

Fallgropar vid mätning av smärta

Visuella analogskalan som smärtskattningsmetod ifrågasatt



HÅKAN ROLFSSON, specialist i anestesi och intensivvård, biträdande överläkare, anestesioch smärtkliniken, Capio S:t Görans sjukhus, Stockholm
hakan.rolfsson@capio.se

Går det att mäta smärta? Vilka problem ställs man inför när man ska mäta ett subjektivt fenomen som smärta? Är VAS, den vanligast använda metoden, den bästa?

Det vanligaste instrumentet inom smärtforskning är VAS, visuell analogskala. VAS introducerades i slutet av 1960-talet och används för att mäta ett stort antal fenomen förutom smärta. Metoden har fått stort genomslag och är nu till och med förankrad hos tonåringar i Sverige – »Hur sur är du på en skala från 0 till 10?«. Metoden är lätt att förstå och genererar siffror som man kan använda för statistiska analyser. Samtidigt har det hela tiden funnits kritiker mot metoden som menat att det är ett alltför enkelt och mekaniskt sätt att mäta subjektiva fenomen. Kritiker har också menat att de statistiska metoder som används inte är tillämpliga på denna typ av data och att en stor del av smärtforskningen bygger på osäkra statistiska analyser.

Vetenskapsteoretisk bakgrund

Inom medicinen möter vi patienter som lider svårt och där vi inte riktigt kan hitta biomedicinska förklaringar till deras tillstånd. Exempel är colon irritabile, psykogen epilepsi och svår lumbago. Inom smärtmedicin finns en stor grupp patienter med kronisk eller långdragen smärta där vi saknar goda biomedicinska förklaringar till patienternas lidande. Ett alternativ för att förstå dessa patienter är den biopsykosociala modellen.

Den biopsykosociala modellen lanserades av den amerikanske psykiatern George L Engel i Science 1977 [1]. En stor del av hans artikel är en kritik av vår rådande biomedicinska modell. Denna är en del av hela vårt naturvetenskapliga tänkesätt som växte fram på 1600-talet med den franske filosofen René Descartes som en av föregångsmännen. Detta tänkesätt är reduktionistiskt, mekanistiskt och bygger på kropp–själdualismen. Med reduktionism menas att ett komplext system inte är något annat än summan av de ingående delarna. Descartes menade att världen kan förstås som ett urverk som man kan plocka sönder, och genom att analysera kugghjulen förstod man hela urverket. Analogt kan sjukdomar förstås på cellulär och subcellulär nivå och slutligen som kemiska och fysiska processer. Descartes menade att det inte finns någon avgörande skillnad mellan den mänskliga kroppen och en maskin. För att få ihop denna modell var Descartes tvungen att placera själen utanför kroppen, det som brukar kallas den cartesianska dualismen. Som Engel påpekar i sin artikel är vår naturvetenskapliga världsbild inte bara en vetenskaplig metod. Den är också det sätt som människor i västvärlden förstår sin verklighet på. Våra smärtpatienter vill att vi ska hitta det trasiga kugghjulet i deras urverk och laga det.

En stor del av problemet när det gäller kronisk värk är samhällets, läkarnas och patienternas låsning vid den biomedicinska modellen. Den biopsykosociala modellen är icke-reduktionistisk och icke-mekanistisk och menar att själen finns i krop-

pen. Smärta och sjukdom måste förstås i en helhet där både kroppen (bio-), själen (psyko-) och samhället (sociala) ingår. Det hela samverkar dialektiskt, och orsakssambanden kan vara riktade åt olika håll. Motsättningen mellan en utpräglad biomedicinsk och en mer biopsykosocial syn kommer fram i litteraturen om VAS och smärtskattning.

Smärta

International Association for the Study of Pain (IASP) definierar smärta som en subjektiv upplevelse, »en obehaglig sensorisk och/eller känslomässig upplevelse förenad med vävnadsskada eller hotande vävnadsskada, eller beskriven i termer av sådan«. Denna definition, som betonar smärtans subjektivitet, medför att smärta i litteraturen om smärtskattning ses som ett konstrukt [2, p. 15]. Ett konstrukt definieras som en etikett man använder för att kategorisera en närliggande grupp observationer. Smärta kan jämföras med andra konstrukt, som depression, ångest eller intelligens. Smärta ska inte blandas ihop med nociception, dvs aktivering av receptorer och nerver som förmedlar information om vävnadsskada.

För att mäta ett konstrukt som smärta väljer man vanligtvis att operationalisera (omforma begrepp till mätbara variabler) den genom att kategorisera nivåer av smärta på en skala (precis det man gör när man skapar en skattningsskala som VAS).

Visuell analogskala

VAS är en 10 cm lång horisontell linje. Vid den vänstra ändpunkten står det »ingen smärta« och vid den högra »värsta tänkbara smärta«. Patienten sätter en markering på linjen som motsvarar den aktuella smärtan. Man mäter sedan var markeringen hamnat. Vanligtvis mäter man i millimeter, vilket ger 101 skalsteg. Hur lång linjen är, om den är horisontell eller vertikal, var man sätter texten och hur man formulerar den, allt detta påverkar resultatet. För att man ska få jämförbara resultat är det viktigt att man är noggrann när man använder VAS. Det finns andra metoder för smärtskattning, exempelvis NRS (Numerical Rating Scale), där patienten uppmanas lämna en siffra, vanligtvis mellan 0 och 10, som motsvarar smärtintensiteten, eller VRS (Verbal Rating Scale). Patienten får där välja mellan ett antal ord som beskriver smärtan, t ex »ingen«, »lindrig«, »måttlig«, »allvarlig«.

Den moderna typen av VAS började användas inom psykiatrin för att mäta graden av depression i slutet av 1960-talet [3]. Rötterna finns dock i 1920-talet. I USA tog man fram en metod för att bedöma underordnades egenskaper [4, 5]. Forskare konstruerade en Graphic Rating Method, som var en VAS med

SAMMANFATTAT

Visuell analogskala, VAS, är den vanligaste metoden för smärtskattning inom både kliniskt arbete och forskning. **Den vanliga uppfattningen** – att VAS är en metod med hög validitet och reliabilitet vars data kan analyseras med

parametriska statistiska metoder – ifrågasätts. **Framför allt kritiserar** användningen av just parametriska statistiska metoder som bl a ligger bakom aktuella evidensbaserade riktlinjer för smärtbehandling.

beskrivande ord under skalan. Man fick på så sätt fram kvantitativa data med hög upplösning. Nu används VAS för att mäta ett otal konstrukt förutom smärta.

Reliabilitet

Inom statistik är reliabilitet ett mått på en metods tillförlitlighet, dvs frånvaro av mätfel. Det ska skiljas från validitet (se nedan). I litteraturen brukar man säga att VAS har god reliabilitet. Ett viktigt sätt att kontrollera en skalas reliabilitet är genom att undersöka test-retest-reliabilitet. Upprepade mätningar under samma förutsättningar ska ge samma resultat. Vid en litteraturgenomgång av VAS vid cancersmärta från 2003 fann man en genomsnittlig test-retest-koefficient på 0,8, vilket är bra [6]. En studie från 2005 [7] visade dock endast 11–26 procents överensstämmelse vid upprepad mätning (i den studien var test-retest-överensstämmelsen för VRS 87–100 procent). Ett problem med mätningar av konstrukt är att man inte har något »sant« värde att bedöma sina mätningar mot. Olika VAS-värden vid olika mätningar kan ju bero på att smärtan ändrats.

Validitet

Validitet är ett mått på att man mäter det man vill mäta. Även nu får man problem när man mäter ett konstrukt som saknar ett objektiva »sant« värde. Inom psykometri finns därför speciella validitetsbegrepp. När det gäller en endimensionell metod som VAS brukar man mäta kriterievaliditet. Man jämför då olika skalor för smärtskattning med varandra, dvs använder dem som kriterier för varandra. Många studier har visat god kriterievaliditet för VAS [2, p. 20]. Men det finns också här nyare studier som visat på motsatsen [7, 8]. Författarna till de senare studierna menar också att man måste använda en icke-parametrisk metod som tar i beaktande VAS-mätningarnas ordnade egenskaper när man bedömer kriterievaliditet (se nedan).

Några kritiker för fram att VAS är ett endimensionellt sätt att mäta smärta på [9]. Smärtmätning blir en signalmätningssuppgift för patienten. Dessa kritiker menar att smärtskattning »snarare ska förstås som ett försök att skapa mening påverkat av och refererande till en mängd inre och yttre faktorer och privata betydelser än uppgiften att koppla ett isolerat inre stimuli till en sträcka eller siffra«. Det finns också gott stöd för detta i deras studie. De frågade 78 kroniska smärtpatienter hur de tänkte när de gav ett värde på VAS eller NRS. Det framkom ett otal funderingar och överväganden som ligger bakom VAS-värdet.

Liknande kritik förs fram i en annan studie [10], där 71 patienter fick skatta sin smärta, ångest och depression på en sexgradig NRS. De fick också svara på en smärtenkät, Multidimensional Affect and Pain Survey (MAPS). MAPS innehåller 101 frågor om smärta och känslor indelade i 30 kluster och tre superkluster. Värdet på smärt-NRS predikterades till 41 procent av poängen i den somatosensoriska smärtsuperklustern i MAPS och till 100 procent i den emotionella smärtsuperklustern. Studien talar för att poäng på en enkel smärtskattningsskala styrs mer av känslor än av en ren sensorisk smärtsupplevelse.

Statistik

Tung kritik som brukar föras fram rör hur VAS behandlas statistiskt. I prestigefyllda medicinska tidskrifter behandlas VAS-värden som kontinuerliga, kvantitativa data. I smärtmedicinens tongivande tidskrift, *Pain*, är detta mycket vanligt. I en aktuell lärobok om smärtskattning anges att VAS »förefaller ha parametriska egenskaper för grupper av människor« [2, p. 20].

»Vi mäter ingenting. Vi lyckas inte få fram några objektiva kvantitativa data med hjälp av ett mätinstrument (den mekaniska VAS-linjalen).«

Samma författare, Mark P Jensen, för fram samma åsikt i bakgrundsmaterialet till ett möte med expertgruppen IMMPACT 2003 [11].

Statistiker redovisar motsatt åsikt [12]. Variabler som kan definieras utifrån fysikens eller kemins lagar (tryck, temperatur, längd, koncentration) ger upphov till kontinuerliga, kvantitativa data. Variabler som bygger på konstrukt och som graderas på en skala ger inte kontinuerliga, kvantitativa data. Värdena från skalan har en ordning, men man kan inte säga något om storlek och avstånd, oberoende av om man har en verbal skattningsskala eller en VAS. Detta medför att man inte kan göra matematiska beräkningar på data. Att räkna ut medelvärde är meningslöst. Vanliga statistiska metoder kan inte användas.

Förespråkare för VAS-värdens kontinuerliga, kvantitativa (»parametriska«) egenskaper brukar referera till studier av Donald D Price från Virginia i USA [13, 14]. Han utsatte försökspersoner för gradvis ökande smärta med en värmeplatta på underarmen och registrerade VAS-värden. Utifrån sambandet mellan det standardiserade stimuli och VAS-värdet tyckte han sig se kontinuerliga, kvantitativa egenskaper. Detta gällde dock bara för VAS, inte för NRS. I sina bilder redovisar han medelvärden för VAS som stöd för sin tes. Han förutsätter alltså redan från början linjära egenskaper, vilket är vad han ska undersöka. Det är också tvivelaktigt att utifrån en grupp försökspersoner i en specifik laboratoriesituation dra slutsatser för alla patienter i alla situationer, något som man gör när man behandlar sina VAS-data som kontinuerliga och kvantitativa. Det finns icke-parametriska studier som visat att kliniska VAS-värden inte är linjära utan endast representerar en ordning och inte absolut storlek och avstånd [7, 8].

Ett annat argument som ibland förs fram i diskussioner är att det inte spelar någon roll om man använder en korrekt eller felaktig statistisk metod – resultatet blir ändå detsamma. Det är svårt att hitta någon som argumenterar på detta sätt i text. Det förs dock fram i en översiktsartikel, som refererar till en studie [15]. Försökspersoner fick höra olika starka ljud och gradera dem på en VAS. Resultaten analyserades sedan med parametriska och icke-parametriska statistiska metoder. Slutsatsen var att det vanligtvis inte spelade någon roll vilken metod man väljer. Men detta sätt att tänka är ju ganska konstigt. Man måste självklart använda korrekta och riktiga metoder när man forskar. Korrekt statistik är lika viktigt som all annan forskningsmetodik.

Det finns icke-parametriska statistiska metoder för att analysera data från VAS och andra skalor. Wilcoxon's test av att jämföra mätningar mellan olika grupper är ett exempel och en metod för upprepade mätningar, t ex för analys av reliabilitet, validitet och förändring, som presenterades i *Läkartidningen* 2007 [16].

IMMPACT

Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials, IMMPACT, är ett initiativ för att ta fram riktlinjer för utfallsmått avseende kliniska studier av kronisk smärta och består av representanter för forskare, patientorganisationer och läkemedelsindustri, främst från USA. De anger smärta som en av sex huvudgrupper som man ska överväga när man

»Frestelsen att använda parametriska statistiska metoder blir också mindre om man måste omvandla sina resultat från ord till siffror före analysen.«

designar en studie [17]. När det gäller smärta rekommenderar de att man dels registrerar användningen av sk rescue-preparat, dels använder en smärtskattningsskala. IMMPACT anser att man i första hand ska använda NRS. Man ger inga riktlinjer när det gäller användning av statistiska metoder. Som nämnts ovan anges i en bakgrundsartikel [11] till IMMPACT:s möte 2003 att VAS men inte NRS kan analyseras med parametriska metoder.

Slutsatser och diskussion

Det mest slående när man läser om VAS är hur man behandlar data statistiskt. Det är troligen så att den rådande metoden är inkorrekt. Trots det är det normala sättet att presentera forskningsresultat VAS-medelvärden. Att läkemedelsföretag vill ha lättfattliga och slående reklammaterial är lätt att förstå. Svårare att förstå är hur förespråkare för evidensbaserad medicin kan använda felaktiga statistiska metoder.

I rapporten om långvarig smärta från SBU [18] skriver man att VAS-data inte bör analyseras med parametriska metoder men att det ändå ofta görs i medicinsk klinisk forskning. Man

anger inte hur man själv värderat olika statistiska metoder när man evidensgraderat forskningsresultat. I Läkemedelsverkets nyligen publicerade riktlinjer för behandling av neuropatisk smärta används också VAS-data analyserade med parametriska metoder [19].

Som kliniker bör man se kritiskt på rekommendationer, forskningsrapporter eller läkemedelsreklam som bygger på VAS-medelvärden. Viktigt är också att se på hur stor den »statistiskt signifikanta« skillnaden mellan placebo och aktiv substans är. I forskning bör man använda en korrekt statistisk metod. Ansvar för det gäller smärtskattningsmetoder ligger inte främst på de enskilda forskarna. Det ligger på aktörer med mer makt: medlemmar i IMMPACT, författare till SBU och LVM-rapporter, redaktörer för tidskrifter som Pain och redaktörer för tongivande läroböcker.

Man kan också titta på andra smärtskattningsskalor. Mycket talar för VRS. Det finns flera studier som visar att patienterna föredrar dessa skalor och att de har större validitet och reliabilitet. Dessutom blir det helt klart vad det är som försiggår, dvs vi frågar patienterna hur ont de har. Vi mäter ingenting. Vi lyckas inte få fram några objektiva kvantitativa data med hjälp av ett mätinstrument (den mekaniska VAS-linjalen). Frestelsen att använda parametriska statistiska metoder blir också mindre om man måste omvandla sina resultat från ord till siffror före analysen.

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

REFERENSER

- Engel GL. The need for a new medical model: a challenge for biomedicine. *Science*. 1977;196(4286):129-36.
- Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, editors. *Handbook of pain assessment*. New York: Guilford Press; 1992.
- Aitken RB. Measurements of feelings using visual analogue scales. *Proc R Soc Med*. 1969;62:989-93.
- Hayes MH, Patterson DG. Experimental development of the graphic rating method. *Psychol Bull*. 1921; 18:98-9.
- Freyd M. The graphic rating scale. *J Educ Psychol*. 1923;14:83-102.
- Jensen MP. The validity and reliability of pain measures in adults with cancer. *Pain*. 2003; 4(1):2-21.
- Lund I, Lundeberg T, Sandberg L, Budh CN, Kowalski J, Svensson E. Lack of interchangeability between visual analogue and verbal rating pain scales: a cross sectional description of pain etiology groups. *BMC Medical Research Methodology*. 2005;5:31.
- Svensson E. Concordance between ratings using different scales for the same variable. *Statist Med*. 2000; 19:3483-96.
- de C Williams AC, Davies MT, Chaudury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain*. 2000;85:457-63.
- Knotkova H, Crawford Clark W, MokRejs P, Padour F, Kuhl J. What do ratings on unidimensional pain and emotion scales really mean? A Multidimensional Affect and Pain Survey (MAPS) analysis of cancer patient responses. *J Pain Symptom Manage*. 2004;28(1):19-27.
- <http://www.immpact.org>
- Svensson E. Vad är behandlingseffekten om patienten blev bättre men ingen vet hur mycket? Analys av förändringar när datamaterialet består av ordnade kategorier. *Läkartidningen*. 2005;102(43):3138-45.
- Price DD, McGrath PA, Rafii A, Buchingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain*. 1983;17(1):45-56.
- Price DD, Long S, Harkins SW. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. *Pain*. 1999;456(2):217-26.
- Maxwell C. Sensitivity and accuracy of the visual analogue scale: a psycho-physical classroom experiment. *Br J Clin Pharmacol*. 1978;6(1):15-24.
- Svensson E. Vad är behandlingseffekten om patienten blev bättre men ingen vet hur mycket? Statistisk metod för parade ordinaldata. *Läkartidningen*. 2007;104(8):596-601.
- Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113(1-2):9-19.
- <http://www.sbu.se>
- <http://www.lakemedelsverket.se>