

Vetenskaplig kvalitet i publicerade effektutvärderingar

**169 svenska utvärderingar av psykologiska
och sociala interventioner**

Du får gärna citera Socialstyrelsens texter om du uppger källan, exempelvis i utbildningsmaterial till självkostnadspris, men du får inte använda texterna i kommersiella sammanhang. Socialstyrelsen har ensamrätt att bestämma hur detta verk får användas, enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk (upphovsrättslagen). Även bilder, fotografier och illustrationer är skyddade av upphovsrätten, och du måste ha upphovsmannens tillstånd för att använda dem.

ISBN 978-91-7555-184-5
Artikelnummer 2014-5-18

Publicerad www.socialstyrelsen.se, maj 2014

Förord

En viktig del i en evidensbaserad praktik är kunskap om hur effektiva olika interventioner är. Inom Socialtjänstområdet är denna kunskap svag på grund av bristen på kvalificerad forskning. Det försvårar för Socialstyrelsen att ge välgrundad vägledning till professionella. I denna rapport beskrivs den vetenskapliga kvaliteten i svensk interventionsforskning.

Rapporten riktar sig till beslutsfattare inom myndigheter och departement.

Författare till rapporten är Knut Sundell och Elizabeth Åhsberg, Socialstyrelsen. Maria Bodin och Johan Glad, Socialstyrelsen, samt Anders Tengström, Karolinska institutet, har medverkat i granskning av underlaget.

Sven Ohlman
avdelningschef
avdelning för kunskapsstyrning

Innehåll

Förord	3
Sammanfattning	7
Summary	8
Bakgrund	9
Syfte	10
Metod	11
Sökprocedur	11
Urval av publikationer	11
Granskning av artiklar	12
Resultat	18
Allmän beskrivning av artiklarna	18
Kvalitetskriterier i artiklarna	18
Redovisat resultat	21
Samlad bild av forskningskvalitet	21
Koppling forskningskvalitet och resultat	23
Diskussion	24
Förklaring till bristande kvalitet	24
Ökad forskarkompetens	26
Tillförlitlighet	26
Referenser	28
Bilaga 1. Granskade artiklar	29
Bilaga 2. Kvalitetsaspekter fördelat på publikationsår (%)	39
Bilaga 3. Kvalitetsaspekter fördelat på startår (%)	40

Sammanfattning

En evidensbaserad praktik baseras på en sammanvägning av brukarens situation, erfarenhet och önskemål, den professionelles expertis samt bästa tillgängliga kunskap om interventioners effekter. Hur informationen från dessa källor vägs samman bestäms av det nationella och lokala sammanhanget [1]. Utan kunskap om interventioners effekter minskar möjligheten att hjälpa klienten och innebörden av en evidensbaserad praktik urholkas. Det säkraste underlaget erhålls från utvärderingar där olika interventioner jämförs, så kallade effektutvärderingar. Syftet med den här rapporten är att identifiera förbättringsområden genom att granska den vetenskapliga kvaliteten i svenska effektutvärderingar av psykologiska och sociala interventioner.

Undersökningen omfattar så vitt känt samtliga 169 svenska effektutvärderingar som publicerats i vetenskapliga tidskrifter med kollegial granskning (peer-review) till och med april 2012. Resultaten bygger på den information som finns i artiklarna.

- Granskningen beskriver en stor variation i rapporterad vetenskaplig kvalitet. Det visar att publicering i en tidskrift med kollegial granskning inte garanterar hög vetenskaplig kvalitet.
- Den vetenskapliga kvaliteten ökade över tid. Det finns en risk att främst äldre utvärderingar redovisar snedvridna resultat som beror på kvalitetsbrister.
- Även nyare artiklar har delvis brister. Det talar för vikten av att förstärka utbildningen för forskare om hur effektutvärderingar bör genomföras och rapporteras.
- Knappt hälften (44 %) av artiklarna redovisar inte någon statistiskt säkerställda skillnader för det primära utfallsmåttet. Det visar på betydelsen av att utvärdera nya interventioner och att inte enbart basera val av insats på teoretiska grunder.
- Bristen på utvärderingar är påtaglig inom individ- och familjeomsorg, skola, äldreomsorg samt försäkringskassa och kriminalvård.

Summary

An evidence-based practice includes an appraisal of: (i) the patient's situation, experience and preferences; (ii) the professional's expertise, and; (iii) the best available research evidence concerning the effectiveness of potential interventions. Both national- and local contexts determine how these different sources of knowledge can and should contribute to such an appraisal. Without scientific knowledge about the effects of interventions, the potential to help the client is reduced and one of the central tenets of evidence-based practice is not met. The aim of this report is to investigate the scientific quality of Swedish efficacy- and effectiveness evaluations of psychological- and social interventions. The results will then be used to identify areas where potential improvement can be made.

To the best of our knowledge, this study includes all Swedish efficacy- and effectiveness evaluations published in scientific peer reviewed journals through April 2012. The results are based on the information reported in the published articles.

- There was substantial variation in scientific quality between studies. This indicates that publishing in a scientific peer reviewed journal does not guarantee scientific quality.
- The scientific quality of reviewed evaluations increased over the time-period examined. There is therefore a risk that the reported results of older evaluations may be biased due to lower scientific quality.
- More recent evaluations, however, still exhibited several areas of concern regarding their scientific quality. This highlights the importance of further training for researchers in the latest scientific methods in order to enable them to conduct high-quality efficacy- and effectiveness evaluations.
- Approximately half (44 %) of the articles did not reported statistically significant effects of the interventions on the primary reported outcome measure. This emphasizes the importance of evaluating new interventions before they are widely implemented rather than basing the choice of intervention solely on theoretical considerations.
- Areas where a lack of published evaluations is noticeable are: social work, education, elder care, corrections and social insurance.

Bakgrund

I Sverige får årligen ungefär 800 000 personer insatser inom ramen för socialtjänsten och 630 000 inom psykiatrin [2]. Det handlar i vissa fall om interventioner som inte är frivilliga och där målgruppen utgörs av samhällets mest utsatta. Få av dessa interventioner är utvärderade i Sverige. Exempelvis har Socialstyrelsen identifierat 103 socialtjänstbaserade öppenvårdsinsatser för att stödja barns psykiska hälsa, varav endast en handfull har ett visst vetenskapligt stöd [3]. Bristen på forskningsstöd är inte ett trivialt problem eftersom det finns exempel på psykologiska och sociala interventioner som orsakat skada [4-6]. Det bör tilläggas att de insatser som inte är utvärderade kan vara effektiva, men att vi inte vet om så är fallet.

Begreppet evidensbaserad medicin myntades i mitten av 1990-talet av den kanadensiske folkhälsoforskaren David Sackett med kollegor. Det skedde som en reaktion på att mycket av den kliniska medicinen inte grundades på de senaste vetenskapliga rönen och att det var ett auktoritärt system där läkarna ensam bestämde [7]. En följd av detta var att patienterna inte fick bästa tänkbara vård eller till och med att skadliga insatser användes.

Evidensbaserad praktik har definierats som ”en noggrann, öppet redovisad och omdömesgill användning av den för tillfället bästa evidensen för beslutsfattande om åtgärder (insatser, metoder) till enskilda personer, kompletterad med professionell expertis och den berörda personens situation och önskemål” [5]. Evidensbaserad praktik kan beskrivas i form av fyra överlappande cirklar (figur 1), där den professionella kompetensen integrerar de olika kunskapskällorna i dialog med den berörda personen (patienten, klienten) och där beslutsgrunderna redovisas öppet [6]. Det är i skärningsfältet mellan dessa fyra informationskällor som den evidensbaserade praktiken finns.

Figur 1. Evidensbaserad praktik [7]



Utan vetenskaplig kunskap om interventioners effekter minskar möjligheterna att hjälpa klienter och innebörden av en evidensbaserad praktik urholkas. Det är framför allt två typer av forskning som ger tillförliga resultat om effekter, så kallade effektutvärderingar: randomiserat kontrollerade studier (RCT) och icke-randomiserade studier som inkluderar jämförelsegrupp och för- respektive eftermätning.

Syfte

Syftet med den här rapporten är att granska den vetenskapliga kvaliteten i svenska effektutvärderingar av psykologiska och sociala interventioner som avrapporterats i vetenskapliga tidskrifter med kollegial granskning (peer-review-förfarande) till och med april 2012. Genom att systematisera denna forskning kan förbättringsområden identifieras.

Metod

Sökprocedur

Sökningen efter vetenskapliga utvärderingar var relativt omfattande eftersom det saknas ett nationellt register för svenska vetenskapliga utvärderingar.

Nedanstående sökprocedurer användes [8]:

- Sökord (på både svenska och engelska) var randomiserad kontrollerad, effekt, utvärdering, implementering, RCT, evidens.
- Sökning gjordes på hemsidor till åtta vetenskapliga discipliner hos samtliga svenska universitet och högskolor: folkhälsovetenskap, kriminologi, omvårdnadsvetenskap, pedagogik, psykiatri, psykologi, socialt arbete och sociologi. Granskningen har omfattat dels publicerade, dels pågående studier.
- Sökning efter projekt som beviljats medel av Forskningsrådet för arbetsliv och socialvetenskap, Forskningsrådet Formas, ERC Europeiska Forskningsrådet, Vetenskapsrådet samt VINNOVA (samtliga 2005-2012), Brottsofferfonden (2003 – 2011).
- Dessutom gjordes sökningar på följande hemsidor:
 - www.avhandlingar.se
 - www.cbti.se (Svenska institutet för kognitiv psykoterapi)
 - www.clinicaltrials.gov (USAs National Institute of Health öppna register över kliniska studier)
 - www.diva-portal.org/smash/search.jsf
 - www.evidens4u.se/
 - www.ncbi.nlm.nih.gov/pubmed
 - www.researchweb.org (tidigare www.fou.nu)
 - www.sanktlukas.se (Förbundet S:t Lukas utbildningsinstitut)
 - www.sapu.se (Stockholms akademi för psykoterapiutbildning)
 - www.sbu.se (Statens beredning för medicinsk utvärdering)
 - www.sofie8.se
 - www.studie.nu
 - www.uppsatser.se
 - www.orebroll.se/sv/Forskning/Forskningsomraden/PFC
- Utvärderingar har även identifierats genom tips från kontaktade forskare som själva genomfört eller genomför effektutvärderingar.

Urval av publikationer

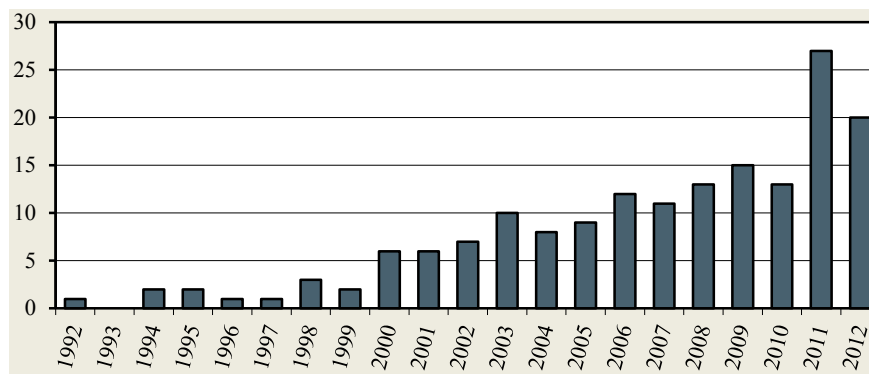
Totalt identifierades 169 utvärderingar, enligt följande kriterier:

- Utvärderingen är publicerad i en vetenskaplig tidskrift med kollegial (peer-review) granskning.
- Det är en randomiserad eller icke-randomiserad utvärdering där en viss intervention har implementerats och testats i försöksliknande eller ordinarie verksamhet.

- Interventionen har utvärderats med minst en förbestämjd jämförelsegrupp.
- De som varit föremål för interventionen har undersökts före och efter att interventionen getts. Det innebar minst två mättillfällen, där det andra tillfället skett efter att interventionen avslutats.
- Utvärderingen omfattar psykologiska eller sociala interventioner (ev. i kombination med medicinska eller farmakologiska). Psykologiska interventioner avser användning av verbal eller ickeverbal kommunikation för att behandla olika typer av känslomässiga, mentala eller beteendemässiga problem. Sociala insatser syftar till socialt eller materiellt stöd, exempelvis stödfamilj och stödgrupper för anhöriga.
- Interventionen är riktad till enskilda eller grupper av klienter med syfte att förbättra deltagarnas psykiska eller fysiska hälsa.
- Utvärderingen omfattar personer i Sverige och den vetenskapligt ansvarige tillhör ett svenskt universitet, högskola eller annan organisation verksam i Sverige.

I de fall flera artiklar publicerats från samma utvärdering har den valts som omfattar den längsta uppföljningstiden¹. De 169 artiklarna är publicerade mellan 1992 och april 2012 (se bilaga 1)². Fördelningen av artiklarna över tid redovisas i figur 2. Det kan noteras att 2012 endast omfattar artiklar som var publicerade under fyra månader.

Figur 2. Publikationsår för effektutvärderingarna till och med april 2012 (n=169)



Granskning av artiklar

I denna rapport avser vetenskaplig kvalitet hur tillförlitliga de resultat är som redovisas. Hög vetenskaplig kvalitet innebär att sannolikheten är låg att resultaten beror på slumpmässiga eller systematiska fel.

De konkreta kriterierna för granskningen fastställdes i en frågemall. Utgångspunkten var CONSORT [9], TREND [10] och Prevention Science [11]

¹ För 45 utvärderingar fanns det minst två publicerade artiklar.

² Några av artiklarna var elektroniskt publicerade i april 2012 och har senare kommit i tryckt skick men då med året 2013.

som kan sägas vara vetenskapssamhällets bedömning av kvalitetskriterier i vetenskapliga studier och dess vidhängande publikationer. Det slutgiltiga valet av kvalitetskriterier utgör en avvägning av vad som var nödvändigt för att kunna värdera genomförande och analys och vad som är möjligt att utläsa ur en artikel utan att konsultera författarna i efterhand. Det senare bedömdes inte som meningsfullt, bland annat på grund av risken att författarna inte längre kommer ihåg detaljer (den äldsta utvärderingen startade 1974).

Varje artikel granskades enskilt av två forskare. Vid oenighet diskuterades varje artikel till konsensus erhöles. Svaren på flertalet frågor i granskningsmallen kategoriserades som antingen ja (den efterfrågade informationen fanns), nej (den efterfrågade informationen fanns inte), ej tillämpligt och osäker (där det var oklart för båda granskarna om den efterfrågade informationen fanns redovisad). Svaret osäker förekom sällan (i genomsnitt 2,3 gånger per kriterium) och omkodades som att information saknades.

Utvärderingarna har bedömts utifrån den information som fanns rapporterad i den vetenskapliga artikeln. Författarna har inte kontaktats för kompletterande information. Nedan presenteras de områden som kodats.

Transparens

Undersökningsprotokoll publicerat. Avser om undersökningsprotokollet publicerats i förväg. I det anges exempelvis vilken målgruppen är, vilka studiealternativ som jämförs och vad som är primärt utfallsmått. Genom att i förväg publicera protokollet motverkas exempelvis att sekundära utfallsmått tillmäts för stor betydelse vid avrapportering.

Redovisning av eventuella intressekonflikter. Beskrivning av om det funnits ekonomiska eller andra intressekonflikter.

Undersökningsgrupp

Inklusions- och exklusionskriterier. Rapporteras vilka urvalskriterier som använts, t.ex. individer, kliniker eller verksamheter. Bedömningen avsåg den nivå man randomiserat/fördelat på.

Demografiska och kliniska variabler. Avser om både demografiska och kliniska uppgifter redovisats, separat för undersökningsgrupperna.

Statistisk styrka (eng. statistical power). Vid planeringen av en studie görs en beräkning som baseras på styrkevärdet, önskad signifikansnivå, och den förväntade effektstorleken. Beräkningen visar hur många deltagare studien behöver för man ska hitta en statistiskt säkerställd effekt givet att en sådan finns.

Datainsamling

Primärt utfallsmått. Avser det mått som författarna angett som primärt för studien, eller det mått som motsvarade syftet i studien. Identifierades genom att i artikeln söka efter termen “primary (outcome)”.

Om inget primärt utfallsmått angivits valdes det först beskrivna utfallsmåttet i artikelns metodavsnitt som utgångspunkt för fortsatta bedömningar om psykometriska egenskaper mm. Om det först nämnda måttet uppenbart var mindre viktigt i relation till populationen/interventionens mål och det finns ett som var tydligt mer relevant användes det senare för fortsatt granskning.

Psykometriska egenskaper hos eventuella bedömningsinstrument. Psykometriska egenskaper kunde t.ex. vara Cronbachs alfa för de studerade grupperna, eller ett generellt uttalande om reliabilitet/validitet samt hänvisning till en artikel som beskriver instrumentets psykometriska egenskaper. Psykometriska uppgifter visar om frågorna i ett instrument är tillförlitliga och adekvata för den målgrupp som avses. Om det primära utfallsmåttet utgjort av registerdata har det kodats som att psykometriska egenskaper saknats.

Om datainsamlare var ovetande om deltagarnas grupptillhörighet. Här avses om datainsamlarna var ovetande (blinda) för deltagarnas grupptillhörighet. Om datainsamlingen genomförts i form av deltagarnas självrapporteringar har det kodats som att ”oblinda”.

Intervention

Omfattning som interventionen implementerats. Finns någon beskrivning av vad som anses utgöra ”treatment completion” och rapporterat antal ”completers”, eller beskrivit antal sessioner etc. Till exempel kan det ha angetts att 76 procent genomfört interventionen som tänkt och att 24 procent avbrutit behandlingen.

Design

Randomisering. Bedömdes genom sökning efter termen “random” eller om det i texten tydligt beskrev att undersökningspersonerna slumpmässigt fördelats (randomiserats) till olika betingelser.

Randomiseringssekvens beskriven (endast randomiserade studier). Avser om randomiseringssekvensen genererats via t.ex. slumptalsgenerator eller lottdragning.

Typ av randomisering beskriven (endast randomiserade studier). Rapporteras vilken typ av randomisering som använts, dvs. enkel randomisering eller randomisering med någon typ av restriktion (t.ex., blockning, stratifiering, minimization) [9].

Dold fördelning till betingelser (endast randomiserade studier). Rapporteras om man förhindrat att de som fördelade personer till betingelser kunde för-

utse vem som skulle hamna i respektive grupp (t.ex. genom att använda personer som var oberoende av projektet, eller i ordningsföljd nummerade, ogenomskinliga kuvert vid lottning).

Fördelningsprincip beskriven (endast icke-randomiserade studier). Rapporteras principen för att fördela deltagare till studiegrupperna? (t.ex. självselektion till grupp, varannan person/klinik, fyller upp interventionsgruppen först sedan kontrollgruppen).

Försök att minska selektionsbias (endast icke-randomiserade studier). Rapporteras försök att minska risken för selektionsbias genom exempelvis matchning?

Analys

Antal personer som fördelats till respektive grupp. Avser om det är beskrivet i flödesschema (flow-chart) eller i text hur många deltagare som fördelades till varje studiegrupp? Om studien har fördelat individerna till kluster (t.ex. skolor) krävs att både antal kluster och antal individer anges per grupp.

Antal personer som föll bort per grupp. Bortfallet avser de som av olika skäl inte fullföljt en insats, eller som avbrutit medverkan i en utvärdering genom att inte medverka i en eftermätning.

Antal personer som analyserats per grupp. Avser det antal som inkluderades i analysen av det primära utfallet.

Intention to treat. ITT-analys innebär att man vid analys av data från en effektstudie (a) inkluderar samtliga deltagare som fördelats till grupperna (b) låter deltagarna behålla sin gruppstillhörighet oavsett om de avbrutit interventionen eller fått annan intervention än den avsedda, samt (c) baserar analyserna på samtliga deltagare. Eftersom bortfall är vanligt i longitudinella studier kräver det sista kriteriet oftast någon form av imputering, dvs. uppskattning av de värden som saknas. ITT-analys anses vara att föredra framför allt i effektutvärderingar. Frågan besvarades genom att i artikeln söka efter ordet ”intention-to-treat”, ”intent-to-treat” och ”ITT”.

Imputering. Att i en statistisk undersökning ersätta värden som saknas med uppskattade värden. Det undersöktes genom att i artikeln söka efter ordet ”input”.

Statistisk justering för eventuella baslinjeskillnader. Avser de skillnader som kunde finnas mellan undersökningsgrupperna vid förmätning. Om sådana skillnader fanns, undersöktes om statistiska metoder använts för att justera för dessa skillnader när interventionens effekt på det primära utfallet beräknades. Om inga skillnader förelåg vid förmätning har artikeln kodats som att justering för baslinjeskillnader fanns.

Undersökt skillnader mellan individer som avbrutit medverkan och de som fullföljt (totalt). Rapporteras statistiska jämförelser av baslinjedata för de som fullföljt en studie och de avbrutit medverkan för samtliga deltagare som helhet (t.ex. hade de som fullföljt mindre problem vid förmätning)?

Undersökt skillnader mellan individer som avbrutit medverkan och de som fullföljt (per grupp). Rapporteras statistiska jämförelser av baslinjedata för de som fullföljt en studie och de avbrutit medverkan för samplet för respektive studiegrupp?

Statistisk justering vid eventuell klusterrandomisering. Avser de studier där grupper av individer randomiserats eller fördelats till olika betingelser och där individdata analyserats. Justering för problemet med korrelerade data inom klustren kunde ha gjorts med hjälp av t.ex. intraklass-korrelation, flernivå regressionsanalys, nästads Anova, hierarkisk linjär modellering eller struktur ekvations modellering. Detta har endast kodats när det i texten har angetts att det funnits kluster och inte exempelvis om flera personer haft samma behandlare.

Resultat

Beskrivande statistik. Avser om t.ex. medelvärde, standardavvikelse eller antal händelser per grupp redovisades för det primära utfallsmåttet för samtliga mätillfällen. Denna typ av information är viktig dels för resultatens generaliserbarhet, dels för att resultat från en studie ska kunna ingå i en metaanalys, dvs. sammanställas med andra liknande studier.

Statistisk signifikans. Avser om det fanns en statistiskt säkerställd skillnad mellan studiegrupperna vid eftermätningen i det primära utfallsmåttet. Om flera mätningar gjorts valdes resultatet från den sista mätningen. Om analysen bygger på en ANOVA med upprepad mätning avses om interaktionen grupp gånger tid är säkerställd.

Effektstorlek. Avser om någon effektstorlek rapporterades, t.ex. Cohens d , odds kvot eller risk kvot. Med effektstorlek menas här att skillnader mellan undersökningsgrupperna har prövats, inte eventuella skillnader över tid för en enskild grupp.

Där medelvärden och standardavvikelser fanns rapporterade för det primära utfallsmåttet beräknades effektstorlek via skillnaden i medelvärde mellan interventionsgruppen och kontrollgruppen, dividerat med ett vägt medelvärde av gruppernas standardavvikelse [12]. För andra typer av information om effekter (t.ex., odds- eller riskkvoter, frekvenser, F - och t -värden) genomfördes omräkning till Cohens d [12]. För studier med flera undersökningsgrupper³ användes den med minst påverkan på undersökningspersoneran (t.ex., väntelista). Om det rapporteras effektstorlekar från flera uppföljningsmätningar, avses den sista mätningen.

³ Det förekom tre grupper i 18utvärderingar och fyra grupper i åtta.

Eventuella oönskade utfall. Avser om forskarna beskriver om insatsen haft oönskade utfall eller medfört komplikationer för deltagarna. Att beskriva eventuella risker med en insats är viktigt för att kunna förstå hur ändamålsenlig insatsen är.

Resultat

Allmän beskrivning av artiklarna

De granskade artiklarna inkluderar interventioner från en rad verksamhetsområden: psykiatri ($n = 64$), hälso- och sjukvård ($n = 43$), missbruk och beroende ($n = 31$), individ- och familjeomsorg ($n = 11$), skola ($n = 9$), äldreomsorg ($n = 8$), kriminalvård ($n = 2$) samt försäkringskassa ($n = 1$).

Bland de interventioner som studerades finns t.ex. kognitiv beteendeterapi, familjeterapi, utbildning, akupunktur och motiverande samtal. Bland utfallsmåtten finns de som beskriver ångest, fobi, livskvalitet, hälsa, depression, stress, smärta, sömn, kriminalitet, narkotika-, alkohol-, och cigarettanvändning, sjukfrånvaro, demenssymptom, fallrisk, ADL-förmåga och mortalitet.

Aspekter som inte direkt kan relateras till utvärderingarnas vetenskapliga kvalitet är att 65 procent av utvärderingarna utgör modellutvärderingar och övriga 35 procent utvärderingar i ordinarie verksamhet. De flesta hade en jämförelsegrupp där personerna fick en annan intervention (66 %), men det förekom också väntelista eller ingen intervention (34 %). Majoriteten (60 %) har en uppföljningstid som var minst 12 månader efter förmätning. Av effektutvärderingarna behandlar 35 procent interventioner som utvecklats i ett annat land än Sverige. Utvärderingarna har hel- eller delfinansierats av något forskningsråd i 54 procent av fallen. Övriga finansiärer var kommuner och landsting, statliga myndigheter och privata stiftelser eller fonder.

Kvalitetskriterier i artiklarna

Nedan beskrivs resultaten som även presenteras i tabell 1. I bilaga 2-3 redovisas hur förekomsten av kvalitetskriterierna förändrats över tid.

Transparens

För att öka transparens är det viktigt att publicera undersökningsprotokollet i förväg. Endast var tionde hade gjort det (i huvudsak ClinicalTrials.gov⁴).

Rapport om eventuella intressekonflikter fanns i 48 av 169 (28 %) artiklar. Av de 48 artiklarna rapporterade sju (15 %) att en intressekonflikt förekommit.

Undersökningsgrupp

I majoriteten (92 %) av artiklarna beskrivs vilka deltagare som valts ut med hjälp av inklusions- och exklusionskriterier.

Något mindre vanligt var att beskriva sociodemografiska och kliniska variabler för undersökningsgruppen (68 %).

För att ha rimlig möjlighet att upptäcka statistisk säkerställda skillnader måste undersökningsgrupperna ha en viss storlek, givet bland annat förvän-

⁴ ClinicalTrial.gov etablerades 2000.

tad effekt. Det var mindre vanligt att det beräknats före studien startade och förekom endast i knappt var fjärde studie.

Tabell 1. Förekomst av rapporterade kvalitetsaspekter

Kvalitetsaspekt	%	N	Ingår i index
<i>Transparens</i>			
Forskningsprotokoll publicerat	10	169	*
Redovisning av ev. intressekonflikt	28	169	*
<i>Undersökningsgrupp</i>			
Inklusions- och exklusionskriterier	92	169	*
Demografiska och kliniska variabler	67	169	*
Beräkning av gruppstorlek för att kunna påvisa skillnader (power)	24	169	*
<i>Datainsamling</i>			
Primärt utfallsmått	36	169	*
Psykometri för bedömningsinstrument	60	169 ^a	*
"Blindad" bedömare	21	169 ^b	*
<i>Intervention</i>			
Omfattning av intervention	62	169	*
<i>Design</i>			
Randomiserad	88	169	*
<i>Vid randomisering</i>			
Randomiseringssekvens	54	149	*
Typ av randomisering	46	149	
Dold fördelning till betingelser	48	149	
<i>Vid icke-randomisering</i>			
Fördelningsprincip	50	20	*
Försök att minska snedvridning (f.ex., via matchning)	50	20	
<i>Analys</i>			
Antal fördelade personer per grupp	86	169	
Antal bortfall per grupp	78	169	
Antal analyserade personer	88	169	
Intent-to-treat (undersökningsgrupper intakta)	56	169	*
Imputering	15	169	
Justerat statistiskt för initiala baslinjeskillnader	66	169	
Undersökt skillnader mellan bortfall och kvarvarande (totalt)	25	169 ^c	*
Undersökt skillnader mellan bortfall och kvarvarande (per grupp)	4	169 ^c	
Statistisk justering vid klusterrandomisering	7	169 ^d	
<i>Resultat</i>			
Beskrivande statistik	77	169	*
Signifikant effekt på primärt utfall	56	169	
Effektstorlek	46	169	*
Eventuella oönskade utfall	14	169	*

^a i 19 studier var det primära utfallsmåttet inte mätt med något bedömningsinstrument

^b i 97 studier var självskattning det primära utfallsmåttet.

^c i 73 studier redovisades att inga initiala skillnader förelåg mellan studiegrupperna.

^d i 142 studier förekom inga tydliga kluster (f.ex., skolklasser).

Datainsamling

Endast i drygt var tredje artikel redovisades vilket utfallsmått som var primärt. Att ett sådant tydligt anges i artikeln anses viktigt för att motverka att mindre centrala utfallsmått tillmäts en för stor betydelse på bekostnad av det som interventionen främst haft för avsikt att påverka.

I 60 % fallen beskrivs de psykometriska egenskaperna för bedömningsinstrument som använts för att mäta det primära utfallsmåttet.

Om möjligt bör datainsamlare sakna förkunskap om vilken grupp som får interventionen och vilken som utgör jämförelsegrupp, för att inte (o)medvetet påverka svaren i förväntad riktning (detta gäller framförallt i intervjusituationer). Inom samhällsvetenskaplig forskning kan detta ibland vara svårt att få till en forskningsdesign som möjliggör ett sådant förfarings sätt. I denna genomgång förekom så kallad ”blindning” i knappt vart fjärde fall.

Intervention

Två av tre artiklar redovisade information om i vilken utsträckning som undersökningspersonerna faktiskt erhållit interventionen.

Design

Av de 169 vetenskapliga studierna var 149 (88 %) genomförda som randomiserade och 20 som icke-randomiserade studier. Med randomisering ges en bättre möjlighet att avläsa effekten av interventionen eftersom den slumpmässiga fördelningen i teorin gör undersökningsgrupperna lika vid förmätningen.

I de 149 randomiserade utvärderingarna finns en beskrivning av hur randomiseringssekvens genererats i ungefär hälften av artiklarna. Typ av randomisering (t.ex., t.ex., blockning, stratifiering) är ungefär lika ofta beskrivet, liksom om fördelning till betingelserna varit dold för den som randomiserat.

I de 20 icke-randomiserade utvärderingarna finns redovisat i exakt hälften av fallen hur deltagarna fördelats till de studerade interventionerna. I hälften av artiklarna redovisas försök att statistiskt justera för olikheter vid förmätningen. Det senare är särskilt viktigt i icke-randomiserade utvärderingar där det kan förväntas finnas en selektion vid förmätningen.

Analys

Antalet personer som fördelats till undersökningsgrupperna redovisades i 86 procent av artiklarna. Information om antalet som avbröt medverkan, respektive som analyserades var 78 respektive 88 procent. I några fall var det först efter ett visst ”detektivarbete” som det gick att räkna fram (t.ex., genom att addera bortfall till analyserade grupper kunde den ursprungliga undersökningsgruppen beräknas). I några fall redovisades inte bortfall per undersökningsbetingelse.

Av de 169 studierna använde drygt hälften termen intent-to-treat, vilket innebär att man behållit individer som avbrutet studien i undersökningsmaterialet. De fall där det inte förekom något bortfall har kategoriserats som intent-to-treat. Om bortfallet är omfattande används ofta så kallad imputering, vilket innebär att man statistiskt uppskattar vilket värde individer som avbrutit skulle ha fått. Det hade förekommit i 15 procent av utvärderingarna. Imputering var inte vanligare när bortfallet av personer var större⁵.

Även om randomisering används för att skapa maximalt lika grupper vid förmätning kan tillfälligheter göra att grupperna skiljer sig åt på väsentliga punkter. Därför är det viktigt att vid behov justera för initiala skillnader vid förmätning, både för hela undersökningsgruppen och för respektive under-

⁵ $F(1,161)=0,03, p > .05$ (bortfallet var i genomsnitt 12 procent när imputering genomförts och 11 procent när det inte skett; för sju artiklar saknas information om bortfall).

sökningsbetingelse. När det fanns säkerställda skillnader vid förmätningen hade statistisk justering skett i två tredjedelar av fallen.

När flera försökspersoner tillhör samma behandlingsgrupp (t.ex., en skolklass) finns risken att de påverkar varandra, vilket gör att behandlingseffekten kan överskattas. Knappt hälften (12 av 27) av de studier som fördelat hela grupper till interventions- respektive kontrollgrupp hade anpassat den statistiska analysen så att hänsyn togs till eventuella gruppeffekter. Resultaten är förmodligen i underkant eftersom det endast redovisar de utvärderingar där kluster explicit angetts av författarna. I flera utvärderingar är det möjligt att det ändå fanns ett beroende mellan personer i utvärderingen.

Redovisat resultat

Flertalet författare (77 %) redovisade beskrivande statistik, vilket är nödvändigt om andra ska kunna kontrollera och jämföra resultaten med annan forskning.

En signifikant effekt på det primära utfallsmåttet förekom i drygt hälften av fallen (56 %). Eftersom tillräcklig statistisk styrka (power) är en förutsättning för att säkert kunna identifiera signifikanta effekter undersöktes om det var relaterat till utfallet. Så var inte fallet; bland de 41 utvärderingarna med powerberäkning var 59 procent statistiskt säkerställda och bland de 128 utan powerberäkning 56 procent⁶.

Det var endast knappt hälften som redovisade effektstorlekar (Cohens *d*). I flertalet av artiklarna, 163 (96 %), gick det dock att extrahera information för att beräkna den. För 110 utvärdering bygger beräkningarna på medelvärden och standardavvikelser och för övriga 53 på andra typer av information (som risk-kvot och odds-kvot). Det fanns en statistiskt säkerställd skillnad, där de förra genererade i genomsnitt högre effektstorlekar ($d = .56$) än de senare ($d = .28$)⁷.

Redovisning av eventuella oönskade utfall förekom endast i drygt var tionde artikel.

Samlad bild av forskningskvalitet

Totalt 16 kriterier på kvalitet adderades per artikel för att få ett övergripande index på vetenskaplig kvalitet (se kriterier med en asterisk i tabell 1)⁸. I genomsnitt uppfyllde utvärderingarna 8,0 ($SD = 2,6$) kriterier, med en variation mellan 2 och 15. Detta samlade kvalitetsmått var tydligt relaterat till publikationsår (1992-2012); ju nyare publikation, desto högre kvalitet (figur 3)⁹. Skillnaden blev tydligare när studiens startår togs som utgångspunkt (figur 4), alltså när datainsamlingen påbörjades (1974-2011)¹⁰. Resultatet är rimligt eftersom det år utvärderingen startar är tydligare kopplat till kunskapen om

⁶ $\chi^2(1) = 0,80, p > 0,5$

⁷ $F(1,162) = 4,63, p < .05$

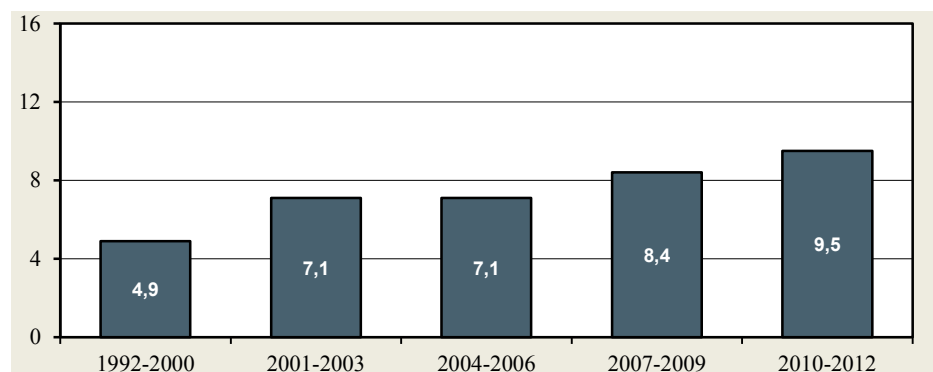
⁸ Flera olika versioner av indexet testades (t.ex., ett bestående av samtliga kriterier i tabell 1), men med ungefär samma resultat som följd.

⁹ $F(4,168) = 12,39, p < .0001$

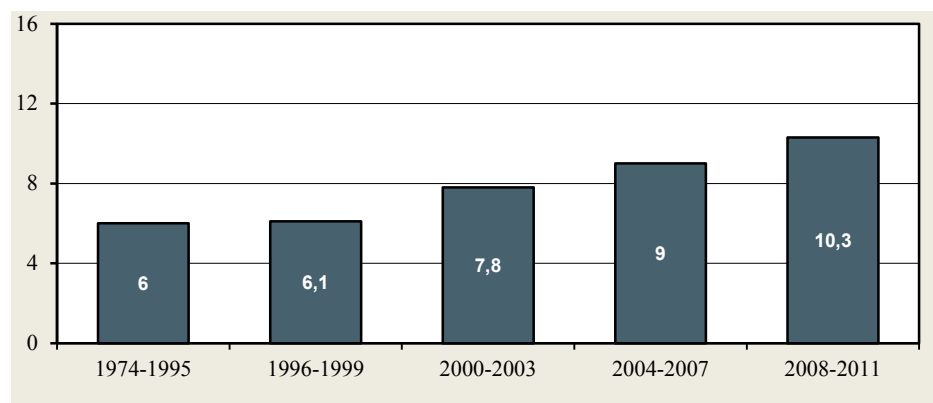
¹⁰ $F(4,166) = 15,70, p < .0001$ (för två artiklar saknas information om startår)

hur en effektutvärdering bör genomföras, än till publikationsår (eftersom genomförandetiden varierat mellan 1 och 28 år)¹¹.

Figur 3. Index för vetenskaplig kvalitet och år studien publicerades



Figur 4. Index för vetenskaplig kvalitet och år studien påbörjades



Test av enkla huvudeffekter visar att den vetenskapliga kvaliteten ökade för artiklar publicerade från och med 2001 samt från och med 2007. Med utgångspunkt från när utvärderingarna startade så fanns statistiskt säkerställda ökningarna i kvalitet 2000, 2004 och 2008.

Även om kvaliteten ökade över tid så finns brister. För de 60 artiklar som publicerades de tre senaste publikationsåren (2010-2012) hade visserligen så gott som samtliga tydliga inklusions- och exklusionskriterier (97 %) och beskrivande statistik (88 %). Majoriteten hade också angett primärt utfallsmått (57 %), beskrivit psykometriska egenskaper (65 %) och i vilken omfattning personerna fått interventionen (65 %), använt intent-to-treat (70 %) samt redovisat effektstorlekar (67 %). Mindre vanligt var att ha beräknat nödvändig gruppstorlek, power (33 %), redovisat eventuella oönskade utfall (20 %) eller använt ”blinda” datainsamlare (17 %).

¹¹ Den genomsnittliga forskningstiden sjönk från 7,2 år (1974-1995), 6,1 (1996-1999), 5,1 (2000-2003), 4,0 (2004-2007) till 2,4 (2008-2011), $F(4,186) = 9,76, p < .0001$. De senare siffrorna kan förväntas öka i takt med att flera påbörjade studier avslutas.

Koppling forskningskvalitet och resultat

En högre vetenskaplig kvalitet (indexet) förklarades statistiskt ($R^2(adj) = .33$) av att den var nyligen startad (betavikt = .52) samt att behandlingen avsåg ett uttalat problem, så kallad indikerad prevention¹² (betavikt = .30). Variabler som inte bidrog till att förklara kvaliteten var om det handlade om en modeltvärdering eller inte, om den ansvarige forskaren kom från en medicinsk disciplin eller psykologi, om den undersökta målgruppen bestod av barn alternativt äldre samt om studien finansierats av ett forskningsråd eller inte¹³.

Det samlade kvalitetsindexet var svagt relaterat till effektstorleken¹⁴; ju högre kvalitet, desto starkare effekt. Det fanns också ett säkerställt samband mellan kvalitetskriteriet och om det primära utfallsmåttet var statistiskt säkerställt¹⁵.

Av de 169 utvärderingarna kom 40 från en grupp forskare som specialiserats sig på utvärderingar av internetbaserad psykoterapi. Dessa artiklar hade högre genomsnittlig vetenskaplig kvalitet ($M=9,4$) än de övriga 129 utvärderingarna ($M=7,6$)¹⁶. Även om de 40 effektutvärderingarna exkluderades så erhöles ett snarlikt resultat som i figur 3 och 4; den rapporterade vetenskapliga kvaliteten ökade med tiden¹⁷.

¹² I motsats till insatser till alla utan undantag (universell prevention) och insatser till riskgrupper (selektiv prevention).

¹³ De nio prediktorerna var inte högt korrelerade ($1,07 \leq VIF \leq 1,31$).

¹⁴ $r(n=163) = .19, p < .05$.

¹⁵ $F(1,168) = 5,68, p < .05$; När det primära utfallsmåttet var statistiskt säkerställt var kvalitetskriteriet i genomsnitt 8,5 ($SD = 2,9$) och när det inte var statistiskt säkerställt 7,4 ($SD = 2,8$).

¹⁶ $F(1,168) = 12,72, p < .0005$

¹⁷ Interaktionen mellan startår enligt gruppering ovan och forskargrupp var inte statistiskt säkerställt: $F(5,118) = 0,05, p > .05$ (eftersom forskargruppen med internetbaserad psykoterapi publicerade sin första artikel ungefär 2000 har tidigare utvärderingar exkluderats i analyserna).

Diskussion

Granskningen av de 169 peer-review-granskade artiklarna visar på en stor variation i rapporterad vetenskaplig kvalitet. Denna variation kunde kopplas till året när utvärderingen startade respektive när den publicerades; ju nyare artikel, desto högre vetenskaplig kvalitet.

Även om antalet rapporterade kvalitetskriterier ökade över tid, finns fortfarande brister i de senaste årens publikationer. För de senast publicerade artiklarna var det genomsnittliga antalet kvalitetskriterier 9,5 av 16 möjliga. De utvärderingar som startats sedan 2008 uppfyllde 10,3 kriterier av 16. Det finns med andra ord en utvecklingspotential.

Ett noterbart resultat är att det förekom statistiskt säkerställda effekter på det primära utfallsmåttet i bara drygt hälften av fallen. Eftersom det resultatet inte enkelt kan förklaras av bristande statistisk styrka (power) åskådliggör det betydelsen av att genomföra effektutvärderingar av interventioner för att undvika att sprida ineffektiva interventioner.

Det är också värt att notera att sju av åtta utvärderingar var randomiserade. Av de som startat sedan 2007 var nästan samtliga, 68 av 70, randomiserade. Det visar att det är möjligt att bedriva randomiserade utvärderingar inom i princip samtliga områden som avser vård och omsorg.

Det är framför allt de äldre utvärderingarna, till exempel de som startade före 2000, som i genomsnitt har lägre vetenskaplig kvalitet. Därmed ökar också risken att de redovisar resultat som drar felaktiga slutsatser om interventioners verkliga effekter.

Förklaring till bristande kvalitet

Det har skett en utveckling av teori och teknik för att genomföra effektutvärderingar med ökad tillförlitlighet. Det i sin tur förändrar innebörden av vad som är hög kvalitet; det som idag är metodologiskt avancerat kommer förmodligen att vara standard om några år. Och det som ansågs vara god kvalitet tidigare är inte alltid det idag. Exempel på det som blivit vanligare i detta material är att forskarna publicerar sina forskningsprotokoll före datainsamlingen startar, att de rapporterar vad som är primärt utfallsmått, antal undersökta respektive analyserade personer, intent-to-treat-analys, imputering samt redovisning av effektstorlek.

Den första kända effektutvärderingen publicerades för bara drygt två decennier sedan, 1992. Fortfarande finns det få svenska studier av psykologiska och sociala interventioner. Frånvaron av studier som kan ge vägledning om vilka interventioner som fungerar bättre respektive sämre kan tyckas märklig, inte minst med tanke på att hundratusentals personer årligen får insatser inom ramen för socialtjänsten och inom psykiatrin. De allra flesta av dessa interventioner används utan direkt vetenskapligt stöd [2]. Inom medicinsk forskning ställs höga krav om att endast introducera nya interventioner som har ett vetenskapligt stöd. I annat fall förväntas en försiktig introduktion som åtföljs

av noggrann dokumentation för att kunna upptäcka eventuella oönskade effekter.

Av de 169 utvärderingarna behandlar majoriteten (81 %) psykiatriska tillstånd, missbruk och beroende samt psykologiska och sociala interventioner inom hälso- och sjukvård. Det är endast 19 utvärderingar som berör äldre samt traditionell individ- och familjeomsorg (inkl. ytterligare några som behandlar missbruksbehandling inom ramen för socialtjänstens arbete). Inom dessa områden finns med andra ord ett stort behov av forskning, inte minst med tanke på att ungefär 800 000 svenskar årligen blir föremål för interventioner inom socialtjänsten. Noterbart är att de 19 utvärderingarna håller hög vetenskaplig kvalitet,

Att det inte publicerats flera effektutvärderingar och att de som finns inte har högre genomsnittlig kvalitet, är dock inte överraskande eftersom det är relativt ovanligt att blivande forskare ges möjlighet att medverka i effektutvärderingar under sin utbildning. En inventering av tio års samtliga avhandlingarna från sju vetenskapliga discipliner [13] identifierade 1 402 avhandlingar. Av dem inkluderade endast fyra procent kontrollerade effektutvärderingar av interventioners effekter. Inom den professionsanknutna disciplinen socialt arbete är det exempelvis endast sex av 319 personer som har disputerat på avhandlingar som innehåller data från kontrollerade effektutvärderingar [14].

En annan förklaring till bristen på effektutvärderingar är att de är relativt kostsamma. En effektutvärdering kostar ofta 2-4 miljoner kronor. De sju effektutvärderingar av sociala interventioner som Vetenskapsrådet (VR) och Forskningsrådet för arbetsliv och socialvetenskap (FAS) beviljade under perioden 2005-2008 fick i genomsnitt 1 900 000. Det är väsentligt mer än VR:s genomsnittliga projektbidrag under samma period som var 760 000 och FAS 1 400 000 [15]. Anledningen till att effektutvärderingar är kostsamma är bland annat att de ofta kräver längre tid att genomföra tack vare för- och eftermätning. Dessutom behövs ofta resurser för att implementera den intervention som ska utvärderas och för att motivera professionella att medverka. Det kan medföra att forskningsråden prioriterar bort dessa relativt sett dyrbara effektutvärderingar i förhållande till andra billigare studier. Det finns de som menar att kommuner själva ska finansiera effektutvärderingar av socialtjänstens interventioner. Bortsett från att det skulle medföra att forskningen skulle fokuseras till främst storstadsregioner som har tillräckliga resurser för denna typ av verksamhet, skulle det också betyda att de närmast berörda skattebetalarna tvingas finansiera angelägen forskning om socialtjänstens interventioner två gånger, dels genom statsskatten som finansierar forskarråden, dels genom kommunalskatten.

Till sist har denna typ av kvantitativ forskning ett lågt akademiskt värde inom delar av svensk socialvetenskaplig forskning. Under flera decennier har stora delar av den svenska sociala och beteendevetenskapliga forskningen utgått från vetenskapsfilosofiska perspektiv som gynnar kvalitativ forskning och socialkonstruktivism. Ett citat belyser inställningen till effektutvärderingar: "Den akademiska forskningen har som sin främsta uppgift att bedriva grundforskning /.../ Tillämpad forskning är problematisk för den akademiska

forskningen generellt. Dels skapar den ett beroende mellan uppdragsgivare och forskare, vilket nästan dagligen uppmärksammas inom till exempel medicinsk forskning. Dels är inte resultaten som uppdragsgivaren vill ha direkt användbara för publicering i någon vetenskaplig tidskrift” [16]. Resultaten i denna rapport ger inget stöd för vare sig påståendet om svårigheter att publicera effektutvärderingar eller att forskare skulle förledas att producera efterfrågade resultat (endast drygt hälften av utfallsmåtten var statistiskt säkert ställda).

Ökad forskarkompetens

Resultaten antyder ett behov av förstärkt forskarutbildning med inriktning mot effektutvärderingar av psykologiska och sociala interventioner. Förutsättningen för att utbilda doktorander är att det finns handledare som behärskar metodiken. En nyckeluppgift är således att fortbilda empiriskt verk-samma forskare. Ett annat sätt är ett internationellt utbyte som ger svenska forskarstuderande möjlighet att utbildas i länder som har den efterfrågade kompetensen. Ett tredje sätt är att möta bristen på handledarkompetens är att göra det via tvärvetenskapliga institutioner där forskare från discipliner med större erfarenhet av effektutvärderingar ingår.

Tillförlitlighet

Vetenskaplig kvalitet är ett komplext begrepp. Den information som extrahe-rats från utvärderingarna är således en kompromiss mellan vad som bedöms som relevant och vad som är möjligt att tillförlitligt kategorisera från publice-rade artiklar. Vetenskaplig kvalitet omfattar så mycket mer än det som granskats i denna studie [17]. Dit hör svårbedömda kriterier som exempelvis obefogade generaliseringar och att negligera eller trivialisera begränsningar. En faktor som är betydelsefull som vi valt att inte kategorisera är implemen-teringskvalitet, eftersom det saknas konsensus om hur det ska beskrivas. De kriterier som använts är inte heller utan förbehåll. En randomiserad utvärde-ring är exempelvis ingen garanti för trovärdiga resultat; ett omfattande eller ett differentierat bortfall mellan de undersökta grupperna kan ge snedvridna resultat. Ett annat exempel är att användningen av bedömningsmetoder med goda psykometriska egenskaper inte är någon garanti för att bedömningsme-toderna använts på avsett sätt, eller för en korrekt målgrupp. Ett tredje exem-pel är om forskaren rapporterat om intressekonflikter, exempelvis att man haft ekonomiska intressen i den utvärderade interventionen. I det fallet är det osäkert om tillförlitligheten i uppgifterna eftersom det finns forskare som uppgivit att det inte funnits någon intressekonflikt när så varit fallet. Mer information om kriterier för vetenskaplig kvalitet finns i den av Socialstyrelsen producerade boken *Att göra effektutvärderingar* [2].

Alla resultat bygger på information i artiklarna. Det är alltså möjligt att författarna inte har angett allt av betydelse för att kunna värdera den vetenskapliga kvaliteten, exempelvis eftersom det inte tidigare har krävts för att kunna publiceras. Det bedömdes inte som meningsfullt att kontakta författar-

na och efterhöra om någon information saknas, bland annat på grund av risken att författarna inte längre kommer ihåg detaljer, en risk som rimligen ökar ju äldre studien är (den äldsta utvärderingen startade 1974).

Transparens är vetenskapens främsta kännetecken; det ska tydligt framgå vilken frågeställning som forskaren undersökt samt hur data samlats in och analyserats. Det betyder att även om all relevant information inte redovisas i artiklarna så är det den som redovisas som forskarsamhället i regel förhåller sig till. Det betyder att utvärderingar som framstår som kvalitetsmässigt sämre riskerar att inte tillmätas något värde. Den forskare som vill att hans eller hennes forskning ska komma till praktisk användning behöver således anpassa sig till de krav som utvecklats för att producera och tolka evidens.

Utöver de 169 publicerade effektutvärderingarna finns ytterligare ett femtiotal tekniska rapporter samt magister- och doktorsavhandlingar. De ingår inte i denna granskning. En anledning är att flera tekniska rapporter historiskt sett har skrivits om och senare publicerats i peer-review-granskade tidskrifter (minst 16 av de 169¹⁸; ytterligare minst sju tekniska rapporter har publicerats efter april 2012, alternativt är inskickade för bedömning).

Sökningen efter effektutvärderingar har komplicerats eftersom det saknas ett nationellt register över dem. På grund av det ökande antalet vetenskapliga studier (som många gånger är otillräckligt eller felaktigt indexerade i databaser) är det mycket arbetskrävande att identifiera svenska effektutvärderingar. Sökningar på universitetsinstitutioner har fångat in en del av effektutvärderingarna, men långt ifrån alla. Som framgår av metodavsnittet har sökningar gjorts brett, men trots det har utvärderingar identifierats på andra sätt (t.ex., tips från kollegor). Det kan alltså finnas ytterligare okända effektutvärderingar från den aktuella tidsperioden.

¹⁸ De 16 tekniska rapporter som senare publicerats i peer-review granskade tidskrifter avvek inte i kvalitet ($M = 7,50$ jämfört med $M = 7,74$ för övriga), $F(1,168) = 0,11, p > .05$

Referenser

1. Jergeby, U, Sundell, K. Evidensbaserat socialt arbete. In: Jergeby U, editor. Evidensbaserad praktik i socialt arbete. Stockholm: Gothia; 2008. p. 11-8.
2. Sundell, K, editor. Att göra effektutvärderingar. Stockholm: Gothia; 2012.
3. Socialstyrelsen. Socialtjänstens öppna verksamheter för barn och unga – en nationell inventering av metoder Stockholm; 2009.
4. Beutler, LE, Malik, M., Alimohamid, S., Harwood, T.M., Talebi, H., Noble, S., & Wong, E., editor. Bergin and Garfield's handbook of psychotherapy and behavior change New York: Wiley; 2003.
5. Lilienfeld, SO. Psychological treatments that cause harm. *Perspectives on Psychological Science*. 2007; 2(1):53-70.
6. Sundell, K, Vinnerljung, B. Goda intentioner kan vålla skada – om etik i socialt arbete. In: Jergeby U, editor. Evidensbaserad praktik i socialt arbete Stockholm: Gothia; 2008. p. 36-46.
7. Sackett, DL, Richardson, WS, Rosenberg, W, Haynes, RB. Evidence-based medicine: How to practice and teach EBM. New York: Churchill Livingstone; 1997.
8. Socialstyrelsen. Svenska effektutvärderingar av psykosociala interventioner. Stockholm: Socialstyrelsen; 2012.
9. Moher, D, Hopewell, S, Schulz, K, Monitori, V, Gøtzsche, P, Devereaux, P, et al. CONSORT 2010 Explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *British Medical Journal*. 2010; 340(c869):doi:10.1136/bmj.c869.
10. Des Jarlais, D, Lyles, C, Crepaz, N, Group, T. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*. 2004; 94(3):361-6.
11. Flay, BR, Biglan, A, Boruch, RF, Castro, FG, Gottfredson, D, Kellam, S, et al. Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*. 2005; 6(3):151-75.
12. Lipsey, M, Wilson, D. Practical meta-analysis. Thousand Oaks: Sage Publications Inc.; 2001.
13. Sundell, K, Stensson, E. Effektutvärderingar i doktorsavhandlingar. Stockholm: Socialstyrelsen; 2010.
14. Brunnberg, E. Om socionomutbildningar, forskarutbildningar och samtliga avhandlingar i socialt arbete 1980-2012. Förteckning över samtliga doktors- och licentiatavhandlingar i socialt arbete i Sverige. Västerås: Mälardalens högskola.
15. Socialstyrelsen. Svensk och internationell forskning om sociala interventioners effekter. Stockholm: Socialstyrelsen; 2011.
16. Blomgren, C. Forskning och praktik: två sidor av socialt arbete. *Socionomen*. 2009; 1:20-5.
17. Ioannidis, J. Why most discovered true associations are inflated. *Epidemiology*. 2008; 19(5):640-8.

Bilaga 1. Granskade artiklar

- Ahmad A, Larsson B, Sundelin-Wahlsten V. EMDR treatment for children with PTSD: results of a randomized controlled trial. *Nord J Psychiatry*. 2007;61(5):349-54.
- Andersson, G, Asmundson, GJG, Denev, J, Nilsson, J, Larsen, HC. A controlled trial of cognitive behavior therapy combined with vestibular rehabilitation in the treatment of dizziness. *Behaviour research and therapy*. 2006; 44:1265-73.
- Andersson G, Bergstrom J, Hollandare F, Carlbring P, Kaldo V, Ekselius L. Internet-based self-help for depression: randomised controlled trial. *The British journal of psychiatry : the journal of mental science*. 2005 Nov;187:456-61.
- Andersson, G, Carlbring, P, Furmark, T, on behalf of the SOFIE Research Group. Therapist experience and knowledge acquisition in Internet-delivered CBT for social anxiety disorder: A randomized controlled trial. *PloS one*. 2012; 7(5):e37411.
- Andersson G, Carlbring P, Holmström A, Sparthan E, Furmark T, Nilsson-Ihrfelt E, et al. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. *Journal of consulting and clinical psychology*. 2006 Aug;74(4):677-86.
- Andersson E, Enander J, Andren P, Hedman E, Ljotsson B, Hursti T, et al. Internet-based cognitive behaviour therapy for obsessive-compulsive disorder: a randomized controlled trial. *Psychological medicine*. 2012 Oct;42(10):2193-203.
- Andersson, G, Green, M, Melin, L. Behavioural hearing tactics: a controlled trial of a short treatment programme. *Behaviour research and therapy*. 1997; 35:523-30.
- Andersson, G, Johansson, C, Nordlander, A, Asmundson, GJG. Chronic pain in older adults: A controlled pilot trial of a brief cognitive-behavioural group treatment. *Behavioural and Cognitive Psychotherapy*. 2012; 40:239-44.
- Andersson, G, Melin, L, Scott, B, Lindberg, P. Behavioural counselling for subjects with acquired hearing loss. A new approach to hearing tactics. *Scandinavian Audiology*. 1994; 23:249-56.
- Andersson, G, Lundström, P, Ström, L. Internet-based treatment of headache. Does telephone contact add anything? . *Headache*. 2003; 43:353-61.
- Andersson, G, Melin, L, Scott, B, Lindberg, P. An evaluation of a behavioural treatment approach to hearing impairment. *Behaviour research and therapy*. 1995; 33:283-92.
- Andersson, G, Paxling, B, Roch-Norlund, P, Östman, G, Norgren, A, Almlöv, J, et al. Internet-based psychodynamic vs. cognitive behavioural guided self-help for generalized anxiety disorder: A randomised controlled trial. *Psychotherapy and Psychosomatics*. 2012; 81:344-55.
- Andersson G, Porsaeus D, Wiklund M, Kaldo V, Larsen HC. Treatment of tinnitus in the elderly: a controlled trial of cognitive behavior therapy. *International journal of audiology*. 2005 Nov;44(11):671-5.
- Andersson, G, Strömgren, T, Ström, T, Lyttkens, L. Randomised controlled trial of Internet based cognitive behavior therapy for distress associated with tinnitus. *Psychosomatic Medicine*. 2002; 64:810-6.
- Andersson E, Walen C, Hallberg J, Paxling B, Dahlin M, Almlöv J, et al. A randomized controlled trial of guided Internet-delivered cognitive behavioral therapy for erectile dysfunction. *The journal of sexual medicine*. 2011 Oct;8(10):2800-9.
- Andreasson S, Hansagi H, Österlund B. Short-term treatment for alcohol-related problems: four-session guided self-change versus one session of advice--a randomized, controlled trial. *Alcohol*. 2002 Aug;28(1):57-62.
- Andrée Löfholm, C, Olsson, T, Sundell, K, Hansson, K. Multisystemic Therapy with Conduct Disordered Youth: Stability of Treatment Outcomes Two Years after Intake. *Evidence & Policy*. 2009; 5:373-97.

-
- Andren S, Elmstahl S. Effective psychosocial intervention for family caregivers lengthens time elapsed before nursing home placement of individuals with dementia: a five-year follow-up study. *International psychogeriatrics*. 2008 Dec;20(6):1177-92.
- Axberg U, Broberg A. Evaluation of "The incredible years" in Sweden: The transferability of an American parent-training program in Sweden. *Scandinavian Journal of Psychology*. 2012;53:224-32.
- Axberg U, Hansson K, Broberg AG, Wirtberg I. The development of a systemic school-based intervention: Marte Meo and coordination meetings. *Family process*. 2006 Sep;45(3):375-89.
- Ballidin J, Berglund M, Borg S, Månsson M, Bendtsen P, Franck J, et al. A 6-month controlled naltrexone study: combined effect with cognitive behavioral therapy in outpatient treatment of alcohol dependence. *Alcoholism, clinical and experimental research*. 2003 Jul;27(7):1142-9.
- Bengtsson-Tops, A, Hansson, L. Clinical and social changes in severely mentally ill individuals admitted to an outpatient psychosis team: an 18-month follow-up study. *Scandinavian journal of caring sciences*. 2003; 17(1):3-11.
- Bergström J, Andersson G, Ljotsson B, Ruck C, Andreewitch S, Karlsson A, et al. Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. *BMC psychiatry*. 2010;10:54.
- Bergström M, Kieler H, Waldenstrom U. Effects of natural childbirth preparation versus standard antenatal education on epidural rates, experience of childbirth and parental stress in mothers and fathers: a randomised controlled multicentre trial. *BJOG : an international journal of obstetrics and gynaecology*. 2009 Aug;116(9):1167-76.
- Berman AH. The Reasoning and Rehabilitation Program: Assessing Short- and Long-Term Outcomes Among Male Swedish Prisoners. *Journal of Offender Rehabilitation*. 2004;40(1/2):85-103.
- Berman AH, Forsberg L, Durbeej N, Källmen H, Hermansson U. Single-session motivational interviewing for drug detoxification inpatients: effects on self-efficacy, stages of change and substance use. *Subst Use Misuse*. 2010 Feb;45(3):384-402.
- Berman AH, Lundberg U. Auricular acupuncture in prison psychiatric units: a pilot study. *Acta psychiatrica Scandinavica Supplementum*. 2002(412):152-7.
- Berman AH, Lundberg U, Krook AL, Gyllenhammar C. Treating drug using prison inmates with auricular acupuncture: a randomized controlled trial. *J Subst Abuse Treat*. 2004 Mar;26(2):95-102.
- Björkman T, Hansson L, Sandlund M. Outcome of case management based on the strengths model compared to standard care. A randomised controlled trial. *Social psychiatry and psychiatric epidemiology*. 2002 Apr;37(4):147-52.
- Bodin M, Leifman H. A randomized effectiveness trial of an adult-to-youth mentoring program in Sweden. *Addiction Research & Theory*. 2011;19(5):438-47.
- Bodin M, Strandberg, AK. The Örebro prevention program revisited: A cluster-randomized effectiveness trial of program effects on youth drinking. *Addiction*. 2011;106:2134-43.
- Brinkborg H, Michanek J, Hesser H, Berglund G. Acceptance and commitment therapy for the treatment of stress among social workers: a randomized controlled trial. *Behaviour research and therapy*. 2011 Jun;49(6-7):389-98.
- Burell G, Öhman A, Sundin, Ström G, Ramund B, Cullhed I, et al. Modification of the Type A behavior pattern in post-myocardial infarction patients: a route to cardiac rehabilitation. *International journal of behavioral medicine*. 1994;1(1):32-54.
- Buhrman, M, Fältenhag, S, Ström, L, Andersson, G. Controlled trial of Internet-based treatment with telephone support for chronic back pain. *Pain*. 2004; 111:368-77.
- Buhrman, M, Nilsson-Ihrfelt, E, Jannert, M, Ström, L, Andersson, G. Guided Internet-delivered cognitive-behavioral treatment for chronic back pain reduces pain catastrophizing: a randomized controlled trial. *Journal of Rehabilitation Medicine*. 2011; 43:500-5.

-
- Carlbring P, Apelstrand M, Sehlin H, Amir N, Rousseau A, Hofmann SG, et al. Internet-delivered attention bias modification training in individuals with social anxiety disorder--a double blind randomized controlled trial. *BMC psychiatry*. 2012;12:66.
- Carlbring P, Bohman S, Brunt S, Buhrman M, Westling BE, Ekselius L, et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. *The American journal of psychiatry*. 2006 Dec;163(12):2119-25.
- Carlbring P, Ekselius L, Andersson G. Treatment of panic disorder via the Internet: a randomized trial of CBT vs. applied relaxation. *Journal of behavior therapy and experimental psychiatry*. 2003 Jun;34(2):129-40.
- Carlbring P, Gunnarsdottir M, Hedensjö L, Andersson G, Ekselius L, Furmark T. Treatment of social phobia: randomised trial of internet-delivered cognitive-behavioural therapy with telephone support. *The British journal of psychiatry : the journal of mental science*. 2007 Feb;190:123-8.
- Carlbring P, Jonsson J, Josephson H, Forsberg L. Motivational interviewing versus cognitive behavioral group therapy in the treatment of problem and pathological gambling: a randomized controlled trial. *Cogn Behav Ther*. 2010;39(2):92-103.
- Carlbring P, Maurin L, Törngren C, Linna E, Eriksson T, Sparthian E, et al. Individually-tailored, Internet-based treatment for anxiety disorders: A randomized controlled trial. *Behaviour research and therapy*. 2011 Jan;49(1):18-24.
- Carlbring P, Maurin T, Sjömark J, Maurin L, Westling BE, Ekselius L, et al. All at once or one at a time? A randomized controlled trial comparing two ways to deliver bibliotherapy for panic disorder. *Cogn Behav Ther*. 2011;40(3):228-35.
- Carlbring P, Nilsson-Ihrfelt, E, Waara, J, Kollenstam, C, Buhrman, M, Kaldo, V, et al. Treatment of panic disorder: Live therapy vs. self-help via Internet. *Behaviour research and therapy*. 2005; 43:1321-33.
- Carlbring P, Smit F. Randomized trial of internet-delivered self-help with telephone support for pathological gamblers. *Journal of consulting and clinical psychology*. 2008 Dec;76(6):1090-4.
- Carlbring P, Westling, BE, Ljungstrand, P, Ekselius, L, Andersson, G. Treatment of panic disorder via the Internet- a randomized trial of a self-help program. *Behavior Therapy*. 2002; 32:751-64.
- Cedereke M, Monti K, Öjehagen A. Telephone contact with patients in the year after a suicide attempt: does it affect treatment attendance and outcome? A randomised controlled study. *European psychiatry : the journal of the Association of European Psychiatrists*. 2002 Apr;17(2):82-91.
- Dahlrup B, Nordell E, Andren S, Elmståhl S. Family caregivers' assessment of symptoms in persons with dementia using the GBS-scale: differences in rating after psychosocial intervention--an 18-month follow-up study. *Clinical interventions in aging*. 2011;6:9-18.
- Duberg A, Hagberg L, Sunvisson H, Möller M. Influencing self-rated health among adolescent girls with dance intervention: a randomized controlled trial. *JAMA pediatrics*. 2013 Jan;167(1):27-31.
- Eberhard S, Nordström G, Höglund P, Öjehagen A. Secondary prevention of hazardous alcohol consumption in psychiatric out-patients: a randomised controlled study. *Social psychiatry and psychiatric epidemiology*. 2009 Dec;44(12):1013-21.
- Egberg Thyme K, Sundin E, Ståhlberg G, Lindström B, Eklöf H, Wiberg B. The outcome of short-term psychodynamic art therapy compared to short-term psychodynamic verbal therapy for depressed women. *Psychoanalytic Psychotherapy*. 2007;21(3):250-64.
- Eklund M, Sandlund M. The life situation of people with persistent mental illness visiting day centers: a comparative study. *Community mental health journal*. 2011 Oct;48(5):592-7.
- Ekman DS, Andersson A, Nilsen P, Stahlbrandt H, Johansson AL, Bendtsen P. Electronic screening and brief intervention for risky drinking in Swedish university students--a randomized controlled trial. *Addictive behaviors*. 2011 Jun;36(6):654-9.

- Ekstrand M, Larsson M, Darj E, Tyden T. Advance provision of emergency contraceptive pills reduces treatment delay: a randomised controlled trial among Swedish teenage girls. *Acta obstetrica et gynecologica Scandinavica*. 2008;87(3):354-9.
- Enebrink P, Högstrom J, Forster M, Ghaderi A. Internet-based parent management training: a randomized controlled study. *Behaviour research and therapy*. 2012 Apr;50(4):240-9.
- Faggiano F, Galanti MR, Bohrn K, Burkhart G, Vigna-Taglianti F, Cuomo L, et al. The effectiveness of a school-based substance abuse prevention program: EU-Dap cluster randomised controlled trial. *Preventive medicine*. 2008 Nov;47(5):537-43.
- Forsberg KA, Björkman T, Sandman PO, Sandlund M. Physical health--a cluster randomized controlled lifestyle intervention among persons with a psychiatric disability and their staff. *Nord J Psychiatry*. 2008;62(6):486-95.
- Forsberg L, Ekman S, Halldin J, Rönnerberg S. Brief interventions for risk consumption of alcohol at an emergency surgical ward. *Addictive behaviors*. 2000 May-Jun;25(3):471-5.
- Forsberg L, Ernst D, Sundqvist K, Farbring, CÅ. Motivational Interviewing delivered by existing prison staff: A randomized controlled study of effectiveness on substance use after release. *Substance Use & Misuse*. 2011;46:1477-85.
- Forster M, Sundell K, Morris RJ, Karlberg M, Melin L. A randomized controlled trial of a standardized behavior management intervention for students with externalizing behavior. *Journal of Emotional and Behavioral Disorders*. 2012;20(3):169-83.
- Furmark T, Carlbring P, Hedman E, Sonnensteing A, Clevberger P, Bohman B, et al. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. *British Journal of Psychiatry*. 2009;195:440-7.
- Geidne S EC. Working with or against the stores? A non-governmental organization working with alcohol purchase attempts. *Health Education*. 2009;109(3):259-78.
- Ghaderi A. Does individualization matter? A randomized trial of standardized (focused) versus individualized (broad) cognitive behavior therapy for bulimia nervosa. *Behaviour research and therapy*. 2006 Feb;44(2):273-88.
- Ghaderi A, Mårtensson M, Schwan H. "Everybody's Different": a primary prevention program among fifth grade school children. *Eating disorders*. 2005 May-Jun;13(3):245-59.
- Ghaderi A, Scott B. Pure and guided self-help for full and sub-threshold bulimia nervosa and binge eating disorder. *The British journal of clinical psychology / the British Psychological Society*. 2003 Sep;42(Pt 3):257-69.
- Granath J, Ingvarsson S, von Thiele U, Lundberg U. Stress management: a randomized study of cognitive behavioural therapy and yoga. *Cogn Behav Ther*. 2006;35(1):3-10.
- Gustafsson, S, Eklund, K, Wilhelmson, K, Edberg, A, Johansson, B, Kronlöf, G, et al. Long-Term Outcome for ADL Following the Health-Promoting RCT--Elderly Persons in the Risk Zone. *The Gerontologist*. 2013; 53(4):654-63.
- Gyllensten AL, Hansson L, Ekdahl C. Outcome of Basic Body Awareness Therapy. A Randomized Controlled Study of Patients in Psychiatric Outpatient Care. *Advances in Physiotherapy*. 2003;5(4):179-90.
- Hallgren MÅ, Källmén H, Leifman H, Sjölund T, Andréasson S. Evaluation of an alcohol risk reduction program (PRIME for Life) in young Swedish military conscripts. *Health Education Journal*. 2009;109(2):155-68.
- Hallgren MÅ, Sjölund T, Källmén H, Andréasson S. Modifying alcohol consumption among high school students: An efficacy trial of an alcohol risk reduction program (PRIME for Life). *Health Education*. 2011;111(3):216-29.
- Hammarberg A, Wennberg P, Beck O, Franck J. A comparison of two intensities of psychosocial intervention for alcohol dependent patients treated with acamprosate. *Alcohol and alcoholism*. 2004 May-Jun;39(3):251-5.
- Hansson K, Cederblad M, Höök B. Funktionell familjeterapi. *Socialvetenskaplig tidskrift*. 2000;3(231-43).

- Hansson K, Johansson P, Drott-Englén G, Benderix Y. Funktionell familjeterapi i barnpsykiatrisk praxis - Om behandling av ungdomskriminalitet utanför universitetsforskningen. *Nordisk psykologi* 2004;56(4):304-20.
- Hansson K, Olsson M. Effects of multidimensional treatment foster care (MTFC): Results from a RCT study in Sweden. *Children and Youth Services Review*. 2012;34(9):1929-36.
- Hansson, H, Rundberg, J, Zetterlind, U, Johnsson, K, Berglund, M. Two-year outcome of an intervention program for university students who have parents with alcohol problems: a randomized controlled trial. *Alcoholism, clinical and experimental research*. 2007; 31(11):1927-33.
- Hedman E, Andersson G, Andersson E, Ljotsson B, Ruck C, Asmundson GJ, et al. Internet-based cognitive-behavioural therapy for severe health anxiety: randomised controlled trial. *The British journal of psychiatry : the journal of mental science*. 2011 Mar;198(3):230-6.
- Hedman E, Andersson G, Ljotsson B, Andersson E, Ruck C, Mörtberg E, et al. Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: a randomized controlled non-inferiority trial. *PloS one*. 2011;6(3):e18001.
- Hesser H, Gustafsson T, Lunden C, Henrikson O, Fattahi K, Johnsson E, et al. A randomized controlled trial of Internet-delivered cognitive behavior therapy and acceptance and commitment therapy in the treatment of tinnitus. *Journal of consulting and clinical psychology*. 2012 Aug;80(4):649-61.
- Hofman-Bang C, Lisspers J, Nordlander R, Nygren A, Sundin O, Öhman A, et al. Two-year results of a controlled study of residential rehabilitation for patients treated with percutaneous transluminal coronary angioplasty. A randomized study of a multifactorial programme. *European heart journal*. 1999 Oct;20(20):1695-74.
- Holländare F, Johnsson S, Randestad M, Tillfors M, Carlbring P, Andersson G, et al. Randomized trial of Internet-based relapse prevention for partially remitted depression. *Acta psychiatrica Scandinavica*. 2011 Oct;124(4):285-94.
- Högberg G, Pagani M, Sundin Ö, Soares J, Åberg-Wistedt A, Tärnell B, et al. On treatment with eye movement desensitization and reprocessing of chronic post-traumatic stress disorder in public transportation workers - A randomized controlled trial. *Nord J Psychiatry*. 2007;6(1):54-61.
- Jansson-Fröjmark M, Lind M, Sunnhed R. Don't worry, be constructive: a randomized controlled feasibility study comparing behaviour therapy singly and combined with constructive worry for insomnia. *The British journal of clinical psychology*. 2012 Jun;51(2):142-57.
- Jansson-Fröjmark M, Linton SJ, Flink IK, Granberg S, Danermark B, Norell-Clarke A. Cognitive-behavioral therapy for insomnia co-morbid with hearing impairment: a randomized controlled trial. *Journal of clinical psychology in medical settings*. 2012 Jun;19(2):224-34.
- Jansson M, Linton S. Cognitive-behavioral group therapy as an early intervention for insomnia: A randomized controlled trial. *Journal of Occupational Rehabilitation*. 2005;15(2):177-90.
- Jensen IB, Bergström G, Ljungquist T, Bodin L. A 3-year follow-up of a multidisciplinary rehabilitation programme for back and neck pain. *Pain*. 2005 Jun;115(3):273-83.
- Jensen IB, Busch H, Bodin L, Hagberg J, Nygren A, Bergström G. Cost effectiveness of two rehabilitation programmes for neck and back pain patients: A seven year follow-up. *Pain*. 2009 Apr;142(3):202-8.
- Jensen J, Lundin-Olsson L, Nyberg L, Gustafson Y. Fall and injury prevention in older people living in residential care facilities. A cluster randomized trial. *Annals of internal medicine*. 2002 May 21;136(10):733-41.
- Jernelöv S, Lekander M, Blom K, Rydh S, Ljotsson B, Axelsson J, et al. Efficacy of a behavioral self-help treatment with or without therapist guidance for co-morbid and primary insomnia--a randomized controlled trial. *BMC psychiatry*. 2012;12:5.

-
- Johansson AC, Linton SJ, Bergkvist L, Nilsson O, Cornefjord M. Clinic-based training in comparison to home-based training after first-time lumbar disc surgery: a randomised controlled trial. *European spine journal*. 2009 Mar;18(3):398-409.
- Johansson, M, Åkerlund, D, Larsen, HC, Andersson, G. Randomised controlled trial of vestibular rehabilitation combined with cognitive-behavioral therapy for dizziness in older people. *Otolaryngology-Head and Neck Surgery*. 2001; 125:151-6.
- Johansson R, Ekbladh S, Hebert A, Lindström M, Möller S, Petitt E, et al. Psychodynamic guided self-help for adult depression through the internet: a randomised controlled trial. *PloS one*. 2012;7(5):e38021.
- Johansson R, Sjöberg E, Sjögren M, Johnsson E, Carlbring P, Andersson T, et al. Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: a randomized controlled trial. *PloS one*. 2012;7(5):e36905.
- Johnsson KO, Berglund M. Comparison between a cognitive behavioural alcohol programme and post-mailed minimal intervention in high-risk drinking university freshmen: results from a randomized controlled trial. *Alcohol and alcoholism (Oxford, Oxfordshire)*. 2006 Mar-Apr;41(2):174-80.
- Johnsson KO, Berglund M. Education of key personnel in student pubs leads to a decrease in alcohol consumption among the patrons: a randomized controlled trial. *Addiction*. 2003 May;98(5):627-33.
- Kakko J, Grönbladh L, Svanborg KD, von Wachenfeldt J, Ruck C, Rawlings B, et al. A stepped care strategy using buprenorphine and methadone versus conventional methadone maintenance in heroin dependence: a randomized controlled trial. *The American journal of psychiatry*. 2007 May;164(5):797-803.
- Kakko J, Svanborg KD, Kreek MJ, Heilig M. 1-year retention and social function after buprenorphine-assisted relapse prevention treatment for heroin dependence in Sweden: a randomised, placebo-controlled trial. *Lancet*. 2003 Feb 22;361(9358):662-8.
- Kaldo V, Cars S, Rahnert M, Larsen HC, Andersson G. Use of a self-help book with weekly therapist contact to reduce tinnitus distress: a randomized controlled trial. *Journal of psychosomatic research*. 2007 Aug;63(2):195-202.
- Kaldo V, Levin S, Widarsson J, Buhrman M, Larsen HC, Andersson G. Internet versus group cognitive-behavioral treatment of distress associated with tinnitus: a randomized controlled trial. *Behav Ther*. 2008 Dec;39(4):348-59.
- Kimber B, Sandell R, Bremberg S. Social and emotional training in Swedish schools for the promotion of mental health: an effectiveness study of 5 years of intervention. *Health education research*. 2008 Dec;23(6):931-40.
- Kling A, Forster M, Sundell K, Melin L. A randomized controlled effectiveness trial of parent management training with varying degrees of therapist support. *Behav Ther*. 2010 Dec;41(4):530-42.
- Klingberg T, Fernell E, Olesen PJ, Johnson M, Gustafsson P, Dahlström K, et al. Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2005 Feb;44(2):177-86.
- Koertge J, Janszky I, Sundin O, Blom M, Georgiades A, Laszlo KD, et al. Effects of a stress management program on vital exhaustion and depression in women with coronary heart disease: a randomized controlled intervention study. *Journal of internal medicine*. 2008 Mar;263(3):281-93.
- Koutakis N, Stattin H, Kerr M. Reducing youth alcohol drinking through a parent-targeted intervention: the Orebro Prevention Program. *Addiction*. 2008 Oct;103(10):1629-37.
- Kristenson H, Osterling A, Nilsson JA, Lindgarde F. Prevention of alcohol-related deaths in middle-aged heavy drinkers. *Alcoholism, clinical and experimental research*. 2002 Apr;26(4):478-84.

-
- Kristensson, J, Ekwall, A, Jakobsson, U, Midlöv, P, Hallberg, I. Case managers for frail older people: a randomised controlled pilot study *Scandinavian journal of caring sciences*. 2010; 24(4):755-63.
- Kyhle Westermarck P, Hansson K, Olsson M. Multidimensional Treatment Foster Care - Results from an independent replication. *Journal of Family Therapy*. 2011;33:20-41.
- Levy Berg A, Sandell R, Sandahl C. Affect-focused body psychotherapy in patients with generalized anxiety disorder: Evaluation of an integrative method. *Journal of Psychotherapy Integration*. 2009;19(1):67-85.
- Lindahl Larsson M, Berglund M, Tonnesen H. Case management in aftercare of involuntarily committed patients with substance abuse. A randomized trial. *Nord J Psychiatry*. 2013 Jun;67(3):197-203.
- Lindström P, Svensson R. Attitudes towards drugs among school youths. An evaluation of the Swedish DARE programme. *Nordisk Alkohol & Narkotikatidskrift*. 1998;15:7-23.
- Linton SJ, Andersson T. Can chronic disability be prevented? A randomized trial of a cognitive-behavior intervention and two forms of information for patients with spinal pain. *Spine*. 2000 Nov 1;25(21):2825-31; discussion 4.
- Linton SJ, Boersma K, Jansson M, Overmeer T, Lindblom K, Vlaeyen JW. A randomized controlled trial of exposure in vivo for patients with spinal pain reporting fear of work-related activities. *European journal of pain (London, England)*. 2008 Aug;12(6):722-30.
- Linton SJ, Hellsing A, Bergström G. Exercise for workers with musculoskeletal pain: does enhancing compliance decrease pain? *Journal of Occupational Rehabilitation*. 1996;6(3):177-90.
- Linton SJ, Ryberg M. A cognitive-behavioral group intervention as prevention for persistent neck and back pain in a non-patient population: a randomized controlled trial. *Pain*. 2001 Feb 1;90(1-2):83-90.
- Lisspers J, Sundin O, Öhman A, Hofman-Bang C, Rydén L, Nygren A. Long-term effects of lifestyle behavior change in coronary artery disease: effects on recurrent coronary events after percutaneous coronary intervention. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*. 2005 Jan;24(1):41-8.
- Littbrand H, Lundin-Olsson L, Gustafson Y, Rosendahl E. The effect of a high-intensity functional exercise program on activities of daily living: a randomized controlled trial in residential care facilities. *Journal of the American Geriatrics Society*. 2009 Oct;57(10):1741-9.
- Ljotsson B, Andersson G, Andersson E, Hedman E, Lindfors P, Andriewitch S, et al. Acceptability, effectiveness, and cost-effectiveness of internet-based exposure treatment for irritable bowel syndrome in a clinical sample: a randomized controlled trial. *BMC gastroenterology*. 2011;11:110.
- Ljotsson B, Falk L, Wibron Vesterlund A, Hedman E, Lindfors P, Ruck C, et al. Internet-delivered exposure and mindfulness based therapy for irritable bowel syndrome - A randomized trial. *Behaviour research and therapy*. 2010;48:531-9.
- Ljotsson B, Hedman E, Andersson E, Hesser H, Lindfors P, Hursti T, et al. Internet-delivered exposure-based treatment vs. stress management for irritable bowel syndrome: a randomized trial. *The American journal of gastroenterology*. 2011 Aug;106(8):1481-91.
- Ljotsson B, Lundin C, Mitsell K, Carlbring P, Ramklint M, Ghaderi A. Remote treatment of bulimia nervosa and binge eating disorder: a randomized trial of Internet-assisted cognitive behavioural therapy. *Behaviour research and therapy*. 2007 Apr;45(4):649-61.
- Lundberg, M, Andersson, G, Lunner, T. A randomized, controlled trial of the short-term effects of complementing an educational program for hearing aid users with telephone consultations. *Journal of the American Academy of Audiology*. 2011; 22:654-62.

-
- Marhold C, Linton SJ, Melin L. A cognitive-behavioral return-to-work program: effects on pain patients with a history of long-term versus short-term sick leave. *Pain*. 2001 Mar;91(1-2):155-63.
- Mörtberg E, Clark D, Sundin O, Åberg Wistedt A. Intensive group cognitive treatment and individual cognitive therapy vs. treatment as usual in social phobia: a randomized controlled trial. *Acta psychiatrica Scandinavica*. 2007 Feb;115(2):142-54.
- Mörtberg E, Karlsson A, Fyring C, Sundin Ö. Intensive cognitive-behavioral group treatment (CBGT) of social phobia: A randomized controlled study. *Anxiety Disorders*. 2006;20:646-60.
- Nohlert E, Tegelberg A, Tillgren P, Johansson P, Rosenblad A, Helgason AR. Comparison of a high and a low intensity smoking cessation intervention in a dentistry setting in Sweden: a randomized trial. *BMC Public Health*. 2009;9:121.
- Nordin S, Carlbring P, Cuijpers P, Andersson G. Expanding the limits of bibliotherapy for panic disorder: randomized trial of self-help without support but with a clear deadline. *Behav Ther*. 2010 Sep;41(3):267-76.
- Ollendick TH, Öst LG, Reuterskiöld L, Costa N, Cederlund R, Sirbu C, et al. One-session treatment of specific phobias in youth: a randomized clinical trial in the United States and Sweden. *Journal of consulting and clinical psychology*. 2009 Jun;77(3):504-16.
- Overmeer T, Boersma K, Denison E, Linton SJ. Does teaching physical therapists to deliver a biopsychosocial treatment program result in better patient outcomes? A randomized controlled trial. *Physical therapy*. 2011 May;91(5):804-19.
- Paunovic N. Exposure Inhibition therapy as a treatment for chronic posttraumatic stress disorder: a controlled pilot study. *Psychology*. 2011;2(6):605-14.
- Paunovic N, Öst LG. Cognitive-behavior therapy vs exposure therapy in the treatment of PTSD in refugees. *Behaviour research and therapy*. 2001 Oct;39(10):1183-97.
- Paxling B, Almlöv J, Dahlin M, Carlbring P, Breitholtz E, Eriksson T, et al. Guided internet-delivered cognitive behavior therapy for generalized anxiety disorder: a randomized controlled trial. *Cogn Behav Ther*. 2011;40(3):159-73.
- Priebe S, McCabe R, Bullenkamp J, Hansson L, Lauber C, Martinez-Leal R, et al. Structured patient-clinician communication and 1-year outcome in community mental healthcare: cluster randomised controlled trial. *The British journal of psychiatry : the journal of mental science*. 2007 Nov;191:420-6.
- Rastad C, Ulfberg J, Lindberg P. Light room therapy effective in mild forms of seasonal affective disorder--a randomised controlled study. *Journal of affective disorders*. 2008 Jun;108(3):291-6.
- Rehnman C, Larsson J, Andreasson S. The beer campaign in Stockholm--Attempting to restrict the availability of alcohol to young people. *Alcohol*. 2005 Oct;37(2):65-71.
- Rosendahl E, Lindelöf N, Littbrand H, Yifter-Lindgren E, Lundin-Olsson L, Haglin L, et al. High-intensity functional exercise program and protein-enriched energy supplement for older persons dependent in activities of daily living: a randomised controlled trial. *The Australian journal of physiotherapy*. 2006;52(2):105-13.
- Sahlen KG, Dahlgren L, Hellner BM, Stenlund H, Lindholm L. Preventive home visits postpone mortality--a controlled trial with time-limited results. *BMC Public Health*. 2006;6:220.
- Salomonsson B, Sandell R. A randomized controlled trial of mother-infant psychoanalytic treatment: I. Outcomes on self-report questionnaires and external ratings. *Infant Mental Health Journal*. 2011;32(2):207-31.
- Sandahl C, Herlitz K, Ahlin G, Rönnerberg S. Time-limited group psychotherapy for moderately alcohol dependent patients: A randomized controlled clinical trial. *Psychotherapy Research*. 1998;8(4):361-78.
- Sandahl C, Lundberg U, Lindgren A, Rylander G, Herlofson J, Nygren A, et al. Two forms of group therapy and individual treatment of work-related depression: a one-year follow-up study. *International journal of group psychotherapy*. 2011 Oct;61(4):539-55.

-
- Sapir-Weise R, Berglund M, Frank A, Kristenson H. Acupuncture in alcoholism treatment: a randomized out-patient study. *Alcohol and alcoholism (Oxford, Oxfordshire)*. 1999 Jul-Aug;34(4):629-35.
- Silfvernagel K, Carlbring P, Kabo J, Edström S, Eriksson J, Manson L, et al. Individually tailored internet-based treatment for young adults and adults with panic attacks: randomized controlled trial. *Journal of medical Internet research*. 2012;14(3):e65.
- Sinadinovic K, Wennberg P, Berman AH. Targeting problematic users of illicit drugs with Internet-based screening and brief intervention: a randomized controlled trial. *Drug and alcohol dependence*. 2012 Nov 1;126(1-2):42-50.
- Sinadinovic, K, Berman, A, Hasson, D, Wennberg, P. Internet-based assessment and self-monitoring of problematic alcohol and drug use. *Addictive behaviors*. 2010; 35:464–70.
- Ström, L, Pettersson, R, Andersson, G. Internet-based treatment for insomnia: A controlled evaluation. *Journal of consulting and clinical psychology*. 2004; 72:113-20.
- Ström, L, Pettersson, R, Andersson, G. A controlled trial of self-help treatment of recurrent headache conducted via the internet. *Journal of consulting and clinical psychology*. 2000; 68:722-7.
- Ståhlbrandt H, Johnsson K, Berglund M. Two-year outcome of alcohol interventions in Swedish university halls of residence: a cluster randomized trial of a brief skills training program, twelve-step-influenced intervention, and controls. *Alcoholism, clinical and experimental research*. 2007 Mar;31(3):458-66.
- Sundell K, Vinnerljung B. Outcomes of family group conferencing in Sweden. A three year follow up. *Child Abuse & Neglect*. 2004;28:267-86.
- Sundin O, Lisspers J, Hofman-Bang C, Nygren A, Ryden L, Ohman A. Comparing multifactorial lifestyle interventions and stress management in coronary risk reduction. *International journal of behavioral medicine*. 2003;10(3):191-204.
- Terjestam Y, Jouper J, Johansson C. Effects of scheduled qigong exercise on pupils' well-being, self-image, distress, and stress. *Journal of alternative and complementary medicine*. 2010 Sep;16(9):939-44.
- Thorell LB. The Community Parent Education Program (COPE): treatment effects in a clinical and a community-based sample. *Clinical child psychology and psychiatry*. 2009 Jul;14(3):373-87.
- Thorén, E, Svensson, M, Törnqvist, A, Andersson, G, Carlbring, P, Lunner, T. Rehabilitative online education vs. internet discussion group for hearing aid users: A randomized controlled trial. *Journal of the American Academy of Audiology*. 2011; 22:274-85.
- Thyme KE, Sundin EC, Wiberg B, Öster I, Åstrom S, Lindh J. Individual brief art therapy can be helpful for women with breast cancer: a randomized controlled clinical study. *Palliative & supportive care*. 2009 Mar;7(1):87-95.
- Tiberg I, Katarina SC, Carlsson A, Hallström I. Children diagnosed with type 1 diabetes: a randomized controlled trial comparing hospital versus home-based care. *Acta paediatrica (Oslo, Norway : 1992)*. 2012 Oct;101(10):1069-73.
- Tillfors M, Andersson G, Ekselius L, Furmark T, Lewenhaupt S, Karlsson A, et al. A Randomized Trial of Internet-Delivered Treatment for Social Anxiety Disorder in High School Students. *Cognitive Behaviour Therapy*. [Article]. 2011;40(2):147-57.
- Tillfors, M, Carlbring, P, Furmark, T, Lewenhaupt, S, Spak, M, Eriksson, A, et al. Treating university students with social phobia and public speaking fears: Internet delivered self-help with or without live group exposure sessions *Depression and Anxiety*. 2008; 25:708-17.
- Tomson Y, Romelsjö A, Aberg H. Excessive drinking--brief intervention by a primary health care nurse. A randomized controlled trial. *Scandinavian journal of primary health care*. 1998 Sep;16(3):188-92.

-
- Wallin E, Andreasson S. Can I have a beer, please? A study of alcohol service to young adults on licensed premises in Stockholm. *Prevention science : the official journal of the Society for Prevention Research*. 2004 Dec;5(4):221-9.
- Vernmark K, Lenndin J, Bjarehed J, Carlsson M, Karlsson J, Öberg J, et al. Internet administered guided self-help versus individualized e-mail therapy: A randomized trial of two versions of CBT for major depression. *Behaviour research and therapy*. 2010 May;48(5):368-76.
- Westman A, Linton SJ, Öhrvik J, Wahlen P, Theorell T, Leppert J. Controlled 3-year follow-up of a multidisciplinary pain rehabilitation program in primary health care. *Disability and rehabilitation*. 2010;32(4):307-16.
- von Thiele Schwarz U, Lindfors P, Lundberg U. Health-related effects of worksite interventions involving physical exercise and reduced workhours. *Scandinavian journal of work, environment & health*. 2008 Jun;34(3):179-88.
- Zetterlind U, Hansson H, Åberg-Örbeck K, Berglund M. Effects of coping skills training, group support, and information for spouses of alcoholics: a controlled randomized study. *Nord J Psychiatry*. 2001;55(4):257-62.
- Zetterqvist K, Maanmies J, Ström L, Andersson G. Randomized controlled trial of internet-based stress management. *Cogn Behav Ther*. 2003;32(3):151-60.
- Zetterqvist Westin, V, Schulin, M, Hesser, H, Karlsson, M, Zare Noe, R, Olofsson, U, Stalby, M, et al. Acceptance and Commitment Therapy versus Tinnitus Retraining Therapy in the treatment of tinnitus distress: a randomized controlled trial. *Behaviour research and therapy*. (49):737-47.
- Öberg, M, Andersson, G, Wänström, G, Lunner, T. The effects of a sound awareness pre-fitting intervention: a randomized controlled trial. *Audiological Medicine*. 2008; 6:129-40.
- Öberg, M, Andersson, G, Wänström, G, Lunner, T. The effects of a pre-fitting intervention on hearing aid benefit: a randomized controlled trial. *Audiological Medicine*. 2009; 7:211-25.
- Öjehagen A, Berglund M, Appel CP, Andersson K, Nilsson B, Skjaeris A, et al. A randomized study of long-term out-patient treatment in alcoholics. *Psychiatric treatment versus multimodal behavioural therapy, during 1 versus 2 years of treatment*. *Alcohol and alcoholism*. 1992 Nov;27(6):649-58.
- Öst LG, Breitholtz E. Applied relaxation vs. cognitive therapy in the treatment of generalized anxiety disorder. *Behaviour research and therapy*. 2000 Aug;38(8):777-90.
- Öst LG, Svensson L, Hellström K, Lindwall R. One-Session treatment of specific phobias in youths: a randomized clinical trial. *Journal of consulting and clinical psychology*. 2001 Oct;69(5):814-24.
- Öst LG, Thulin U, Ramnerö J. Cognitive behavior therapy vs exposure in vivo in the treatment of panic disorder with agoraphobia (corrected from agrophobia). *Behaviour research and therapy*. 2004 Oct;42(10):1105-27.
- Öst LG, Westling BE. Applied relaxation vs cognitive behavior therapy in the treatment of panic disorder. *Behaviour research and therapy*. 1995 Feb;33(2):145-58.

Bilaga 2. Kvalitetsaspekter fördelat på publikationsår (%)

	Publikationsår:					stat sign
	-00 (n=18)	01-03 (n=23)	04-06 (n=30)	07-09 (n=38)	10-12 (n=60)	
<i>Transparens</i>						
Forskningsprotokoll publicerat	0	4	3	5	22	.007
Redovisning av ev. Intressekonflikt	6	9	21	26	48	.0003
<i>Undersökningsgrupp</i>						
Inklusions- och exklusionskriterier	84	87	83	92	97	-
Demografiska och kliniska variabler	50	57	66	72	73	-
Beräkning av gruppstorlek (power)	0	30	17	23	33	.05
<i>Datainsamling</i>						
Primärt utfallsmått	0	26	21	38	57	.0001
Psykometri bedömningsinstrument	44	52	55	67	65	-
"Blindad" bedömare	11	39	24	18	17	-
<i>Intervention</i>						
Omfattning av intervention	56	52	69	59	65	-
<i>Design</i>						
Randomiserad	88	96	79	82	93	-
<i>Vid randomisering</i>						
Randomiseringssekvens	13	41	48	75	63	.0004
Typ av randomisering	19	50	39	56	50	-
Dold fördelning till betingelser	19	32	35	63	59	.005
<i>Vid icke-randomisering</i>						
Fördelningsprincip	0	100	33	71	50	-
Försök att minska selektionsbias	50	0	67	43	50	-
<i>Analys</i>						
Fördelade personer per grupp	50	87	86	84	98	.0001
Bortfall per grupp	61	91	62	74	88	.008
Analyserade personer	67	91	83	92	93	.03
Intent-to-treat	17	48	41	69	70	.0002
Imputering	0	9	7	23	22	-
Statistiska metoder för justering	44	70	59	77	68	-
Statistisk justering för baslinjeskillnader (totalt)	22	30	24	23	25	-
Statistisk justering för baslinjeskillnader (grupp)	6	0	0	5	5	-
Statistisk justering vid klusterrandomisering	6	4	7	13	5	-
<i>Resultat</i>						
Beskrivande statistik	61	78	66	74	88	-
Signifikant effekt på primärt utfall	39	61	52	51	63	-
Effektstorlek	6	17	41	54	67	.0001
Eventuella oönskade utfall	6	4	17	10	20	-

Bilaga 3. Kvalitetsaspekter fördelat på startår (%)

	Studiens startår ¹⁹ :					stat sign
	74-95 (n=21)	96-99 (n=27)	00-03 (n=47)	04-07 (n=41)	08-11 (n=31)	
<i>Transparens</i>						
Forskningsprotokoll publicerat	0	7	4	10	26	.01
Redovisning av ev. Intressekonflikt	14	19	19	44	42	.02
<i>Undersökningsgrupp</i>						
Inklusions- och exklusionskriterier	95	85	87	93	100	-
Demografiska och kliniska variabler	57	56	68	58	90	.03
Beräkning av gruppstorlek (power)	10	15	19	32	39	.05
<i>Datainsamling</i>						
Primärt utfallsmått	14	15	26	46	71	.0001
Psykometri bedömningsinstrument	52	56	57	66	65	-
"Blindad" bedömare	14	33	23	12	19	-
<i>Intervention</i>						
Omfattning av intervention	62	48	55	71	71	-
<i>Design</i>						
Randomiserad	86	67	87	95	100	.001
<i>Vid randomisering</i>						
Randomiseringssekvens	17	39	56	69	65	.002
Typ av randomisering	17	44	56	46	52	-
Dold fördelning till betingelser	17	28	46	51	74	.0009
<i>Vid icke-randomisering</i>						
Fördelningsprincip beskrivet	33	33	67	100	-	-
Försök att minska selektionsbias	67	44	50	50	-	-
<i>Analys</i>						
Fördelade personer per grupp	62	81	89	93	97	.003
Bortfall per grupp	67	70	74	78	97	-
Analyserade personer	71	78	85	98	100	.003
Intent-to-treat	33	37	57	68	74	.005
Imputering	5	0	17	29	13	.01
Justerat statistiskt för initiala baslinjeskillnader	43	59	72	78	68	-
Undersökt skillnader mellan bortfall (totalt)	0	37	32	22	26	.04
Undersökt skillnader mellan bortfall (gruppvis)	0	4	4	5	3	-
Statistisk justering vid klusterrandomisering	5	4	13	7	0	-
<i>Resultat</i>						
Beskrivande statistik	52	63	81	80	97	.001
Signifikant effekt på primärt utfall	43	41	57	51	77	.04
Effektstorlek	14	7	53	61	71	.0001
Eventuella oönskade utfall	14	4	11	12	29	-

¹⁹ För två utvärderingar saknas information om startår.