ovenfor omtalte stort anlagte undersøgelse kompromitteres således betydeligt af det store frafald.

Terapeuteffekten

Medens medicinsk behandling i almindelighed kan specificeres ret præcist ud fra præparat og dosis, er psykoterapi langt vanskeligere at standardisere, da behandlingen i betydelig grad afhænger af terapeuten. En række undersøgelser har sandsynliggjort, at terapeutfaktorer kan spille en større rolle for psykoterapiens resultat end metodefaktorer (Wampold, 2001; Kim, Wampold & Bolt, 2006), selv om resultatet ikke er helt entydigt (Elkin et al., 2006).

Man har søgt at tage højde for problemet med terapiens forskelligartede udførelse ved at kræve, at der anvendes terapimanual og kvalitetskontrol af behandlingen; et krav, der blev markant fremført som betingelse for evidens for en særlig behandlingsmetode af bevægelsen for empirisk valideret psykoterapi (Task Force, 1995). Terapimanual anvendes for at sikre undersøgelsens teoretiske validitet (Shadish et al., 2002) ved at tjekke, om behandlingen udføres forskriftsmæssigt og kompetent; dvs., om det faktiske er den terapiform, undersøgelsen drejer sig om, der anvendes af terapeuterne. Man har fundet, at anvendelsen af terapimanual mindsker terapeutfaktorens indflydelse, men den fjerner den ikke (Crits-Christoph et al., 1991).

Gruppering af terapiformer

Et andet problem, der følger af, at psykoterapi er vanskeligere præcist at indholdsbestemme end medicin, består i at rubricere psykoterapiformer sammen i forbindelse med evidensoversigter. Der er store forskelle mellem forskellige varianter af fx psykodynamisk psykoterapi eller kognitiv adfærdsterapi. Således kan kognitiv adfærdsterapi ved depression omfatte traditionel kognitiv terapi ad modum Beck et al. (1979), hvor de kognitive metoder har prioritet, adfærdsaktiveringsterapi (Jacobson, Martell & Dimidjian, 2002), der slet ikke anvender traditionelle kognitive metoder, og adfærdsanalytisk terapi (»Behavioral Analysis System of Psychotherapy«; McCullough, 2000), der nærmest kan betragtes som en kreativ syntese af adfærdsterapi og interpersonel psykodynamisk terapi (jf. Sundhedsstyrelsen

2007b). Er det rimeligt at gruppere disse metoder sammen som forskellige varianter af kognitiv adfærdsterapi, eller skal de snarere betragtes som særskilte terapiformer?

Inden for bevægelsen for empirisk valideret psykoterapi forsøgte man at løse problemet med de uklare grænser mellem terapiformer ved at angive evidens for den specifikke manualbaserede version af psykoterapi (Task Force, 1995). Denne løsning har imidlertid den hage, at der som regel kun er udført få studier af psykoterapi baseret på en særlig manualudgave af terapien, hvorfor der ikke kan være høj grad af evidens for den (bevægelsen for empirisk valideret psykoterapi opstillede ret lempelige krav til evidens; jf. fx Chambless & Hollon, 1998). Dertil kommer så, at meget få psykoterapeuter anvender manual i daglig terapeutisk praksis (Hougaard, 2007a), hvorfor der alligevel kommer et fortolkningsmoment ind i spørgsmålet om udøvelsen af evidensbaseret psykoterapi.

Bløde effektmål

I de fleste former for medicinsk behandling findes der objektive mål for effekt ud fra fysiologiske målinger eller klart konstaterbare kropslige ændringer, herunder død/overlevelse. Udbyttet af psykoterapi måles derimod helt overvejende ud fra patientens subjektive tilkendegivelser, uanset om der er tale om patientspørgeskemaer eller kliniske bedømmerskalaer (Strupp, Horowitz & Lambert, 1997). Det er vigtigt, at effektmålene har tilfredsstillende psykometriske egenskaber (jf. Tabel 2, pkt. 7). Men selv psykometrisk tilfredsstillende skalaer savner klare oplysninger om, hvad et givet resultat betyder ud fra objektive standarder; de har, som det udtrykkes »arbitrære metrikker« (Kazdin, 2006). Værdien af et fald på fx 10 point på Hamilton Depression Skala eller Beck Depression Inventorium er ikke intuitivt forståelig, og der er heller ikke nogen objektiv logik, som begrunder, at fx 50% reduktion af depressive symptomer på Hamilton Depression Skala eller 35% reduktion af obsessive-kompulsive symptomer på Y-BOCS er udtryk for et tilfredsstillende resultat, selv om der kan opnås konsensus om sådanne grænseværdier (hvilket det også kniber med). Dertil kommer den meget store variation i skalaer, der skal måle det samme. En håndbog over skalaer vedrørende angst og angsttilstande angiver således flere hundrede forskellige målemetoder, alle med rimeligt tilfredsstillende psykometriske egenskaber (Antony et al., 2001).

Skalaer med arbitrær metrik er en hæmsko for tilvejebringelse af evidens (Kazdin, 2006), og den store variation i målemetoder vanskeliggør sammenligning af forskellige undersøgelser. Variationen vanskeliggør også opstilling af fokuserede spørgsmål efter det såkaldte PICO-format (se ovenfor) med klar specifikation af, hvilket »outcome« metoden vurderes i forhold til. I det ene af to nylige danske referenceprogrammer opgav man at formulere fokuserede spørgsmål (Sundhedsstyrelsen, 2007b), medens det andet for-

søgte sig med et tilnærmet format, jf. følgende spørgsmålsformulering ved psykoterapi:

»Hvilke psykoterapiformer er dokumenteret virksomme over for voksne patienter med den særlige angsttilstand primært over for lidelsens centrale symptomer (angst- og anspændingssymptomer, undgåelse, obsessioner, kompulsioner eller påtrængende erindringer) eller i forhold til mål for global ændring?« (Sundhedsstyrelsen, 2007a, p. 70).

Denne formulering er næppe helt i PICO-kravets ånd, men det var den bedste tilnærmelse, man i arbejdsgruppen kunne finde frem til. I et af NICE-programmerne udtrykkes et fokuseret spørgsmål i følgende, nærmest parodisk brede formulering: »What works for whom, and how is it best delivered?« (NICE, 2004b; p. 395).

Ekstern validitet

RCTs yder den største kontrol af interne validitetstrusler i undersøgelsen, medens der som regel gøres væsentlig mindre ud af at sikre undersøgelsens eksterne validitet og dermed muligheden for at generalisere til almindelig klinisk praksis. Nogle af de metodiske strategier, der anvendes for at maksimere undersøgelsens interne validitet, kan direkte tænkes at formindske dens eksterne validitet (Hougaard, 2003; 2004). Det gælder således tilfældig fordeling til behandling (hvor klinikerens vurdering af egnet behandling og patientens præferencer negligeres), forhåndsfastlagt behandlingsform og varighed af behandling samt selektion af diagnostisk homogene patientgrupper, som relativt sjældent forekommer i klinisk praksis. Også krav om behandlingsmanual og monitorering af terapiens udførelse, som især blev fremført inden for bevægelsen for empirisk valideret psykoterapi (Task Force, 1995), vil kunne medvirke til at fjerne »forskningspsykoterapi« fra psykoterapi i den almindelige kliniske dagligdag.

Statistisk kan man øge muligheden for at generalisere til en særlig population ved at anvende en tilfældig stikprøve af populationen. Dette sker imidlertid praktisk talt aldrig i RCTs, og det store frafald, der ofte sker allerede inden randomiseringen, ødelægger under alle omstændigheder en statistisk tilfældigt sikret repræsentativitet for patienternes vedkommende (jf. Westen et al., 2004). I undersøgelsen af behandling af OCD-patienter (Foa et al., 2005), som er refereret ovenfor, var frafaldet inden randomiseringen således på 684 af 833 patienter svarende til 82%.

Shadish et al. (2002) opstiller en »grounded theory« for, hvordan forskere i almindelighed argumenterer i forbindelse med generalisering af forskningsresultater ud fra 1) overfladelighed, 2) udelukkelse af irrelevante forskelle, 3) diskrimination mht. særlige variabler, 4) ekstrapolation ud fra principper eller 5) ud fra teoretiske mekanismer. Det er sådanne kvalitative

argumenter, vi gør brug af, når vi argumenterer for eller i mod anvendeligheden af resultater fra særlige undersøgelser, og det sker i praktisk talt alle tilfælde ud fra plausibilitet snarere end stringent evidens.

Den Amerikanske Psykolog Forening (2006) anbefaler som nævnt, at validering af psykologiske interventioner anvender en tostrengt forskningsstrategi med primært fokus på hhv. den interne og den eksterne validitet. Herved bliver naturalistisk forskning under almindelige kliniske betingelser et ekstra led i tilvejebringelsen af praktisk relevant evidens: Efter at der er dokumenteret effekt (efficacy) af interventionen i RCTs, skal det tillige godtgøres, at den er klinisk brugbar under almindelige kliniske forhold (effectiveness). Naturalistisk forskning kan nogle gange anvende eksperimentelle metoder, men mere typisk gøres der brug af »benchmarking« af effekt (dvs. sammenligning af ESs på tværs af undersøgelser eller behandlingssteder) eller lignende metoder, der almindeligvis anvendes til kvalitetssikring af behandling i daglig klinisk praksis (Lambert, 2007). Nogle har argumenteret for, at naturalistisk udforskning af psykoterapi er mindst lige så vigtig som RCTs (Lambert, 2007; Seligman, 1996).

Er der evidens for psykoterapiens virkning?

Det fremgår af de ovenfor beskrevne problemer, at evidensfastsættelse af psykoterapi efter stringente kvantitative forskningsmetodologiske forskrifter ikke er et simpelt anliggende. Men kan der alligevel opnås tilfredsstillende evidens i besvarelsen af nogle relevante psykoterapeutiske spørgsmål?

Bestræbelser for at tilgodese de særlige metodologiske krav ved RCTs inden for psykoterapiens område er af nyere dato. Det første forsøg på at efterligne medicinske RCTs var den berømte komparative undersøgelse af kognitiv adfærdsterapi, interpersonel psykoterapi, medicin (imipramin) og placebo – det såkaldte NIMH-depressionsprojekt (egentlig NIMH-TCRCP: National Institute of Mental Health Treatment of Depression Collaborative Research Project) (Elkin et al., 1989). Undersøgelsen, der omfattede 240 patienter ved tre forskellige behandlingscentre med et anslået budget på omkring 10 millioner dollars, blev planlagt i 1977, gennemført i perioden 1980-86 med den første rapport om resultater i 1989. Senere er der gennemført en række lignende undersøgelser med langt de fleste i USA. En lang række RCTs, de fleste væsentligt mindre omfattende end NIMH-depressionsprojektet, har dokumenteret effekt af nogle psykoterapiformer, overvejende varianter af kognitiv adfærdsterapi, med den højeste grad af evidens (Ia i Tabel 5) over for især angsttilstande og depressive lidelser (fx NICE 2004a;b; 2005; 2006; SBU, 2004; 2005; Sundhedsstyrelsen, 2007a; b). Der findes også en rimelig dokumentation for, at psykoterapi her i de fleste tilfælde opnår symptomreduktion i samme størrelsesorden som, men mere holdbart end, medicinsk behandling, selv om evidensniveauet for

denne konklusion ved mange særlige lidelser eller undergrupper af lidelser fortsat er beskedent (ibid.).

Der er til gengæld yderst få velgennemførte RCTs, der sammenligner forskellige psykoterapiformer over for velafgrænsede lidelser. Der er således ikke solid evidens for, at særlige psykoterapiformer virker bedre end andre, og denne konklusion gælder for alle lidelser, inklusive angst og depression. På den anden side gælder det så også, at den såkaldte »Dodo-kendelse«, at forskellige psykoterapiformer har omtrent samme virkning, også bygger på et særdeles spinkelt evidensgrundlag. Almene metaanalyser understøtter Dodo-kendelsen (fx Lambert & Ogles, 2004; Wampold, 2001), men resultatet kan skyldes dårligt kontrolleret forskning, at det overvejende er dokumenteret virksomme metoder, der sammenlignes, eller at hovedparten af den eksisterende forskning sammenligner uspecifikke terapiformer over for blandede lidelser (Hougaard, 2004). Der er *indicier* for, at kognitiv adfærdsterapi opnår noget bedre virkning ved en række lidelser, især blandt angsttilstande, men der findes ingen solid evidens.

Med de mange praktiske vanskeligheder ved at gennemføre velkontrollerede RCTs inden for psykoterapiområdet er det ikke overraskende, at der ikke forekommer »mur- og nagelfast« dokumentation til besvarelse af ret mange spørgsmål om terapiformernes relative effekt. Der er optalt mere end 400 forskellige navne på psykoterapiformer (heraf mindst 108 manualbaserede, empirisk underbyggede terapiformer for voksne patienter [Chambless & Ollendick, 2001]), og der findes mere end 300 forskellige diagnostiske kategorier i ICD-10 eller DSM-IV. Blot at parre en tiendedel af disse terapiformer og psykiske lidelser på kryds og tværs ville kræve et astronomisk antal komparative undersøgelser.

Kan evidenslister vejlede psykoterapiens praksis?

En simpel opfattelse af evidensbaseret psykoterapi, der mest forfægtes af kritikere, er omtalt som »matrix-modellen« (Hougaard, 2007a). Hvis man opstiller en tabel eller matrix og anfører fx klientens problemer eller lidelser i kolonnerne og terapiformer i rækkerne, så vil man kunne afkrydse de felter, hvor der foreligger evidens. Forestiller man sig videre, at man rubricerer klientproblemer som psykiatriske diagnoser og slavisk anvender særlige terapimanualer, så har man den udgave af matrixmodellen, som især har været skydeskive for kritik. Som nævnt er der ikke ret mange (om overhovedet nogen), der forestiller sig evidensbaseret psykoterapi efter så firkantede principper. Det er imidlertid et relevant spørgsmål, i hvor høj grad evidenslister kan vejlede klinikere i deres praktiske udførelse af terapien, og der kan også her nævnes en række problemer.

Utilstrækkelig eller forældet viden

Som nævnt er det meget få spørgsmål, som er relevante for psykoterapiens praksis, der i øjeblikket kan besvares med udgangspunkt i RCTs. Der er evidens for virkning af særlige metoder over for særlige lidelser, men metodernes relative virkning, virkning over for særlige patienttyper inden for en diagnostisk kategori, eller hvilke metoder man bør anvende i tilfælde af manglende virkning af en særlig terapiform, er yderst dårligt belyst. NICE-rapporterne, der tager hensyn til mange praktisk relevante spørgsmål, er fyldt med såkaldte GPP-anbefalinger (se ovenfor), der alene bygger på klinisk konsensus. Det forhold, at kun meget få anbefaling i NICE-rapporten om depression (2004a) (6 af 142)³ var tilknyttet den højeste grad af evidens (niveau A i Tabel 5 svarende til at konklusionen understøttes af mindst ét RCT), blev kommenteret under overskriften »A NICE mess« i et medicinsk tidsskrift (Gupta & Warner, 2007). Godt 3/4 af anbefalingerne understøttedes højst af niveau 4 evidens i Tabel 5. Tilsvarende er det påpeget, at psykiatriske praksisvejledninger generelt set har ringe gennemslagskraft i klinikken (Weinmann, Koesters & Becker, 2007) – måske ikke kun på grund af træghed mht. forandring blandt klinikere, men også fordi praksisvejledningerne savner klare svar på praktisk relevante spørgsmål.

Dertil kommer, at det kan tage lang tid at gennemføre et tilfredsstillende RCT vedrørende psykoterapi, hvor ti år ikke er en usædvanlig tidsramme. Der gik således 12 år fra starten af planlægningen af NIMH-depressionsprojektet til den første rapportering fra projektet (Elkin et al., 1989), medens Foa et al.'s (2005) ovenfor omtalte OCD-projekt var hele 15 år undervejs. Alle psykoterapiformer udvikler sig, og som Salkovskis (2002a) har udtrykt det, er de fleste kognitiv adfærdsterapeutiske behandlingsprogrammer forældede efter bare 5-10 år. Der er stor interesse i øjeblikket for såkaldte »mindfulness and accept«-stratier (Hayes, Folette & Linehan, 2004) som supplement til, eller erstatning for traditionelle kognitiv metoder i kognitiv adfærdsterapi, men der er meget begrænset evidens for disse nyere terapeutiske metoder. Baggrunden for interessen er dels, at der ikke er dokumentation for virkning af kognitive metoder i kognitiv adfærdsterapi (Kazdin, 2007; Longmore & Worrell, 2006), dels almene emotionsteorier (fx LeDoux, 1996), ifølge hvilke rent kognitive metoder har begrænset indflydelse på emotionelle systemer. Strengt taget er psykoterapi direkte baseret på evidensoversigter som regel forældet ud fra aktuelle kliniske og teoretiske standarder.

3 I teksten anføres fejlagtigt, at det gælder for 4 af 32 anbefalinger (Gupta & Warner, 2007, p. 442).

Viden om hvad?

APA (2006, p. 274) opregner blandt de åbne, ubesvarede spørgsmål i forbindelse med evidensbaseret psykoterapi, om praksis skal vejledes mht. 1) specifikke terapiformer eller terapiprotokoller/manualer, 2) særlige terapeutiske faktorer/variabler eller 3) terapeutiske principper eller processer/mekanismer? Evidensbestræbelser har hidtil hovedsageligt været rettet mod at dokumentere virkningen af særlige terapiformers »pakkelsesninger« af forskellige terapeutiske metoder svarende til den gængse fremgangsmåde ved effektforskning. Der har imidlertid også været argumenteret for, at de centrale byggestene i evidensbaseret psykoterapi bør være terapeutiske enkeltfaktorer, i form af teknikker (fx Beutler, 1983) eller ikke-tekniske faktorer (Hubble, Miller & Duncan, 1999), eller terapeutiske forandringsprocesser eller principper relateret til særlige problemmekanismer (fx Barlow, Allen & Choate, 2004; Ingram, Hayes & Scott, 2000; Rosen & Davison, 2003; Salkovskis, 2002a;b). Viden om særlige terapeutiske faktorer virkning eller om ændringsmekanismer i psykoterapi tilvejebringes normalt i såkaldt proces-effektforskning (Tabel 1 pkt. f). Stringent procesforskning er et kompliceret forehavende, og her foreligger der indtil nu meget begrænset viden (Kazdin, 2007). Også her er power-problemet prekært, eftersom virkningen af enkeltteknikker normalt må anses for mindre end virkningen af pakkelsesninger af teknikker.

Forskellige måder at bruge viden på

Det er almindeligt at skelne mellem på den ene side instrumentel eller teknologisk vidensanvendelse og på den anden konceptuel eller teoretisk vidensanvendelse (Wenneberg, 1999). Medens teknologiske vidensanvendelse kan bygge på empiriske generaliseringer, fordrer konceptuel vidensanvendelse indsigt i indhold, strukturer, mekanismer og processer, således at sammenhængen mellem empiriske variabler kan forklares.

En anden adskillelse skelner mellem regelstyret og case-baseret analyse (Prentzas & Hatzilygeroudis, 2007). Regler er udtryk for almen viden inden for et område, medens cases er specifik viden baseret på eksempler. Case-baseret analyse består i at matche et nyt tilfælde, en person eller særlig situation med et repertoire af tidligere oplevede eller beskrevne enkelttilfælde, dvs., der trækkes på specifik viden knyttet til en særlig kontekst. Begge former for vidensanvendelse har været genstand for udforskning inden for artificial intelligens, hvor det er en almindelig antagelse, at en kombination af metoderne er at foretrække (Prentzas & Hatzilygeroudis, 2007; Wang et al., 2007). Case-baseret analyse er i overensstemmelse med den måde, professionsudøvere ofte tænker på i praksis (Schön, 1987) og med teorier om praksisviden baseret på fænomenologi eller pragmatisk filosofi (Brinkmann & Tanggaard, 2007).

Evidensbaseret praksis med udgangspunkt i kliniske praksisvejledninger har eksklusivt fokuseret på teknologisk viden og regelstyret analyse, hvorved man udelader vidensformer, der er almindelige i klinisk praksis (Hougaard, 2004; 2007a).

Psykoterapi mellem naturvidenskab og humanisme

Der har været rejst tvivl om, hvorvidt det er berettiget at opfatte psykoterapi som sjælelig behandling i analogi til medicinsk behandling af legemlige sygdomme. (Hougaard, 2004). Nogle har argumenteret for, at psykoterapi basalt set mere er en humanistisk disciplin end en naturvidenskabelig disciplin (fx Wampold, 2001; 2007), eller at psykoterapi låner træk såvel fra naturvidenskab som humanvidenskab (Hougaard; 2004). Barlow (2004) foreslår at skelne mellem »psykologisk behandling«, hvor der er videnskabelig dokumentation for virkningen af særlige metoder over for bestemte lidelser, og »psykoterapi«, der virker i kraft af nonspecifikke faktorer, som er fælles for mange eller de fleste former for psykisk hjælp. Alle psykoterapeuter er antageligt opmærksomme på, at der foregår særdeles meget uden for protokol – selv ved manualbaseret psykoterapi – og sådanne uspecificerede faktorer er en central virksom bestanddel af alle former for psykoterapi (Hougaard, 2004; Wampold, 2001).

Konklusion

Stringent forskning er vigtig, for at den psykologiske profession kan udvikle sig. Det gælder også, selv om at psykoterapi først og fremmest er en humanistisk disciplin (Wampold, 2007). Principielt er der ikke forskel på de optimale spilleregler for medicinsk og psykologisk forskning; reglerne er metodologisk velbegrundede, og der er empiriske eksempler på, at deres fravigelse kan resultere i praktisk betydningsfulde fejlslutninger. Psykologer bør kende til de metodologiske principper for velgennemført effektforskning og ikke bare affærdige dem som udtryk for medicinsk ideologi.

Det er næppe et tilfælde, at nogle af de bedste psykoterapeutiske effektstudier er komparative undersøgelser, hvor psykoterapi sammenlignes med medicin. Evidensspillereglerne er bedre etablerede inden for medicinsk forskning, og spørgsmål om »psykoterapi og/eller medicin?« må nødvendigvis blive afgjort efter de etablerede vurderingskriterier inden for sundhedsområdet. Der er betydelige praktiske og teoretiske gevinster af den viden, som er indhentet gennem komparative RCTs af medicin og psykoterapi inden for angst- og depressionsområdet.

Selv om det ikke er muligt at gennemføre RCTs for alle terapiformer over for alle psykiske lidelser, vil det være yderst relevant at gennemføre metodisk stringente undersøgelser til afklaring af forholdet mellem de vigtigste

terapiformer, fx psykodynamisk terapi og kognitiv adfærdsterapi, over for hovedformer af psykiske lidelser. Hvis der blev investeret lige så meget i udforskningen her som i udforskningen af forholdet mellem psykoterapi og medicin inden for angst- og depressionsområdet, ville vi blive væsentligt klogere på psykoterapi – uanset om undersøgelserne vil støtte eller modsige Dodo-kendelsen.

Det er på den anden side også vigtigt at huske på, at psykoterapi er et særdeles komplekst fænomen, og at ambitionerne for tilvejebringelse af evidens må være mindre her end inden for medicin. Evidens ud fra RCTs spiller en relativt beskeden rolle for psykoterapiens udvikling, hvor andre forskningstyper antageligt er vigtigere (Salkovskis, 2000a;b). Hvis jeg i bakspejlet skulle beskrive den strategi, jeg selv har anvendt i mine bestræbelser for en teoretisk og empirisk indholdsbestemmelse af psykoterapi (Hougaard, 2004), kunne den benævnes som psykoterapi på basis af empirisk, teoretisk og klinisk plausibilitet. Det er efter min vurdering vigtigere for psykoterapiens udvikling, at den baseres på psykologisk viden om psykiske problemer og terapeutiske forandringsprocesser end på teknologiske evidensalgoritmer.

Al forskning må foregå inden for de praktiske muligheds rammer, og evidensbaseret praksis fordrer ikke, at der er evidens for alt, man foretager sig, men blot, at man inddrager de bedste, eksisterende forskningsresultater i beslutningsprocessen. Tilsvarende tages der afstand fra en simpel anvendelse af kogeboogsprincipper i behandlingen. Som det fremhæves i den Amerikanske Psykolog Forenings principper for evidensbaseret psykologisk praksis, må al psykoterapi tilstræbe en effektiv balance mellem konsistens i brugen af interventioner og fleksibilitet under hensyntagen til feedback fra patienten (APA, 2006, p. 276). På denne baggrund mener jeg, det er hensigtsmæssigt at psykologer tilslutter sig principperne for individualiseret evidensbaseret behandling. Det er imidlertid vigtigt at understrege, at der er forskelle på medicinsk og psykologisk behandling og på kravene, der realistisk kan stilles til evidensbaseret psykologisk og medicinsk praksis.

Der er fortsat mange uafklarede problemer vedrørende forskning af relevans for evidensbaseret psykoterapi (APA, 2006), herunder vægtningen af de forskellige forskningsmetoder; fx RCTs i forhold til mindre kontrolleret naturalistisk forskning, effektforskning i forhold til procesforskning, case-baseret forskning i forhold til gruppesammenligninger eller kvantitativ forskning i forhold til kvalitativ forskning (jf. Norcross, Beutler & Levant, 2006). Forskningsmetodologi er, ligesom psykoterapi, et felt i stadig udvikling. De første RCTs i 1950'erne mødte skepsis blandt datidens psykiatere (Healy, 1997), og Smith & Glass' (1977) metaanalyse af psykoterapeutisk effektforskning blev affærdiget som »mega-silliness« (Eysenck, 1978). Det, som er »god latin« i dagens psykoterapiforskning, er ikke nødvendigvis det samme, som vil være det om 50 år.

Der er også mange uafklarede spørgsmål og faldgruber i forbindelse med den praktiske udformning af evidensbaseret psykoterapi (jf. Møhl, 2008). Artiklen her har alene forholdt sig til forskningsgrundlaget for dokumentation af effekt. Vi er kommet langt, siden Eysenck i 1952 konkluderede, at forskningen tyder på, at psykoterapi er virkningsløs.

REFERENCER

- AANONSEN, A.-M. (2007). Evidensbaseret psykologisk praksis. Leder. *Tidsskrift for Norsk Psykologforening*, 44, 802.
- ANTONY, M.M., OSILLO, S.M. & ROEMER, L. (2001). *Practitioner's guide to empirically based measures of anxiety*. New York: Kluwer Academic/Plenum Publishers.
- APA (American Psychological Association) (1995). Template for developing guidelines: Interventions for mental disorders and psychosocial aspects of physical disorders. Albany: American Psychological Association.
- APA (American Psychological Association) (2002). Criteria for evaluating treatment guidelines. *American Psychologist*, 57, 1052-1059.
- APA (American Psychological Association) (2006). American Psychological Association Presidential Task Force on Evidence-based Practice: Evidence-based practice in psychology. *American Psychologist*, 61, 271-285.
- BARLOW, D.H. (2004). Psychological treatments. *American Psychologist*, 59, 869-879.
- BARLOW, D.H., ALLEN, L.B. & CHOATE, M.L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy*, 35, 205-230.
- BECK, A.T., RUSH, A.J., SHAW, B.F. & EMERY, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- BEUTLER, L.E. (1983). *Eclectic psychotherapy: A systematic approach*. New York: Pergamon Press.
- BOHART, A.C., O'HARA, M. & LEITNER, L.M. (1998). Empirically violated treatments: Disenfranchisement of humanistic and other psychotherapies. *Psychotherapy Research*, 8, 141-157.
- BRINKMANN, S. & TANGGAARD, L. (red.) (2007). *Psykologi: Forskning og profession*. København: Reitzel.
- CHAMBLESS, D.L. & HOLLON, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 60, 904-908.
- CHAMBLESS, D.L. & OLLENDICK, T. (2001). Empirically supported psychological interventions: Controversis and evidence. *Annual Review of Psychology*, 52, 685-716.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.
- CONSORT Group (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine*, 134, 663-694.
- CRITS-CHRISTOPH, P., BARANACKIE, K., KURCIAS, J. et al. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research*, 1, 81-91.
- ELKIN, I., SHEA, T., WATKINS, J.T. et al. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry*, 46, 971-982.
- ELKIN, I., FALCONNIER, L., MARTINOVICH, Z. & MAHONEY, C. (2006). Therapist effects in the national institute of mental health treatment of depression collaborative research program. *Psychotherapy Research*, 16, 144-160.

- EYSENCK, H.J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, 16, 319-324.
- EYSENCK, H.J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517, 1978.
- FOA, E.B., LIEBOWITZ, M.R., KOZAK, M.J. et al. (2005). Randomized, placebo-controlled trial of exposure and ritual prevention, clomipramine and their combination in the treatment of obsessive-compulsive disorder. *American Journal of Psychiatry*, 162, 151-161.
- GUPTA, S. & WARNER, J. (2007). A NICE mess. *Journal of the Royal Society of Medicine*, 100, 442-443.
- GUYAT, T. G., CAIRNS, J., CHURCHILL, D. et al. (Evidence-Based Medicine Working Group) (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA*, 268, 2420-2425.
- HAYES, S.C., FOLETTE, V.M. & LINEHAN, M.M. (eds.) (2004). *Mindfulness and acceptance: Expanding the cognitive-behavioural tradition*. New York: Guilford.
- HEALY, D. (1997). *The antidepressant era*. Cambridge, MA: Harvard University Press.
- HIGGINS, J.P., THOMPSON, S.G., DEEKS, J.J. et al. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557-560.
- HOUGAARD, E. (2003). Evidensbaseret psykoterapi: Baggrund, problemer og implementeringsformer. *Matrix: Nordisk Tidsskrift for Psykoterapi*, 20, 3-26.
- HOUGAARD, E. (2004). *Psykoterapi: Teori og Forskning*. København: Dansk psykologisk Forlag.
- HOUGAARD, E. (2007a). Evidensbaseret psykoterapi: Hvordan foregår det i praksis? I: S. Brinkmann & L. Tanggaard (red.), *Psykologi: Forsknings og profession*. København: Reitzel.
- HOUGAARD, E. (2007b). Evidens: Noget for psykologer? *Psykolog Nyt*, 61 (20), 14-21.
- HUBBLE, M.A., MILLER, S.D. & DUNCAN, B.L. (1999). *The heart and soul of change: What works in therapy*. New York: APA.
- INGRAM, R.E., HAYES, A. & SCOTT, W. (2000). Empirically supported treatments: A conceptual analysis. I: C.R. Snyder & R.E. Ingram (eds.), *Handbook of psychological change: Psychotherapy processes and practices for the 21st century* (pp. 40-60). New York: Wiley.
- JACOBSON, N.S., MARTELL, C.R. & DIMIDJIAN, S. (2002). Behavioral activation treatment for depression: Returning to contextual roots. *Clinical psychology: Science and Practice*, 8, 255-270.
- JONSSON, H. & HOUGAARD, E. (submitted). Outcome of Group Cognitive Behavioural Therapy for Obsessive Compulsive Disorder (OCD): A systematic review and meta-analysis.
- KAZDIN, A.E. (2006). Arbitrary metrics: Implications for identifying evidence-based treatments. *American Psychologist*, 61, 2006, 42-49.
- KAZDIN, A.E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, 3, 1-27.
- KIM, D-M., WAMPOLD, B.E., BOLT, D.M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research*, 16, 161-172.
- LAMBERT, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, 17, 1-14.

- LAMBERT, M.J. & OGLES, B.M. (2004). The efficacy and effectiveness of psychotherapy. I: M.J. Lambert (ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.) (pp. 139-193). New York: Wiley.
- LEDoux, J.E. (1996). *The emotional brain: The mysteriously underpinnings of emotional life*. New York: Simon & Schuster.
- LONGMORE, R.J. & WORRELL, M. (2006). Do we need to challenge thoughts in cognitive behaviour therapy? *Clinical Psychology Review*, 27, 173-187.
- MCCULLOUGH, J.P. (2000). *Treatment for chronic depression: Cognitive-behavioral analysis system of psychotherapy*. New York: Guilford.
- MITTE, K. (2005a). A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *Journal of Affective Disorders*, 88, 27-45.
- MITTE, K. (2005b). Meta-analysis of cognitive-behavioral treatments for generalized anxiety disorder: A comparison with pharmacotherapy. *Psychological Bulletin*, 131, 785-795.
- MONCRIEFF, J., CHURCHILL, R., DRUMMOND, D.C. & MCGUIRE, H. (2001). Development of a quality assessment instrument for trials of treatments for depression and neurosis. *International Journal of Methods in Psychiatric Research*, 10, 126-133.
- MØHL, B. (2008). Alternativer til evidensbaseret klinisk praksis. *Psyke og Logos*, dette nr.
- NICE (National Institute of Clinical Excellence) (2004a). Clinical guideline for the management of anxiety: Management of anxiety (panic disorder, with or without agoraphobia, and generalised anxiety disorder) in adults in primary, secondary and community care. National Clinical Practice Guideline Number 22. London: National Institute for Clinical Excellence (www.nice.org.uk).
- NICE (National Institute of Clinical Excellence) (2004b). Depression: Management of depression in primary and secondary care. National Clinical Practice Guideline Number 23. London: National Institute for Clinical Excellence (www.nice.org.uk).
- NICE (National Institute of Clinical Excellence) (2005). Post-traumatic stress disorder: The management of PTSD in adults and children in primary and secondary care. National Clinical Practice Guideline Number 26. London: National Institute for Clinical Excellence (www.nice.org.uk).
- NICE (National Institute of Clinical Excellence) (2006). Obsessive-compulsive disorder: Core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder. National Clinical Practice Guideline Number 31. London: National Institute for Clinical Excellence (www.nice.org.uk).
- NORCROSS, J.C., BEUTLER, L.E. & LEVANT, R.F. (2006). *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions*. Washington, DC: American Psychological Association.
- PRENTZAS, J. & HATZILYGEROUDIS, I. (2007). Categorizing approaches combining rule-based and case-based reasoning. *Expert Systems*, 24, 97-121.
- ROSEN, G.M. & DAVISON, G.C. (2003). Psychology should list empirically supported principles of change (ESP's) and not credential trademarked therapies or other treatment packages. *Behavior Modification*, 27, 300-312.
- ROSENTHAL, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills: Sage.
- SACKETT, D.L., ROSENBERG, W.M.C., GRAY, J.A.M. et al. (1996). Evidence based medicine: What it is and what it isn't. *BMJ*, 312, 71-72.
- SACKETT, D.L., STRAUS, S.E., RICHARDSON, W.S. et al. (2001). *Evidence-based medicine: How to practice and teach EBM*. London: Churchill Livingstone.
- SALKOVSKIS, P. (2002a). Empirically grounded clinical interventions. *Behavioural and Cognitive Psychotherapy*, 30, 1-2.

- SALKOVSKIS, P. (2002b). Empirically grounded clinical interventions: Cognitive-behavioural therapy progresses through a multi-dimensional approach to clinical science. *Behavioural and Cognitive Psychotherapy*, 30, 3-9.
- SBU (Statens Beredning För Medicinsk Utvärdering) (2004). *Behandling av depressionssjukdomar: En systematisk litteraturoversikt, Vol. 1-3*. Stockholm: Statens Beredning För Medicinsk Utvärdering, Publ. Nr. 166/1-3. (www.sbu.se).
- SBU (Statens Beredning För Medicinsk Utvärdering) (2005). *Behandling av ångestsyndrom: En systematisk litteraturoversikt, Vol. 1-2*. Stockholm: Statens Beredning För Medicinsk Utvärdering, Publ. Nr. 171/1-2 (www.sbu.se).
- SCHÖN, D.A. (1983). *The reflective practioner: How professionals think in action*. New York: Basic Books.
- SELIGMAN, M.E.P. (1996). Science as an ally of practice. *American Psychologist*, 51, 1072-1079.
- SHADISH, W.R., COOK, T.D. & CAMPBELL, D.T. (2002). *Experimental and quasi-experimental design for causal inference*. New York: Houghton Mifflin.
- SMITH, M.L. & GLASS, G.V (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.
- SPIELMANS, G.I, PASEK, L.F. & MCFALL, J.P. (2007). What are the active ingredients in cognitive and behavioral psychotherapy for anxious and depressed children: A meta-analytic review. *Clinical Psychology Review*, 17, 642-654.
- SPRING, B. (2007). Evidence-based practice in clinical psychology: What it is; why it matters; what you need to know. *Journal of Clinical Psychology*, 63, 611-631.
- SPRING, B., PAGOTO, S., KNATTERUD, G. et al. (2007). Examination of the analytic quality of behavioral health randomized clinical trials. *Journal of Clinical Psychology*, 63, 53-71.
- STRUPP, H.H., HOROWITZ, L.M. & LAMBERT, M.J. (eds.) (1997). *Measuring patient changes in mood, anxiety and personality disorders: Toward a core battery*. Washington: American Psychological Association.
- SUNDHEDSSTYRELSEN (2007a). Referenceprogram for angstlidelser hos voksne. København Sundhedsstyrelsen. (www.sst.dk).
- SUNDHEDSSTYRELSEN (2007b). Referenceprogram for unipolar depression hos voksne. København Sundhedsstyrelsen. (www.sst.dk).
- TASK FORCE ON EMPIRICALLY VALIDATED TREATMENTS (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *Clinical Psychologist*, 48, 3-23.
- WANG, W.M., CHEUNG, C.F., LEE, W.B. & KWOK, S.K. (2007). Knowledge-based treatment planning for adolescent early intervention of mental healthcare: A hybrid case-based reasoning approach. *Expert Systems*, 24, 232-251.
- WAMPOLD, B.E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.
- WAMPOLD, B.E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist*, 63, 857-873.
- WEINMANN, S., KOESTERS, M. & BECKER, T. (2007). Effects of implementation of psychiatric guidelines on provider performance and patient outcome: Systematic Review. *Acta Psychiatrica Scandinavica*, 115, 420-433.
- WENNEBERG, S. (1999). Forskningsanvendelse: Hvad er det? Papir præsenteret på den 20. Nordiske Sociologikonference i Bergen.
- WESTEN, D., NOVOTNY, C.M. & THOMPSON-BRENNER, H. (2004). Empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631-663.
- WIERZBICKI, M., & PEKARIK, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice*, 24, 190-195.