

Evidensbaseret behandling af bevægeapparatet

Virker den pågældende behandling, og hvornår kan den anbefales?

Robin Christensen, cand. scient., ph.d. stud., forskningsassistent på Parker Institutet, H:S Frederiksberg Hospital og Else Marie Bartels, dr. scient., fysiologi og biofysik, seniorforsker på Københavns Universitetsbibliotek

Christensen R, Bartels EM (2006, 12. oktober). Evidensbaseret behandling af bevægeapparatet – virker den pågældende behandling, og hvornår kan den anbefales? *Forskning i Fysioterapi (online)*, 4. årg.) s. 1-8: URL: <http://www.ffy.dk/sw10577.asp>

Det at behandle en person for en lidelse kan ofte give indtryk af at skabe en positiv effekt af den specifikke behandling, selvom det i praksis meget vel kan dreje sig om effekten af, at der bliver gjort noget aktivt ved en oplevet håbløs situation. Derfor er randomiserede og kontrollerede studier nødvendige, hvis en behandling skal evalueres.

Personer med symptomgivende sygdom i bevægeapparatet vil oftest få stillet en diagnose og efterfølgende blive behandlet med medicin eller fysioterapi for at forbedre tilstanden eller blot vente, til den for behandleren velkendte tilstand går over af sig selv.

Visse lidelser kræver, at patienten selv tager aktiv del i behandlingen for at opnå succes. Således må patientens biomekaniske/somatiske forhold understøtte den lægeordnede behandling. Eksempler på dette er type 2 Diabetes Mellitus (T2DM) og osteoartrose (OA) i vægtbærende led, der begge med fordel kan behandles med teoretisk velunderbygget ernæringsterapi og/eller træning. Karakteristisk for denne slags "sund fornuft" terapi er, at patienten skal anerkende, at det kræver stor egenindsats, og at denne indebærer livslang ændring i den enkeltes hverdag, da der er tale om kroniske lidelser. Begrebet *coping* dækker patientens evne til at acceptere og tage ansvar for denne type livslange ændringer i levevis som en del af behandlingen. Dårlig *coping* medfører erfaringsmæssigt også

dårlig "*compliance*" (det, at man kan affinde sig med sin tilstand). Personer med dårlig *coping*, og dermed følgende dårligere tilstand og respons til anden behandling, er tilbøjelige til at se sig om efter andre behandlingsmuligheder, og ofte vælges her metoder af mere alternativ karakter.

Der synes at være en ide i store dele af befolkningen om, at læger (og anerkendte behandlere i det hele taget) er konservative og ikke er åbne for nye lovende behandlinger, "*som alle ved virker*". Hvis kun disse behandlinger var uden effekt, ville de ikke kunne gøre skade, men næsten enhver behandling, der har en effekt, har samtidig en mulighed for at have bivirkninger. Skal man kunne argumentere for og imod forskellige behandlingstyper, er det nødvendig at kende til evidensen for behandlingerne.

Formålet med denne artikel er at belyse, hvad begrebet evidens dækker, og efterfølgende at give konkrete rettesnore til, hvornår og hvorvidt nye og sparsomt dokumenterede behandlinger kan anbefales som et led i den daglige klinik. Det er således et mål med nærværende artikel, at fysioterapeuter, der er involveret i forskning eller i beslutninger omkring valg af behandlingstyper, får redskaber til at tage aktivt del i debatten og på baggrund af matematisk-statistiske beregninger kan fortolke og præsentere argumenter for og imod behandlinger.

Evidensbaseret medicin

Evidensbaseret medicin tager udgangspunkt i veldefinerede kliniske forsøg. Disse studier følger en defineret international standard, der medfører en høj kvalitet. Resultatet af sådanne studier bliver i evidensbaseret behandling sammenholdt med klinikerens erfaring og patientens ønsker. Det er vigtigt at tage med, at begrebet evidensbaseret medicin ikke kun vedrører effektivitet af diverse behandlinger til "korrekt" diagnosticerede sygdomme, men også gælder korrekt diagnostik, prognose og afvikling af klinik i bred forstand, herunder også økonomiske overvejelser.

Evidensbaseret medicin har tidligere været gennemgående beskrevet i 'Ugeskrift for Læger' af Tom Pedersen m.fl. der konkluderede: *"Brugen af evidensbaserede kliniske retningslinjer i en klinisk integreret forløbsbeskrivelse vil give en mere omkostningseffektiv behandling. Det burde være en national strategi, at sundhedsvæsenet skal være evidensbaseret"* (1).

Når det skal vurderes, om en behandling virker (eller har bivirkninger), er den bedste evidens den, der stammer fra veludførte lodtrækningsundersøgelser, hvor patientpræferencer og personlige coping-evner antages at være sat ud af spil. Disse studier benævnes randomiserede (tilfældig fordeling af de deltagende personer) kontrollerede (behandlingsgruppen måles mod en kontrolgruppe) undersøgelser (på engelsk: randomised controlled trials, RCT). Denne klasse af studier indeholder såkaldt kategori 1b evidens (2).

Når evidensen systematisk kombineres for alle disse RCT'ere med hensyn til en enkelt lidelse (f.eks. osteoartrose) og et enkelt behandlingstiltag (f.eks. glukosamin), så er der tale om et systematisk *review*. Den umiddelbare opfølgning hertil er kvantitativt at summere den samlede evidens - med de dertil udviklede formler. Når man kvantitativt samler de tidligere studiers resultater, er der tale om en metaanalyse (på engelsk: meta-analysis).

En metaanalyse giver som resultat en overordnet behandlingseffekt, der er både sagligt og kvantitativt velbegrundet. Således er det ypperligste niveau af evidens (kategori 1a) det, der stammer fra et systematisk review (2), hvor dette bliver understøttet af en metaanalyse.

I slutningen af 2004 udkom bogen *"Evidence-Based Rheumatology"* (3), der beskriver en struktur for, hvorledes man i *"The Cochrane Musculoskeletal Group"* (CMSG) vil have kategoriseret evidens i forbindelse med systematiske reviews omkring behandlinger af lidelser i bevægeapparatet. Rangordningen i denne struktur er forsøgt gengivet på dansk i tabel 1, side 3.

Evidensfortolkning

Vurdering, anbefaling og implementering

Uanset hvordan man vender og drejer begrebet evidens, så er der én ting, der er gældende: Patienten skal erfare en additiv effekt. Denne effekt skal både økonomisk og i form af tålelige bivirkninger stå mål med, hvor meget patienten får ud af indsatsen.

Det er en udbredt misforståelse, at kliniske effekter, der ikke er eftervist i et RCT, er utilregnelige. Et RCT belyser alene, om der er en "ekstra bonus" ved at tilføje denne ekstra behandling, hvis man sammenligner med kontrolinterventionen. Det er af samme grund, at den behandling, der benyttes i kontrolgruppen, skal være repræsentativ for *"best clinical practice"*. Det er den ekstra gevinst af den tilføjede behandling, et randomiseret studie kan og sigter på at belyse. Et resultat heraf er, at det kun medfører "bronze-evidens", når der i en ikke-kontrolleret undersøgelse er fundet statistisk signifikant effekt ($P < 5\%$) ved sammenligning af før- og efterværdien af en valgt parameter for behandlingen (typisk analyseret ved parret t-test).

Vurdering

Det enkelte studies kvalitet

Skal vi vurdere, hvorvidt et nyt behandlingstiltag kan anbefales, kan man ikke regne med, at der er lavet et systematisk *review*, eventuelt med en metaanalyse, da der som oftest ikke findes ret mange randomiserede undersøgelser af effekten af den nye behandlingsstrategi. Derfor må man om nødvendigt kunne uddrage en kritisk konklusion af blot én enkelt (publiceret) randomiseret undersøgelse, således at både evidensens kvalitet og effektiviteten af interventionen fremstår utvetydigt.

Læsere af videnskabelige artikler og kliniske forskere er nødt til at have et indgående kendskab til læsning og skrivning af randomiserede inter-

Table 1: Rangordning af evidensbaseret reumatologi

PLATIN evidens

Platinniveau af evidens er forbeholdt interventioner der har været underlagt følgende metodiske kriterier:

- Publiceret** Som minimum indeholder dette resultat fra mindst 2 uafhængige RCT der opfylder:
- Systematisk review:**
- Mindst 100 personer er randomiseret til 2 behandlinger (dvs. minimum 50 pr. gruppe). Det vil indebære en tilstrækkelig styrke til at kunne detektere en 20% relativ forskel.
 - Gensidig blinding af deltagere, hvor muligt, (således at både patienter og personale, der foretager målingerne, er blandede).
 - Udførlig beskrivelse af hvorledes frafald fra studiet er forekommet (bivirkninger, men også uforklarede forhold). Frafald (drop-out) bør ikke overstige 20%. Der bør være foretaget en intention-to-treat (ITT) overvejelse (beregning).
 - Det skal være belyst, hvorvidt personalet, der inkluderede patienterne, og patienterne selv godkendte deres deltagelse, før de kunne gennemskue hvad de blev randomiseret til (= concealment of treatment allocation) - med efterfølgende ITT ansvar.

GULD evidens

Guldniveau af evidens kræver at interventionen har været underlagt de følgende metodiske kriterier:

- Publiceret:** Som minimum skal dette opfylde følgende kriterier:
- Mindst 100 personer er randomiseret til 2 behandlinger (dvs. minimum 50 pr. gruppe). Det vil indebære en tilstrækkelig styrke til at kunne detektere en 20% relativ forskel.
 - Gensidig blinding af deltagere, hvor muligt, (således at både patienter og personale, der foretager målingerne, er blandede).
 - Udførligt beskrevet hvorledes frafald fra studiet er forekommet (bivirkninger, men også uforklarede forhold). Frafald (drop-out) bør ikke overstige 20%. Der bør være foretaget en intention-to-treat (ITT) overvejelse (beregning).
 - Det skal være belyst, hvorvidt personalet, der inkluderede patienterne, og patienterne selv godkendte deres deltagelse, før de kunne gennemskue, hvad de blev randomiseret til (= concealment of treatment allocation) - med efterfølgende ITT ansvar.

SØLV evidens

Sølvniveau af evidens gives til systematiske *reviews* eller randomiserede undersøgelser, der ikke i sig selv opfylder de ovennævnte kriterier, eller i prioriteret rækkefølge:

- Publiceret:**
- Resultater fra mindst en ikke-randomiseret kohorte, hvor det med rimelighed kan dokumenteres, hvor stor en andel af personerne, der fik den aktive intervention, og hvor mange personer der ikke gjorde.
 - Mindst et velkontrolleret Case-kontrol studie.

BRONZE evidens

Bronze-niveau af evidens gives til:

- Publiceret:**
- Mindst 1 patientserie uden kontrolgruppe (herunder også det simple forhold, test af før og efter intervention)
 - Ekspertvurdering med baggrund i klinisk erfaring alene underbygget af en fysiologisk hypotese.

ventionsstudier. Biblen om dette emne er "the CONSORT statement" (4). Der er en umiddelbar opfølgning til "The CONSORT statement" med en mere uddybende beskrivelse af emnet af bl.a. statistiker D. G. Altman, der går i dybden med en fortolkning af udsagnene i *The Consort Statement* (5).

Taler man om evidensbaseret medicin, er Cochrane organisationen hovedforangsmænd. I Cochrane sammenhæng er den nok mest udbredte måde at kvantificere kvalitet på at lave en såkaldt Jadad score (6), hvor et kontrolleret effektstudie kan opnå en karakter mellem 0 og 5. Jadad scoren kan fortolkes meget enkelt: *jo højere score – jo lavere chance for bias*.

Det er en elegant finesse ved Jadad-indeksering, at læseren af en videnskabelig artikel alene med udgangspunkt i artiklen skal kunne uddrage, hvorvidt studiet er beskrevet som randomiseret (Ja/Nej); har forfatterne beskrevet deres studie som dobbelt-blindet? (Ja/Nej); samt, er antallet af frafaldne patienter (pr. gruppe) beskrevet i detaljer? (Ja/Nej). Hvert af de to første spørgsmål kan, hvis svaret var ja (rando-

miseret og dobbelt-blindet), yderligere tillægges et point, hvis beskrivelsen virker overbevisende og tilstrækkelig. Hvis det modsatte er tilfældet, kan de to "optjente point" fratrækkes igen. Jadad-scoren kan dog ikke blive lavere end 0. I praksis vil det sige, at jo højere Jadad-score (højest 5) des mere kan man regne med de fundne resultater. Under forholdene omkring frafald pr. gruppe er det i dag en standard, at alle randomiserede undersøgelser skal udføre sine statistiske analyser både på data fra patienter, der faktisk gennemgik interventionen (og mødte op til måling sidste dag), men i særdeleshed også analysere på data fra de patienter, det var intentionen skulle have gennemført. Denne sidste metode kaldes overordnet for *intention-to-treat* (ITT) analyse. I praksis kan disse manglende-data analyser (ITT-analyser) nemt udføres ved at man i sin endelige database (f.eks. i Excel®), manuelt indsætter det først målte tal for patienten ved de blanke celler, der måtte forekomme. Dette kalder man for en *Non-responder-* eller *baseline observation carried-forward* analyse. Sammentælling/beregning af Jadad kvalitetsscore er beskrevet i tabel 2 på denne side.

Tabel 2. Beregning af Jadad kvalitetsscoren (6)

Basis spørgsmål (0-3 point)		Ja	Nej
1.*	Er studiet beskrevet som randomiseret?		
2.	Er studiet beskrevet som dobbelt-blindet?		
3.†	Er der redegjort for drop-outs? • Kan man læse hvor mange personer, der er udgået pr. gruppe, siden randomiseringen • Kan man nøjagtigt læse, hvor mange personer der indgik i analysen • Er der foretaget intention-to-treat (ITT) analyse		
Addition	Tillæg 1 [+1] hvis det pågældende spørgsmål er overbevisende beskrevet [ja]:		
1a*	Var den beskrevne randomisering eksplicit forklaret?		
2a	Virker det sandsynligt at dobbelt-blindingen har virket?		
Subtraktion	Fratræk 1 [-1] hvis udsagnet virker utroværdigt [ja]:		
1b	Virker det som en "tilsnigelse" at kalde studiet randomiseret?		
2b	Er det usandsynligt, at patienter og/eller sundhedspersonale ikke har vidst, hvilken behandling patienterne blev udsat for?		

*: Et ja til dette spørgsmål kan kun retfærdiggøres hvis der også er tale om "concealed allocation"
†: Dette spørgsmål indebærer kun mulighed for at give ét point; det er derfor vigtigt at overveje samtlige af de skitserede underpunkter.

Anbefaling af en ny behandling

At der gennemgående kan være stærk (for eksempel 'platin', jf. tabel 1) evidens for, at en behandling har en positiv effekt på patienter med en defineret lidelse, indebærer ikke nødvendigvis, at patienten skal underkastes denne behandlingsform. Behandleren må gennemskue, om det er prisen og ulejligheden værd at implementere en ny behandlingsstrategi i den pågældende patientgruppe.

Der findes gode simple tommelfingerregler til at bedømme, om en behandling kan anbefales, hvis man antager, at den publicerede randomiserede undersøgelse, man ligger inde med, gennemgående er repræsentativ for ens egen patient. Herunder ses også på, om kvaliteten af undersøgelsen er overbevisende (se ovenfor, tabel 2). Generelt vil det primære slutresultat (*primary outcome*) med rimelighed kunne fortolkes som enten en kontinuer variabel (ratio-interval skala), eller som den simplest tænkelige diskrete datafordeling, en binær variabel (ja eller nej).

Der kan også forekomme studier, hvor de præsenterede data principielt er diskrete med flere udfaldsrum end 2, og som sådan skal håndteres med andet end en *binomialfordeling* (f.eks. *Poissonfordeling* eller *multinomialfordeling*). I forhold til fortolkning af et publiceret randomiseret studie, behøver det ikke at medføre store vanskeligheder. Hvis man af artiklen kan udtrække, hvor mange personer der har opnået det studerede udfald, kan dette fortolkes som binært. Modsat dette: Hvis der er givet medianværdi for hver studeret gruppe (med tilhørende P-værdi til vurdering af, om grupperne er ens), kan man med rette udregne en antagelig spredning (*SD*) for den tilhørende teoretiske *t*-fordeling.

I praksis betyder ovenstående: Når vi fra studier skal sammenfatte, hvorvidt en ny behandling kan anbefales, skal vi vurdere evidensen for en positiv effekt af den pågældende behandling. Vi ser derfor på følgende: 1) er der et systematisk *review*/en metaanalyse af mere end ét RCT, der giver et utvetydigt billede af, om behandlingen virker?; eller 2) Er der et velkontrolleret (randomiseret) studie af en passende kvalitet, eventuelt bedømt ved Jadad-score med mindst 2 ud af 5 mulige (6), som passer utvetydigt på netop den patient, der ønskes behandlet.

Et ofte tilbagevendende emne inden for klinisk forskning i bevægeapparatet er, hvorvidt en behandling/intervention er klinisk signifikant. Med dette menes: Er den statistisk signifikante (publicerede) effekt også relevant i klinikken? Dette emne er blevet gennemgået systematisk i de såkaldte "*Philadelphia Panel evidence-based practice guidelines on selected rehabilitation guidelines*" (7). Disse guidelines belyser kort fortalt, hvilke forhold der bør være opfyldt, før diverse interventioner kan betragtes som brugbare i den brede rehabiliterende klinik. Mere konkret har denne ekspertgruppe fokuseret på lænderyg-, nakke-, knæ- og skuldersmerter. Ved sund kritisk tilgang til stoffet kan disse guidelines generaliseres til øvrige lidelser i bevægeapparatet. Holdning til, om noget er evident, er den samme i Philadelphia-panelet som hos "*The Cochrane Collaboration*", og er således i tråd med det tidligere præsenterede (se tabel 1).

Konklusionen er, at hvis der er evidens (via RCT publicerede signifikante resultater) for, at en intervention har en betydelig effekt på patientrelaterede selvrapporterede slutresultater, såsom for eksempel smerte målt ved visuel analog skala (VAS), så kan behandlingen muligvis anbefales. Philadelphia-panelet definerer en betydelig klinisk effekt som mindst 15 % bedring i forhold til kontrolgruppen og relativt til startniveauet (formlen er præsenteret i faktaboks på side 6). Det er således altafgørende i forhold til panelets rettesnor, at interventionen har været underkastet et veludført RCT, hvor det for patienten væsentlige slutresultat er statistisk signifikant ($P < 0,05$).

Det er en væsentlig fordel for læseren af resultaterne fra et klinisk studie, at den kliniske effekt af de kontinuere data bliver oplyst som procentvise eller relative kliniske effekter. Dette indebærer, at behandleren får en bedre mulighed for at generalisere og fortolke uden alt for mange forbehold.

J Cohen har tidligere beskrevet en tilgang til præsentation af kontinuerte data, for hvilke de faktiske enheder er abstrakte eller direkte uforståelige ved fortolkningen af selv den største statistiske signifikans, kaldet *Effect Size (ES)* (8). Formlen til den simpleste måde at kvantificere denne *Cohen*

Faktaboks

Nyttige estimater, der kan uddrages ud fra de publicerede resultater fra et veludført randomiseret, kontrolleret studie (RCT).

Variationen (støj) inden for hver gruppe er oplyst som

SE (= stikprøve spredning) eller

SD (= standard afvigelse).

Disse kan med rimelig sikkerhed omregnes ud fra følgende formel:

$$SD_A = SE_A \times \sqrt{N_A}$$

$$SE_A = SD_A / \sqrt{N_A}$$

- og tilsvarende for SD_B og SE_B

a:

Relativ procentvis forskel (RPF)

$$RPF = (GNS_A - GNS_B) / SN \times 100 \%$$

b:

Effect Size (ES)

$$ES = (GNS_A - GNS_B) / SD_{A\&B}$$

c:

Number Needed to Treat (NNT)

$$NNT = 1 / (\text{Succes}_A - \text{Succes}_B)$$

Forkortelser i øvrigt:

N_A = Antal personer i den A'te gruppe!

GNS_A = Gennemsnittet i den A'te gruppe

SN = Startniveauet, gennemsnitsværdien

ved baseline på tværs af alle grupper

$SD_{A\&B}$ = Dette er den gennemsnitlige standard deviation

Et "tænkt" eksempel - brug formlerne

Vi forestiller os et 2-armet RCT hvor der er 105 patienter, der er blevet randomiseret til 2 behandlinger (A hhv. B). Startniveauet (SN) for variabelen af interesse: SN=900 score-point.

Ud af artiklen kan vi læse antallet af patienter, der er randomiseret til hver af de to behandlinger ($N_A=52$ patienter hhv. $N_B=53$ patienter). Disse vil resten af vejen være intention-to-treat (ITT) patienter.

Behandlingseffekten pr. gruppe er givet som gennemsnit (alternativt median-værdier), hvilket vi noterer som GNS_A hhv. GNS_B : $GNS_A=334,5$ og $GNS_B=115,2$.

I det følgende sætter vi $SD_A = SD_B = SD_{A\&B} = 370$.

Tilsvarende kan man af "artiklen" udtrække, hvor mange patienter der beskrev sig selv som raske efter interventionsperioden: $Rask_A(25)$ hhv. $Rask_B(13)$.

Dette medfører naturligt, at vi beregner succes-raten: $\text{Succes}_A = Rask_A / N_A (25/52=0,48)$ hhv. $\text{Succes}_B = Rask_B / N_B (13/53=0,25)$.

Vi kan beregne den absolutte risiko reduktion (ARR); $ARR = \text{Succes}_A - \text{Succes}_B = 0,23$.

$$RPF = (334,5 - 115,2) / 900 \times 100 \% = 24,4 \%$$

$$ES = (334,5 - 115,2) / 370 = 0,59$$

$$NNT = 1 / 0,23 = 4,3 \text{ patienter}$$

ES er vist i faktaboks på denne side. Finessen ved denne tilgang til stoffet er, at man med rimelighed kan fortolke den kliniske effekt via ES til:

- $ES \geq 0,2$: lille klinisk effekt
- $ES \geq 0,5$: moderat klinisk effekt
- $ES \geq 0,8$: stor klinisk effekt

I praksis benyttes dog oftest den såkaldte *Hedges'* justerede g-værdi, der er det beregningsgrundlag, "The Cochrane Collaboration" benytter via den såkaldte *Standardised Mean Difference*, SMD (9). Forskellene ved disse to metoder er ikke af den store betydning, men

den justerede værdi, g-værdien, har i højere grad taget højde for mulige usikkerheder ved små studier. Vi har foretaget en simulationsanalyse af denne mulige diskrepans, som følge af *Hedges'* justering eller ej: Ved studier med færre end 10 patienter pr. gruppe vil justeringen kunne påvirke den kliniske fortolkning (>4,2 % diskrepans); med mere end 25 patienter pr. gruppe vil justeringen være moderat (<1.6 %) og endelig vil justeringen være minimal, hvis der er mere end 50 patienter pr. gruppe, hvilket er kravet for et studie med guldevidens (<0.8 %).

Når et statistisk signifikant RCTs primære slutresultat er et *binært endepunkt*, dvs. antal opnåede "raske" ved behandlingen i forhold til det antal, der kunne være blevet det (ja/nej), så vil det fortolkningsmæssigt være mest fornuftigt at beregne det såkaldte "*Number Needed to Treat*" (*NNT*).

NNT er blevet et populært effektmål, fordi det er relativt nemt at forstå og intuitivt meningsfuldt. Ydermere foreligger der en klar anbefaling om, at reumatologisk forskning præsenteres med *NNT* (3). Formlen til den simpleste måde at kvantitere *NNT* på er vist i faktaboks side 6. Det skal retfærdigvis nævnes, at effekt-begrebet *NNT* bedst fortolkes/ implementeres, når der er tale om klinisk "hårde end-points", såsom: infarkt, apopleksi, ortopædkirurgi eller død. Alligevel kan man for selv-rapporterede data (som f. eks. VAS-smerte) sagtens definere et binært succeskriterium og ud fra dette beregne, hvor mange patienter der har opnået dette efter endt intervention. Man kan umiddelbart fortolke *NNT* således: Hvis vi forestiller os, at en intervention medfører en *NNT* på 5 (beregnet mod den randomiserede kontrolgruppe, se faktaboks), så vil dette kunne fortolkes som 5 patienter skal behandles med den pågældende intervention for at behandle (eng. *Treat*) 1 patient med positiv effekt.

Det har vist sig, at der er ganske god overensstemmelse mellem *ES* og *NNT*, og at dette praktisk talt er uafhængigt af, hvad man definerer som succeskriterium ved beregning af *NNT* (10). Vi er i skrivende stund ikke bekendt med nogen egentlig klinisk konsensus omkring, hvordan diverse niveauer af *NNT* kan fortolkes. Derfor vil de følgende anbefalinger til fysioterapeutisk praksis alene bygge på en vurdering af den empiriske evidens baseret på percentiler i en normalfordeling, der for nylig blev fremlagt på den europæiske reumatologi kongres, *EULAR* (11) - med baggrund i de oprindeligt præsenterede *ES* (8):

- *NNT* < 13 patienter: lille klinisk effekt
- *NNT* < 6 patienter: moderat klinisk effekt
- *NNT* < 4 patienter: stor klinisk effekt

En praktisk måde at anskue *NNT* på er at huske, at en *NNT*-værdi på 2 ville svare til, at hver anden patient ville få en mærkbar forbedring af sin lidelse efter at være udsat for den pågældende intervention, uafhængig af den øvrige behandling. Succeskriteriet for *NNT*'s beregning kræver såle-

des ikke, at patienten bliver "rask", men at der *a priori* er et veldefineret behandlingsmål. På dette grundlag skal man naturligvis antage en mere pragmatisk holdning til sådanne anbefalinger, end hvis man betragter et selv-forklarende "*hårdt endepunkt*" som *død/ikke-død* patient i kardiologisk- eller cancer-mæssig sammenhæng.

Valg af behandling – for den travle kliniker

Enhver intervention, en patient udsættes for, kan ved første bedømmelse give indtryk af at have en positiv effekt. Det kan ganske simpelt være en effekt af, at der er gjort noget, der giver håb om bedring. Der er i dette tilfælde tale om en placeboeffekt i den bredest tænkelige forstand.

Evidensbaseret medicin har ikke til hensigt at forkaste gode menneskelige tiltag, men objektivt at kvantificere, hvorvidt en ny behandling/intervention kan oppebære en ekstra omkostning for patienten eller samfundet. Hvis en intervention er blevet undersøgt i én for klinikerens relevante patientgruppe, så skal klinikerens være bevidst om, hvorvidt resultatet er blevet til i et studie af rimelig kvalitet.

Drejer det sig om implementering af nye behandlinger, bør de alle have deres effekt belyst i et randomiseret kontrolleret studie af en rimelig kvalitet. Et minimumskrav er en *Jadad-score* på mindst 2. Når man i sin klinik har med patienter at gøre, bør behandleren kunne stilles til ansvar for, hvorfor denne har anbefalet en specifik behandling. Hvis der for et kontinuert slutresultat er fundet en statistisk signifikant effekt, bør den relative kliniske forskel være på mindst 15%. Det generelle billede belyses til gengæld bedre, hvis man betragter *Effect Size (ES)*, der skal være mindst 0,2, før man med rimelighed kan regne med, at effekten kan monitoreres og dermed kan indføres i klinikken med rimelig begrundelse.

Hvis en behandling for eksempel medfører en *NNT* på 13 eller derover, må interventionen anses for irrelevant for patienten. Klinikerens må derudover gøre sig sine egne erfaringer med den pågældende behandling til den relevante patientgruppe, hvis der foreligger acceptable publicerede studier, der som minimum har vist moderat effekt ($ES \geq 0,5$ og/eller $NNT < 6$).

Litteratur

1. Pedersen T, Gluud CN, Gotzsche PC, Matzen P, Wille-Jorgensen PA. Hvad er evidensbaseret medicin? *Ugeskr. Læger* 2001;163:3769-72.
2. Eccles M, Freemantle N, Mason J. North of England evidence based guidelines development project: methods of developing guidelines for efficient drug use in primary care. *BMJ* 1998;316:1232-5.
3. Tugwell P, Shea B, Boers M et al. *Evidence-Based Rheumatology*. London: BMJ books, 2004.
4. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
5. Altman DG, Schulz KF, Moher D et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern.Med.* 2001;134:663-94.
6. Jadad AR, Moore RA, Carroll D et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin.Trials* 1996; 17:1-12.
7. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions: overview and methodology. *Phys.Ther.* 2001; 81:1629-40.
8. Cohen J. *Statistical power analysis for the behavioural sciences*. Lawrence Earlbaum Associates, 1988.
9. The Cochrane Collaboration. *RevMan Analyses [Computer program]*. Review Manager (RevMan) 4.2. Oxford, England: 2003.
10. Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med.Care* 2001;39:1039-47.
11. Christensen R, Dougados M, Holm CC, Bliddal H. The number needed to treat (NNT) in relation to effect size (ES): preliminary standards for clinical efficacy in evidence-based rheumatology. *Ann. Rheum. Dis.* 2006;65II:277-8 (Abstract).

Abstract

Evidensbaseret behandling af bevægeapparatet - virker den pågældende behandling og hvornår kan den anbefales?

*Robin Christensen, cand. scient., ph.d. stud. og
Else Marie Bartels dr. scient., seniorforsker.*

Hvis et nyt behandlingstiltag er blevet undersøgt og fundet gunstigt i en publiceret randomiseret og kontrolleret undersøgelse, skal man desuden være bevidst om, hvad man kan forvente af klinisk effekt. Hvis der for et kontinuert endepunkt er fundet en statistisk signifikant effekt, bør den relative procentvise forskel være på mindst 15 %. For en kliniker er det værd at gøre sig sine egne

erfaringer med en behandling, hvis der er kvalificerede, publicerede værker, der som minimum har vist en moderat effekt, dvs. en Effekt Size (ES) $\geq 0,5$ og/eller Number Needed to Treat (NNT) < 6 . Den kliniske "bagatelgrænse" for diverse fysioterapeutiske behandlinger bliver her til praktiske formål sat til at være: ES $\geq 0,2$ og/eller NNT < 13 patienter.