Hur utvärdera behandlingsresultat vid artros?

-med fokus på patient-relevanta instrument

EWA ROOS

Introduktion och sammanfattning

Riktlinjerna för kliniska studier av artros innefattar utvärdering av såväl patient-relevanta aspekter som utvärdering av ledstruktur med röntgen, MR eller motsvarande [1]. Ledstruktur utvärderas främst i studier med mer än ett års uppföljning. I denna sammanställning ligger fokus på patient-relevanta utvärderingsinstrument som kan användas såväl inom forskning som i klinisk praxis.

I kliniska artrosstudier utvärderas primärt patient-relevanta aspekter som symptom och funktion. Det vanligaste symptomet vid etablerad artros är smärta. Vid tidiga stadier är dock funktionsnedsättning ofta ett väl så påtagligt symptom som smärta. Vanligen mäts effekten i behandlingsstudier av artros med hjälp av väldokumenterade frågeformulär som utvärderar såväl patientens upplevda smärta, stelhet och övriga symptom som funktionsnedsättning. Ibland utvärderas också vilken effekt smärta och funktionsnedsättning har på patientens livskvalitet.

Historisk bakgrund

Utvärdering av behandling, eller bedömning av status, kan ske på olika nivåer. Traditionellt har ledsjukdom utvärderats med objektiva mått på organskada (röntgen, stabilitetsmätning) eller klinisk undersökning då svullnad, rodnad, palpationsömhet etc. har utvärderats. Sällan användes patientens åsikter för att bedömma behandlingsresultatet. Även om patientens åsikter ansågs viktiga bedömdes de också som subjektiva och opålitliga. Då symptom och funktion utvärderades skedde det istället med så kallade ledscorer. För knäartros användes t ex the Hospital for Special Surgery Score (HSS) och för höftartros Harris Hipscore. Gemensamt för dessa scorer är att de är utvecklade för att stödja läkaren i hans/hennes uppfattning om patientens symptom och funktion. Sålunda är frågorna konstruerade av läkare, och också poängsatta och viktade olika beroende på konstruktörens uppfattning om hur mycket olika symptom och funktionsnedsättningar bidrar till en försämrad led. En totalscore beräknas för alla frågor, dvs. symptom och funktion adderas till varandra, tillsammans med till exempel rörlighet och mekaniska problem. Inte förvånande har det visat sig att resultaten skiljer sig åt då samma patientgrupp utvärderas med de olika instrumenten. Ett skäl till detta är att sambandet mellan strukturella förändringar och patientens symptom är svagt, inte minst i tidiga skeden av sjukdomen. Ingen av dessa scorer testades ursprungligen med avseende på pålitlighet (reliabilitet), giltighet (validitet) eller känslighet för kliniskt relevant förändring.

I takt med att patienten satts i centrum för behandlingsprocessen och ekonomiska åtstramningar genomförts, har behovet av att undersöka om våra behandlingar tillfredställer patienternas behov ökat. Under 90-talet har antalet instrument (frågeformulär) som mäter patient-relevanta aspekter ökat dramatiskt och finns nu tillgängliga för de flesta vanliga patientgrupper.

Metodologiska egenskaper hos utvärderingsinstrument

Validitet, reliabilitet och känslighet för förändring över tid är de grundläggande egenskaperna hos alla typer av utvärderingsinstrument. Ju bättre egenskaper ett instrument har desto mer användbart är det.

Gemensamt för den nya generationen av utvärderingsinstrument är att de är utvecklats i väl strukturerade processer inkluderande intervjuer med i första hand patienter men också sjukvårdspersonal. Detta förfarande borgar för god validitet, det vill säga att kvaliteter viktiga för avsedd grupp verkligen undersöks. Till exempel vet vi att patienter anser att frågor om förändringar i livsstil och hur ofta man tänker på sitt knä är väl så väsentliga som frågor kring smärtans utbredning och duration. Den nya generationen instrument är valida mått på patientens syn, aspekter som inte korrelerar med vare sig röntgenbilder, stabilitetsmätning eller klinisk undersökning. Observera att det ibland inte räcker att instrumentet är validerat för diagnosen i fråga (artros) utan också måste vara testat på aktuell patientgrupp då symptom och krav på fysisk funktion skiljer sig åt mellan olika grupper av patienter med artros.

God reliabilitet borgar för att resultatet är reproducerbart och därmed pålitligt. Patient-relevanta frågeformulär har visat sig avsevärt mer pålitliga (test-retest stabila) än både upprepade läsningar av röntgenbilder och klinisk undersökning.

God känslighet över tid borgar för att en verklig förändring i status verkligen är mätbar. Det sistnämnda innebär att färre patienter behövs i kliniska studier för att påvisa behandlingsresultat om patient-relevanta utvärderingsinstrument används istället för traditionella mått på ledsjukdom. Effect size (förändring i poäng delat med standard deviationen av ursprungspoängen) är ett statistiskt mått på effekten av en behandling. Om samma utvärderingsinstrument används kan effekten av olika behandlingar jämföras (Figur 1). Denna typ av jämförelser kan bland annat användas för att styra vårdresurser.

Vilket instrument ska jag välja?

Det råder idag konsensus om att det primära utvärderingsinstrumentet vid utvärdering av artrosbehandling bör vara ett frågeformulär som utvärderar för patienten adekvata symptom och funktionsnedsättningar. I kliniska studier rekommenderas komplettering med ett generellt instrument för att få bredare täckning, upptäcka oväntade effekter och kunna jämföra med andra sjukdomar [1]. Vilket instrument man bör välja styrs av vilken eller vilka leder som ska studeras och avsikten undersökningen. För knän, höfter och händer finns ledspecifika instrument framtagna. För generaliserad artros kan det vara lämpligare med ett generellt utvärderingsinstrument som tar hänsyn till hela kroppen. Samma rekommendationer kan gälla för kvalitetssäkring i vården av enskilda diagnoser. Om en specifik enhets verksamhet kring flera diagnoser ska utvärderas kan ett generellt (t ex SF-36) och eventuellt ett patient-specifikt (t ex Patient-specific Functional Scale, PSFS [2]) instrument användas. Figur 2 utgör en översikt över publicerade patient-relvanta utvärderingsinstrument som kan användas vid artros medan några av de mer vanligt förekommande utvärderingsinstrumenten beskrivs nedan.

Vanligt förekommande patient-relevanta utvärderingsinstrument

Generellt patient-relevant instrument: SF-36

Instrumenten kan vara generella, det vill säga att de kan användas för alla oavsett diagnos. Det mest använda generella instrumentet idag är SF-36 som mäter hälsostatus i åtta olika delskalor och täcker såväl fysisika som mentala aspekter [3]. SF-36 finns översatt och anpassat till svenska förhållanden [4]. SF-36 har använts i

1000-tals studier på friska och olika patientgruppen och är ett lämpligt instrument om man vill göra jämförelser med någon annan grupp. Till exempel finns referensvärden från 8 930 personer som representerar Sveriges befolkning att tillgå. Nackdelen med de generella instrumenten är att de är just generella och därmed kanske inte tillräckligt känsliga för sjukdoms-specifika symptom eller funktionsnedsättningar.

Knä- och höftspecifika patient-relevanta instrument: WOMAC, KOOS, HOOS

Flest instrument finns det i den sjukdoms-specifika gruppen. För knä- och höftartros används ofta WOMAC (Western Ontario and MacMaster Universities Osteoarthritis Index) som mäter smärta, stelhet och funktion i dagliga livet [5]. För yngre och fysiskt aktiva med knäartros är motsvarade instrument KOOS (Knee Injury and Osteoarthritis Outcome Score) som innehåller alla WOMAC-frågor men som i tillägg utvärderar sport- och fritidsfunktion och knä-relaterad livskvalitet [6, 7]. En adaptation av KOOS för höftartros, Hip dysfunction and osteoarthritis outcome score (HOOS), har nyligen publicerats [8].

Samtliga instrument är översatta och anpassade för svenska förhållanden [8-10]. KOOS-formulär på olika språk, manual och program för poängberäkning kan hämtas ner från Internet (www.koos.nu). De flesta instrumenten, inklusive de ovan nämnda, kräver poäng-beräkning, för hand eller vanligast med dator. Dessa instrument lämpar sig väl för användning i kliniska studier.

Hand-specifika patient-relevanta instrument:AUSCAN

Motsvarigheten till WOMAC för handartros heter AUSCAN (Australian/Canadian) osteoarthritis index. Instrumentet består av 10 smärtfrågor, 2 frågor om stelhet och 30 frågor om handfunktion. Såväl själva instrumentet som en första test av dess grundläggande egenskaper publicerades 2002 [11, 12]. Instrumentet är nytt och data saknas ännu från andra användare.

Patient-specifikt instrument lämpligt för kliniskt bruk: PSFS

Om man på en mottagning önskar utvärdera behandlingseffekten hos samtliga patienter med muskuloskeletala besvär är det opraktiskt att ha ett instrument för varje diagnos, och då kan patient-specifika instrument användas. Ett exempel är Patient-Specific Functional Scale [2] som låter patienten själv ange de mest besvärande funktionstörningarna och sen med hjälp av en visuell analog skala bedöma förändring efter behandling. Denna typ av instrument är dock utanför fokus för denna sammanställning.

Sekundära utvärderingsparametrar

Aktivitetsnivå

Bedömning av aktivitesnivå ska göras för att spåra förändringar efter skada eller behandling hos en individ. Jämförelser mellan patienter bör inte göras eftersom en lägre aktivitesnivå inte är liktydigt med ett sämre knä.

De flesta aktivitetsskalor bygger på gradering av olika idrotter eller yrken med avseende på tyngd eller knäprovocerande moment. För artrospatienter kan aktivitetsnivån bedömmas enligt Edworthy. I dessa skalor bedöms fritid, arbete och gångförmåga separat på en skala från 0-6. Dessa skalor sträcker sig från inget hushållsarbete, sjukskriven/förtidspensionär och behöver käpp till tävlingsidrott, tungt arbete och går obehindrat i kuperad skogsterräng. Skalorna har god test-retest pålitlighet och god känslighet för förändring vid skada respektive intervention [13].

Funktionella test

Det är väl känt att att styrke- och konditionsträning med måttlig intensitet minskar symptomen och förbättrar funktionen vid knäartros. Test av muskelstyrka och kondition kan därför användas dels för att motivera patienten till träning och dels för att objektivt dokumentera resultatet av träning. Enkla funktionella test för muskelstyrka i quadriceps har visat sig praktiskt möjliga att utföra av allmänläkare. Ett exempel på ett praktiskt genomförbart test är det så kallade "Klivtestet" då patienten ombedes kliva upp med ett ben i taget på en så hög nivå som möjligt. Klivlådan har en löstagbar skiva som kan flyttas mellan sju olika nivåer [14]. Nivå II och nivå IV motsvarar höjden av ett respektive två trappsteg. I stället för klivlåda kan pallar och stolar av olika höjd användas. Tillse dock att pallen/stolen är fastgjord, tänk på fallrisken! Eftersom klivhöjden är beroende inte bara av patientens längd och vikt utan också av höft-, knä-, och fotledsrörlighet kan det vara svårt inte bara att jämföra med andra patienter utan också att jämföra mellan sidorna. Därför rekommenderas att varje patient är sin egen kontroll. Ett mått på god klivförmåga är om patienten kan kliva upp på den höjd där det initialt är cirka 90 graders vinkel i knäleden. Detta klaras upp till 80-85 års ålder.

Referenser

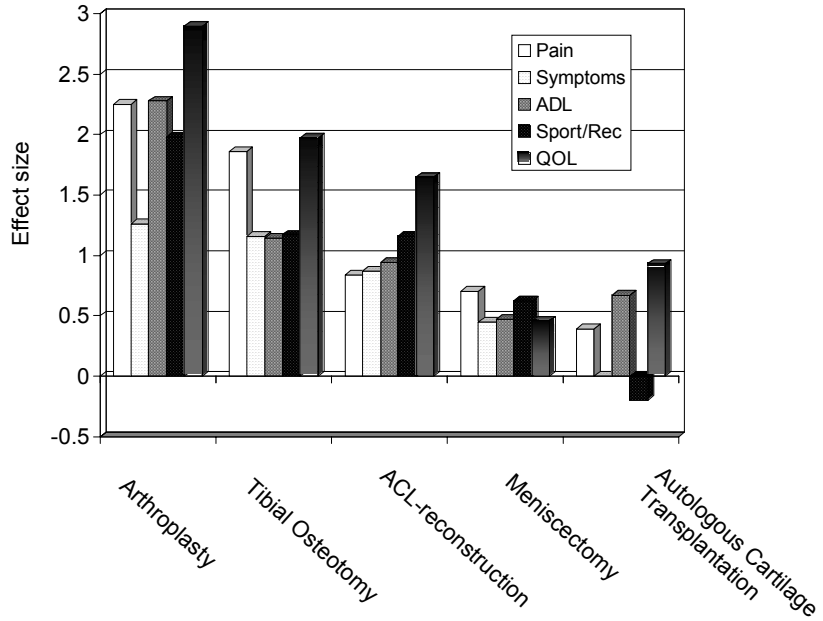
1. N Bellamy, J Kirwan, M Boers, P Brooks, V Strand, P Tugwell, R Altman, K Brandt, M Dougados, M Lequesne: Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997, 24:799-802.
2. MD Westaway, PW Stratford, JM Binkley: The patient-specific functional scale: validation of its use in persons with neck dysfunction. *J Orthop Sports Phys Ther* 1998, 27:331-8.
3. JE Ware, Jr., CD Sherbourne: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992, 30:473-83.
4. M Sullivan, J Karlsson: SF-36 Hälsoenkät: Swedish Manual and Interpretation Guide. Gothenburg, Sweden: Health Care Unit, Sahlgrenska Hospital; 1994.
5. N Bellamy, WW Buchanan, CH Goldsmith, J Campbell, LW Stitt: Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988, 15:1833-40.
6. EM Roos, HP Roos, LS Lohmander, C Ekdahl, BD Beynnon: Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998, 28:88-96.
7. EM Roos, LS Lohmander: Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes* 2003, 1:64.
8. AK Nilsdotter, SL Lohmander, M Klassbo, EM Roos: Hip Disability and Osteoarthritis Outcome Score (HOOS) - Validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003, 4:10.
9. EM Roos, M Klassbo, LS Lohmander: WOMAC osteoarthritis index. Reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and MacMaster Universities. *Scand J Rheumatol* 1999, 28:210-5.
10. EM Roos, HP Roos, C Ekdahl, LS Lohmander: Knee injury and Osteoarthritis Outcome Score (KOOS)--validation of a Swedish version. *Scand J Med Sci Sports* 1998, 8:439-48.
11. N Bellamy, J Campbell, B Haraoui, E Gerecz-Simon, R Buchbinder, K Hobby, JC MacDermid: Clinimetric properties of the AUSCAN Osteoarthritis Hand

Index: an evaluation of reliability, validity and responsiveness. *Osteoarthritis Cartilage* 2002, 10:863-9.

12. N Bellamy, J Campbell, B Haraoui, R Buchbinder, K Hobby, JH Roth, JC MacDermid: Dimensionality and clinical importance of pain and disability in hand osteoarthritis: Development of the Australian/Canadian (AUSCAN) Osteoarthritis Hand Index. *Osteoarthritis Cartilage* 2002, 10:855-62.
13. EM Roos, HP Roos, LS Lohmander: WOMAC Osteoarthritis Index--additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and MacMaster Universities. *Osteoarthritis Cartilage* 1999, 7:216-21.
14. E Roos: Hur utvärdera behandlingsresultat vid knäsjukdom? In: *Knäledens sjukdomar och skador* Edited by J Karlsson. pp. 120-130. Södertälje: Astra; 2000: 120-130.
15. NA Johanson, ME Charlson, TP Szatrowski, CS Ranawat: A self-administered hip-rating questionnaire for the assessment of outcome after total hip replacement. *J Bone Joint Surg Am* 1992, 74:587-97.
16. JN Katz, CB Phillips, R Poss, JJ Harrast, AH Fossel, MH Liang, CB Sledge: The validity and reliability of a Total Hip Arthroplasty Outcome Evaluation Questionnaire. *J Bone Joint Surg Am* 1995, 77:1528-34.
17. MG Lequesne, C Mery, M Samson, P Gerard: Indexes of severity for osteoarthritis of the hip and knee. Validation--value in comparison with other assessment tests. *Scand J Rheumatol Suppl* 1987, 65:85-9.
18. EM Roos, S Toksvig-Larsen: Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes* 2003, 1:17.
19. J Dawson, R Fitzpatrick, A Carr, D Murray: Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br* 1996, 78:185-90.
20. J Dawson, R Fitzpatrick, D Murray, A Carr: Questionnaire on the perceptions of patients about total knee replacement.*J Bone Joint Surg Br* 1998,80:63-9.
21. M Bergner, RA Bobbitt, WE Pollard, DP Martin, BS Gilson: The sickness impact profile: validation of a health status measure. *Med Care* 1976, 14:57-67.
22. SM Hunt, SP McKenna, J Williams: Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthrosis. *J Epidemiol Community Health* 1981, 35:297-300.

Figur 1 Effect size vid olika ortopediska artrosinterventioner

KOOS effect sizes 6 månader efter olika knäoperationer. Effect size definieras som förändring i poäng delat med standard deviationen av den pre-operativa poängen. Resultaten från knäledsartroplastik (n=105, medelålder 71.3 år) och främre korsbandsrekonstruktion (n=21, medelålder 32 år) är publicerade. Resultaten för tibia osteotomi baseras på 50 patienter (medelålder 54 år) som genomgått hemicallotasis osteotomi (W-Dahl A, Lund University, personal communication 2001). Resultaten för brosktransplantation baseras på 50 patienter (medelålder 34 år) som genomgått autolog brosktransplantation (Olsson M, University of Gothenburg, personal communication 2002). Resultaten för meniskektomi baseras på 66 patients (medelålder 44 år) som genomgått artroskopisk partiell meniskektomi (Roos E, Lund University, personal communication 2003). Resultatet av placebobehandling har inte tagits hänsyn till vid beräkningen av effect sizes. Det bör noteras att effect size är beroende av den pre-operativa standard deviationen som kan variera mellan grupperna. Likaså är en högre effect size möjlig i grupper som har ett sämre utgångsläge. Alla resultat visas med tillstånd från författarna.



Figur 2 Översikt över patientrelevanta utvärderingsinstrument

Instruments	# of questions	#of domains	Time to complete	Measurement level			Administration method		Assesses measurement levels in separate scores
				Impair-ment	Disa-bility	Handi-cap	Patient Admini-stered	Observer admini-stered	
Disease and joint specific									
Hip Rating Questionnaire [15]	15	1	10	•	•		•		
Total Hip Arthroplasty Outcome Evaluation Questionnaire [16]	15			•	•	•			
Lequesne Indices [17]	10 knee 10 hip	1	10	•	•			•	
WOMAC [5]	24	3	5	•	•		•		•
KOOS [6, 10, 13, 18]	42	5	10	•	•	•	•		•
Oxford [19, 20]	12	1	5	•	•		•		
AUSCAN [11, 12]	42	3	10	•	•		•		•
Generic									
SIP [21]	138	7	30	•	•	•	•	•	•
NHP [22]	38	6	10	•	•	•	•		•
SF-36 [3]	36	8	10	•	•	•	•		•

From: Bischoff HA, Roos EM, Liang MH. Outcome Assessment In Osteoarthritis: a guide for research and clinical practice. In: Osteoarthritis, 2nd edition. Editors: Brandt, Lohmander and Doherty, Oxford University Press, 2003.