



Föreningen Sveriges Habiliteringschefer

Rikstäckande nätverk för barn- och ungdomshabiliteringen i Sverige. Grundad 1994

2

Evidensvärdering

Gunilla Bohlin

Ur rapporten

Mångsidiga intensiva insatser
för barn med autism i förskoleåldern

Sammanfattning

- Begreppet evidensbaserad medicin (EBM) eller habilitering innebär att man fattar beslut om insatser och behandling utifrån ”bästa tillgängliga vetenskap, sammanställd på ett systematiskt sätt”. Viktiga grundprinciper för evidensbaserad medicin eller habilitering är:
 - att god vård bygger på objektiva data
 - att det finns en evidenshierarki, alltså att vissa bevis är starkare än andra
 - att enbart vetenskapliga data inte är tillräckliga för god vård, utan att klinisk bedömning måste integreras.
- Systematiska översikter utgör en utvecklad metodik för att på ett vetenskapligt, uttömmande och systematiskt sätt identifiera, värdera och sammanställa kliniska studier inom ett specifikt område för att kunna dra så starka och välunderbyggda slutsatser som möjligt om en insats effekt.
- Kliniska studier kan beroende på hur de är upplagda och genomförda ge olika säker kunskap om hur effektiv en insats är. Denna kunskap kan graderas hierarkiskt. Starkaste bevis ger sådana upplägg som effektivast minimerar olika förvrängande faktorer (bias) inverkan på resultatet.
- Vanliga felkällor och förvrängande faktorer är exempelvis att en förbättring sker helt oberoende av insatsen, att den som bedömer förändringen är påverkad av sina egna förväntningar om insatsens effektivitet, att urvalet av dem som undersöks inte är representativt eller att den som ger insatsen inte följer programmet som undersöks utan lägger till egna moment.
- Sådant som kontrollgrupp, blinda bedömare (=som är ovetande om vilken insats den undersökta personen fått), att använda kvantitativa mått som analyseras statistiskt, att genomföra både före- och eftermätningar och att man har uttalad kontroll över att det tänkta behandlingsupplägget följs är sätt att minimera förvrängande faktorer inverkan på resultatet.
- Randomiserade kontrollerade studier där deltagarna slumpmässigt fördelas till experiment- och kontrollgrupp är högst i hierarkin.
- Klinisk erfarenhet där var och en enbart förlitar sig till sina egna erfarenheter och konsensus där experter kommer överens om vad de anser vara bäst (med otillräcklig vetenskapligt underlag) är lägst i hierarkin.
- För att formulera rekommendationer är vetenskapliga data inte tillräckliga utan dessa måste integreras med bland annat klinisk bedömning och erfarenhet, etiska överväganden samt samhällliga, ekonomiska, kulturella och kompetensmässiga aspekter.

Principer vid evidensvärdering

När man ställer krav på att åtgärder inom vård och omsorg ska vara "evidensbaserade" innebär det att man menar att det ska finnas skäl att tro att åtgärderna har en positiv effekt. Att sådana förväntningar om en positiv effekt måste grundas på systematisk kunskap har blivit alltmer klart sedan man kunnat visa att vår vardagliga kunskapsbildning, den som vi får genom egen erfarenhet, generellt är utsatt för en rad förvrängande faktorer. Så har vi t ex en tendens att minnas det som fungerat, snarare än det som inte fungerat, att söka efter information som bekräftar snarare än ifrågasätter våra idéer, och i beslutssituationer har vi en benägenhet att överskatta hur säker information är.

Evidensvärdering av behandlingsmetoder tillämpar en uppsättning principer som ska vara till hjälp när man vill dra slutsatser om ett behandlingsprogramms effektivitet. Det man vill veta är om de personer som behandlats inom programmet har förbättrats under behandlingens gång och i så fall om det är själva behandlingen som orsakat detta. Man vill alltså dra slutsatser om orsaken till en eventuell förändring - inte bara beskriva den. Även inom ramen för mer systematisk kunskapsbildning finns en uppsättning möjliga felkällor som hotar att leda till att man drar felaktiga slutsatser och principerna för evidensvärdering ska fungera som vägledning för att öka möjligheten att dra riktiga slutsatser. Några av de vanligaste felkällorna är att en förbättring sker helt oberoende av behandlingen, s k självläkning, att den som bedömer förbättringen har förväntningar om vilken effekten ska bli och påverkas av detta i sina bedömningar, att de behandlade personerna inte är typiska för den kategori klienter som behandlingen är avsedd för, och att behandlaren inte följer programmet utan inkluderar egna moment.

Bland värderingsprinciperna utgör den experimentella metoden "the golden standard" (Rossi, 1997), d v s det är den som har det tyngsta bevisvärdet. Några grunddrag i denna metod är

1. att antaganden om behandlingseffekter och mekanismer grundas i teori
2. att de som ska behandlas fördelas slumpmässigt till experiment- och kontrollgrupp
3. att experimentledaren kontrollerar interventionen så att den genomförs på ett standardiserat sätt
4. att effekter mäts kvantitativt och analyseras statistiskt för att avgöra om effekter är säkerställda, d v s sannolikt skulle uppträda igen i en likadan studie.

I behandlingsstudier inom medicin tillkommer för den högsta evidensnivån att behandlingen är dubbel-blind, d.v.s. varken behandlare eller patient vet vilken grupp som får behandling med placebo. Detta krav kan naturligtvis vara omöjligt att uppfylla i psykologiskt inriktade behandlingsstudier.

Att den experimentella metoden har det högsta bevisvärdet innebär inte att studier som gjorts med andra metoder är helt utan bevisvärde. Ett relativt sett högre bevisvärde bör tillmätas icke-experimentella studier som har en kontrollgrupp som på något sätt visats initialt vara likvärdig med behandlingsgruppen, studier som har både för- och eftermätning och som använder ett beprövat kvantitativt utvärderingsinstrument. Ett lägre bevisvärde har studier där fördelningen till behandlings- och jämförelsegrupp helt tillkommit via självselektion (vanligast när man jämför två behandlingar) och studier som saknar kontrollgrupp och i stället förlitar sig till att jämföra för- och eftermätning i behandlingsgruppen. Klinisk erfarenhet och enstaka fallstudier har ett lågt bevisvärde, men i en situation där kunskapsnivån inom området är låg är även sådana studier av värde och deras evidens värderas efter hur välgjord och

systematiserad beskrivningen är vad gäller de behandlade personerna, typ av initiala problem, procedur för behandlingen och utfall.

Sammanfattningsvis kan man ställa upp några ”kvalitetsindikatorer” som kan fungera som vägledning när man bedömer hur bra kunskap man har om effekterna av ett visst behandlingsprogram:

1. Det finns studier med kontrollgrupp – åtminstone ett par studier där fördelning av personer på behandlings- och kontrollgrupp skett slumpmässigt.
2. Både för- och eftermätningar finns inkluderade i studier.
3. Man har haft kontroll över hur behandlingen har gått till.
4. De utvärderingsinstrument som använts är meningsfulla.
5. Det finns långtidsuppföljning av behandlingseffekter.
6. Behandlingseffekterna har kunnat påvisas i studier i klinisk praktik.

Statens beredning för medicinsk utvärdering (SBU) har gjort en indelning i 4 olika grader av evidens enligt följande:

Om flera stora studier, från olika centra och med en för frågan lämplig design och högt bevisvärde, givit samma resultat blir det vetenskapliga underlaget starkt och slutsatsen pålitlig (*evidensstyrka 1*). Det gäller även då det finns en god systematisk översikt av flera välgjorda studier.

Om endast en studie med högt bevisvärde finns, men resultaten stöds av flera studier med medelhögt bevisvärde, blir evidensstyrkan måttligt stark (*evidensstyrka 2*).

Begränsad styrka (*evidensstyrka 3*) föreligger när ingen studie med högt bevisvärde men minst två med medelhögt bevisvärde finns.

Under denna nivå är det vetenskapliga underlaget otillräckligt; man måste ange vad eventuella slutsatser i så fall baserar sig på (*evidensstyrka 4*).

Förutsättningen för det angivna är att studierna pekar i samma riktning och att inget talar emot slutsatsen. Vid mindre divergenser kan styrkegraden sänkas.

Författare MONA BRITTON, professor, medicinskt sakkunnig, SBU, Stockholm
Utdrag ur Läkartidningen 2000;97:4414-15
www.sbu.se

Särskilda riktlinjer för värdering av metoder för psykologisk intervention med barn har givits av American Psychological Association (APA). De följer i huvudsak de allmänna riktlinjerna och ger vägledning för vad som kan betraktas som väletablerade metoder och de som kan ses som ”troligen effektiva” metoder:

Kriterier för väletablerade interventioner

1. Åtminstone två välgjorda gruppstudier, utförda av olika forskarteam som visar att interventionen antingen är
 - a. bättre än pillerplacebo eller en alternativ intervention
 - b. likvärdig med en redan etablerad intervention i studier med adekvat statistisk power.

ELLER

2. Ett stort antal ”single-case”-studier (det vill säga, $n > 9$) som både

- a. använder god experimentell design OCH
- b. jämför interventionen med en annan intervention.

OCH

3. Att manual för interventionen helst används.

OCH

4. Att deltagarnas karaktäristika tydligt redovisas.

Kriterier för troligen effektiva (probably efficacious) interventioner

1. Två studier som visar att interventionen är effektivare än en kontrollgrupp som inte får någon intervention (till exempel vänteliste-kontrollgrupp).

ELLER

2. Två gruppstudier som uppfyller kriterierna för en väletablerad intervention men är utförda av samma forskargrupp.

ELLER

3. Ett mindre antal "single-case"-studier (det vill säga, $n > 3$) som i övrigt uppfyller kriterium 2 för väletablerade interventioner.

OCH

4. Att manual för interventionen helst används.

OCH

5. Att deltagarnas karaktäristika tydligt redovisas.

Absoluta kriterier:

Randomisering i gruppstudier.

Adekvat statistisk power.

Källa: Lonigan, 1998
Översättning: Eric Zander

Systematiska översikter och evidensbaserad medicin (EBM)

Det synsätt som kallas för evidensbaserad medicin (EBM) handlar alltså enkelt uttryckt om att fatta beslut om behandling, insatser och liknande utifrån "bästa tillgängliga vetenskap, sammanställd på ett systematiskt sätt" (Eliasson, 2000). De viktigaste grundprinciper är att:

1. God vård bygger på objektiva data, vilket medför:
 - a. behov av kliniskt relevant info.
 - b. att nya rön förändrar vården.
 - c. ett livslångt lärande för berörda.
2. Det finns en evidenshierarki, alltså att vissa bevis starkare än andra.
3. Enbart vetenskapliga data inte är tillräckliga för god vård, utan att data måste integreras med kliniskt omdöme.

(Nordenström, 2004).

För att få fatt i ”bästa tillgängliga vetenskap” inom ett specifikt område krävs att man på ett systematiskt sätt går igenom totaliteten av det kunskapsunderlag som föreligger inom området ifråga. I det här avseendet spelar den systematiska översikten (systematic review) en väsentlig roll (Eliasson, 2000). I den systematiska översikten utgår man från en utvecklad metodik (Centre for Reviews and Dissemination (CRD), 2001; Clarke & Oxman, 2003) där kunskapsunderlaget i form av primärstudier, t ex kliniska (utfalls)studier om tidiga insatser för barn med autism, identifieras, värderas och beskrivs på ett vetenskapligt, systematiskt, uttömmande och genomskinligt vis på många nivåer. Syftet är att göra det tydligt hur man gått tillväga och resonerat i översikterna samt göra det möjligt för andra att reproducera undersökningen, precis som när det gäller annan vetenskaplig produktion. En systematisk översikt skiljer sig så tillvida från de traditionella översikterna (narrative reviews) som oftast inte är uttömmande, innehåller systematiska fel (bias) av olika slag och, har det visat sig, därmed ofta överskattar eller i alla fall missbedömer effekterna (Antman *et al.*, 1992; Kunz & Oxman, 1998). En systematisk översikt kan sägas vara vetenskap på metanivå eller vetenskap om vetenskap.

För att kunna dra tillförlitliga slutsatser utifrån en systematisk översikt är man naturligtvis beroende av att ingående data i form av primärstudier är av hög kvalitet: styrkan i slutsatserna är direkt avhängig primärstudiernas kvalitet. I det här hänseendet ger systematiska översikter som utgår från flera välgjorda, randomiserade, dubbelblinda kliniska studier av högsta kvalitet det bästa underlaget (starkast evidens), därför att det har visat sig att studier med sådan design, till skillnad från andra typer av design, kontrollerar för systematiska fel på ett effektivt sätt (Harrington *et al.*, 2002). Man har visat att systematiska översikter av icke-randomiserade studier eller inte tillräckligt välgjorda randomiserade studier är problematiska; de missbedömer ofta grovt effekten vid behandlingar på ett systematiskt sätt, oftast åt det positiva hållet (Clarke & Oxman, 2003; Kunz & Oxman, 1998; Moher *et al.*, 1998; NRSMG, 2001). I många fall finns dock endast icke-randomiserade studier att tillgå och då blir följaktligen även slutsatserna om effekt svagare/svårare att dra.

Systematiska översikter finner man huvudsakligen på tre håll (Eliasson, 2000):

1. I Cochranebiblioteket som är en internationell organisation som bygger på att genom samarbete samla ihop, värdera och presentera den bästa befintliga vetenskapen inom olika medicinska/hälsoområden (www.cochrane.com; Gotszsche, 2000). De metodologiska bedömningarna utförs som regel av flera forskare oberoende av varandra, och uppdateras kontinuerligt utifrån nya data och andras kritik. Dessa översikter finns tillgängliga on-line på Internet.
2. I olika länders organisationer för medicinsk utvärdering (Health Technology Assessment, HTA), som vår egen Statens beredning för medicinsk utvärdering (SBU).
3. I referentbedömda tidskrifter.

Tio kriterier på god vetenskaplig metodik i en systematisk översikt

1. Inriktad på en fokuserad fråga
2. Metoden för litteratursökning väl beskriven
3. Explicita kriterier för inklusion och exklusion av studier
4. Ingående studiers metodik kvalitetsbedömd
5. Reproducerbar granskning och bedömning av originalstudier
6. Framtida forskningsbehov utpekade
7. Motstridiga resultat (heterogenitet) utforskat och tolkat
8. Kvantitativ syntes utförd där så är lämpligt och möjligt (metaanalys)
9. Viktiga patientrelaterade effektmått studerade
10. Fyndens generaliserbarhet diskuterad (Eliasson, 2000)

Det finns olika sätt att mer exakt gradera evidens (Clarke & Oxman, 2003; Lonigan *et al.*, 1998). Ofta handlar det om att man graderar evidensstyrkan i sina slutsatser i en fyragradig skala eller liknande, men där de exakta formuleringarna för varje grad kan variera (se bilaga 4).

Nästa steg i EBM-processen är att formulera rekommendationer. I den delen av processen har inte bara kunskapsläget betydelse utan även andra, ickevetenskapliga faktorer måste vägas in, t ex etiska överväganden, samhällsstruktur, kulturella aspekter, ekonomi, kompetens, etc (Chorpita, 2002; Chorpita *et al.*, 2002; Rehnqvist, 2001).

Även om det evidensbaserade synsättet av många anses vara det mest tillförlitliga sättet att värdera behandlingsmetoder och liknande och är en metodik som ständigt utvecklas har kritiska stämmor ibland höjts mot detta synsätt (t ex Stolt, 2001). En del har berott på missförstånd som att endast de åtgärder som det finns evidens för skulle få användas. Med tiden har EBM kommit att innebära integration av ”kliniskt kunnande med bästa tillgängliga underlag som framtagits med hjälp av systematiskt utförd klinisk forskning” (Aspelin, 2004).

Referenser

- Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *Jama*, 268(2), 240-248.
- Aspelin, P. (2004). Förord. In Jörgen Nordenström, *Evidensbaserad medicin i Sherlock Holmes fotspår* (pp. 7). Stockholm: Karolinska University Press.
- Centre for Reviews and Dissemination (CRD). (2001). *Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews*. CRD Report Number 4 (2 ed.). York: York Publishing Services.
- Chorpita, B. F. (2002). Treatment Manuals for the Real World: Where Do We Build Them? *Clin. Psychol.*, 9(4), 431-433.
- Chorpita, B. F., Yim, L. M., Donkervoet, J. C., Arensdorf, A., Amundsen, M. J., McGee, C., et al. (2002). Toward Large-Scale Implementation of Empirically Supported Treatments for Children: A Review and Observations by the Hawaii Empirical Basis to Services Task Force. *Clin. Psychol.*, 9(2), 165-190.
- Clarke, M., & Oxman, A. D. (Eds.). (2003). *Cochrane Reviewers' Handbook 4.2.0 [updated March 2003]*. Oxford: Update.
- Eliasson, M. (2000). Den systematiska översikten grundval i evidensbaserad medicin. Ett av de viktigaste tillskotten till den kliniska medicinen under senaste årtiondet. *Läkartidningen*, 97(22), 2726-2728.
- Gotzsche, P. C. (2000). Cochranesamarbetets betydelse för evidensbaserad medicin. *Läkartidningen*, 97(26-27), 3180-3182.
- Harrington, R. C., Cartwright-Hatton, S., & Stein, A. (2002). Annotation: randomised trials. *J Child Psychol Psychiatry*, 43(6), 695-704.
- Kunz, R., & Oxman, A. D. (1998). The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ*, 317(7167), 1185-1190.

- Lonigan, C. J., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: an overview. *J Clin Child Psychol*, 27(2), 138-145.
- Moher, D., Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., et al. (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*, 352(9128), 609-613.
- Nordenström, J. (2004). *Evidensbaserad medicin i Sherlock Holmes fotspår*. Stockholm: Karolinska University Press.
- NRSMSG. (2001). *Guidelines on Non-randomised studies in Cochrane reviews, Draft chapters*. Retrieved 28/11, 2003, from <http://www.cochrane.dk/nrsmsg/guidelines.htm>
- Rehnqvist, N. (2001). Lång väg från evidens till kliniska rekommendationer. *Läkartidningen*, 98(12), 1355-1356.
- Rossi, P. H. (1997). Advances in quantitative evaluation 1987 – 1996. In D. Rog & D. Fournier (Eds.), *New Directions for Program Evaluation*, 76, 57-68. San Francisco: Jossey-Bass.
- Stolt, C. M. (2001). Evidensbaserad medicin: tio kritiska invändningar mot en värdefull metod. *Läkartidningen*, 98(28-29), 3255-3256.