

Att jämföra kliniker, en grannliga hantering

Det är modernt att jämföra resultat vid olika kliniker. Tyvärr görs detta många gånger på ett statistiskt primitivt sätt, menar Adam Taube. Ett exempel är en jämförelse mellan ortopederna i Huddinge och ryggkirurgiska kliniken i Strängnäs, beträffande viss ryggkirurgi. Grupperna i studien är inte randomiserade, materialet är litet och opartiskheten tveksam.

Det är för närvarande på modet att jämföra resultaten vid olika kliniker. Sådana jämförelser anses av många utgöra en viktig komponent i det alltmer omhuldade kvalitetsarbetet. Så har till exempel Stockholm läns landsting satsat 800 000 kronor per år under några år på en utvärdering av viss ryggkirurgi vid ortopederna i Huddinge och ryggkirurgiska kliniken i Strängnäs, dit vissa Stockholmspatienter har remitterats. Resultat från denna utvärdering presenterades helt kortfattat vid Läkarstämman nyligen och ger anledning till kritiska funderingar.

Utvärdering av kirurgisk behandling

Studien gällde 52 patienter med kronisk ryggsmärta varav 30 hade opererats i Strängnäs och 22 i Huddinge. Alla hade en svår smärtproblematik utan säker orsak men misstanke om segmentrelaterad rörelsesmärta förelåg. Patienterna hade själva valt var de ville opereras.

Enligt undersökarna var »grupperna likartade: De patienter som opererades i Strängnäs hade i medeltal haft ont i ryggen i 117 månader (SD 150) medan Huddingepatienterna hade haft besvär i 122 månader (SD 114) i medeltal. Strängnäspatienterna hade i medeltal varit sjukskrivna i 54 månader (SD 64), Huddingepatienterna i 36 månader (SD 38).»

Resultatet presenterades sålunda:

Författare

ADAM TAUBE
professor, institutionen för statistik,
Uppsala universitet.

»Det fanns ingen säker statistiskt påvisbar skillnad i behandlingsresultat mellan grupperna men enbart 50 procent upplevde sig förbättrade, något fler i den grupp som behandlades på Huddinge sjukhus. De stora skillnaderna som påvisades låg i att antalet ingrepp var fler i Strängnäsgruppen. Snittkostnaden för hela behandlingen var på ryggkirurgiska kliniken i Strängnäs ca 600 000 kr, mer än fem gånger högre än motsvarande snittkostnad på Huddinge sjukhus.»

Författarna drog så slutsatsen: »Även om vissa patienter förbättrades av den kirurgiska behandlingen är det tveksamt om de stora insatserna, inte minst de ekonomiska och den ibland tunga konvalescensen uppvägs av de måttliga möjligheterna till förbättring.»

Man har alltså kommit fram till en allmän tveksamhet till behandlingen – oavsett vid vilken klinik den ges. Eftersom de medicinska resultaten tycks ha varit ungefär likvärdiga följer sedan obevekligt den inte rakt ut formulerade slutsatsen att det är dålig affär att skicka patienter till Strängnäs där antalet operationer var större och kostnaderna högre. Att antalet operationer blivit större i Strängnäs sammanhänger förmodligen med att man där gör ingreppen stegvis för att inte steloperera fler segment än nödvändigt.

Ej randomiserade grupper

Det bör observeras att patienterna fördelats på Huddinge respektive Strängnäs efter eget val. Detta medför att grupperna a priori inte kan förutsättas vara jämförbara. I en sådan situation bör man vid den statistiska analysen så långt som möjligt ta hänsyn till möjliga bakomliggande så kallade prognostiska faktorer. Det räcker inte med att påstå att grupperna »är likartade».

Vid presentationen på Läkarstämman viftades den kritiska frågan om (frånvaron av) randomisering bort – det anfördes att eftersom författarna inte hävdade att studien skulle vara randomiserad så behövde den invändningen ej diskuterats. Dessvärre löser inte detta problemet.

Eftersom patienterna fick välja klinik själva är det inte omöjligt att de från

DEBATT

”Kohortstudier med data från diverse specialregister eller insamlade på annat sätt bearbetas av glada amatörer. Här visar sig läkarnas statistiskt metodologiska medvetenhet ibland vara chockerande låg.”

början skiljer sig åt mellan klinikerna. Om till exempel väntetiderna var desamma vet vi ej. Man observerar i detta sammanhang även att fördelningen av patienterna (remisserna) gått via Huddinge. Om det till exempel vore så att svårare fall hamnade i Strängnäs skulle även detta kunna bidra till det större antalet reoperationer där. Ett relevant mått på den initiala svårighetsgraden borde alltså vara en väsentlig prognostisk faktor att inkludera i analysen.

Informationen om hur länge patienterna haft ryggbesvär och hur länge de varit sjukskrivna bör föranleda en djupare penetration. Dessa variabler – vilka ej kan anta negativa värden – visade sig ha standarddeviationer av ungefär samma storlek som sina respektive medelvärden. Det är alltså fråga om extremt sneda fördelningar. Förmodligen omfattar materialet dels en kategori (färska) fall med förhållandevis korta tidsförlopp, dels sådana som har en sjukhistoria som pågått i decennier. Denna heterogenitet kan beaktas vid analysen, till exempel med hjälp av regressionsanalys. Såsom materialet nu presenteras ger det hela ur metodologisk synvinkel ett synnerligen amatörmässigt intryck.

Små material

Varje patient har klassificerats som »förbättrad» eller »ej förbättrad». Ungefär 50 procent har blivit förbättrade utan att grupperna skiljer sig nämnvärt åt. Låt oss som ett räkneexempel anta att det var till exempel 13 av 22 (59 procent) som ansågs förbättrade i Huddingegruppen och 14 av 30 (47 procent)

i Strängnäsgruppen. Differensen skulle i så fall bli 59 procent – 47 procent = 12 procent. Ett konfidensintervall (95 procent) för differensen går då från – 13 procent till + 39 procent. Detta onekligen ganska breda intervall innehåller alla hypotetiska skillnader som ej kan förkastas på grundval av tillgängliga data.

För att påstå att det i någon mening föreligger likhet behövs stora material så att aktuella konfidensintervall »stänger in» möjliga hypotetiska skillnader. Att hävda likhet mellan grupper på grundval av att de ej uppvisar en statistiskt signifikant skillnad är ett elementärt fel som dessvärre inte är ovanligt i medicinska arbeten.

Opertisk värdering?

Författarna till den aktuella utvärderingen kommer från olika institutioner på Huddinge och Karolinska institutet – det sticker i ögonen att ryggekliniken i Strängnäs inte finns representerad. Efter vad jag har kunnat inhämta inviterades Strängnäs-kliniken (1992) till ett samarbete i studien, men man ville där i så fall att själva utvärderingen skulle handhas av någon tredje, oberoende part. Så blev det ej. Strängnäs-kliniken

tvingades ändå tillhandahålla patientmaterialet – annars hade några patienter inte fått remiss dit.

Det är säkerligen möjligt att utvärderarna trots allt varit självständiga visavi Huddinges ortopediska klinik. Detta ändrar dessvärre ingenting – hela företaget hade fått en helt annan trovärdighet om det känsliga uppdraget att bedöma slutresultaten från de bägge involverade klinikerna inte förmedlats via den ena av dessa.

Varför så amatörmässigt?

När ett nytt läkemedel skall prövas och jämföras med till exempel någon standardbehandling kan detta numera endast göras med en studie vars uppläggning (studieprotokoll) i förväg granskats och godkänts av speciella instanser. I detta dokument måste finnas precisering av grundläggande hypoteser, erforderlig materialstorlek, randomiseringsförfarande, metoder för neutral bedömning av behandlingsresultat, efterföljande statistisk analys med mera. Inom det området har det under de senaste decennierna successivt tvingats fram ett starkt statistiskt metodologiskt medvetande bland de involverade läkarna.

Ifråga om randomiserade prövningar av olika operationsmetoder eller jämförelser mellan till exempel strålning och operation tycks motsvarande medvetenhet ej nått samma styrka även om den logiska problemstrukturen här är densamma. Det har här inte förelegat samma tryck från utomstående instanser.

Vid jämförelser mellan kliniker, där de studerade grupperna med nödvändighet ej är randomiserade, tycks verksamheten på många håll bedrivas än mer primitivt. Kohortstudier med data från diverse specialregister eller insamlade på annat sätt bearbetas av glada amatörer. Här visar sig läkarnas statistiskt metodologiska medvetenhet ibland vara chockerande låg.

Den ovan skildrade studien av ryggoperationer grundas på ett häpnadsväckande litet antal observationer vilka kommer från icke randomiserade grupper. Den statistiska argumentationen tyder på en statistisk metodologisk aningslöshet. De myndigheter som beställer och finansierar »evalueringar» av detta slag borde fundera mera på vad det är de verkligen får ut för pengarna – varken patienter, sjukvård eller politiker kan ha nytta av detta slag av amatörmässiga »cost benefit»-studier. •

Replik:

Vissa resultat kunde presenteras

Föredraget baserades på data som nyligen framkommit i en större studie. De skillnader som beskrevs mellan de båda klinikerna baserades inte på statistiska jämförelser – det var tydliga resultat som kunde värderas i och för sig.

Adam Taubes inlägg kan kanske tas som en illustration till den inflammatoriska stämning som ofta försvårat analysen av data när Strängnäs-kliniken berörs. För att läsarna skall få en bättre uppfattning om vad saken gäller har vi summerat resultat och konklusioner ur det aktuella föredraget.

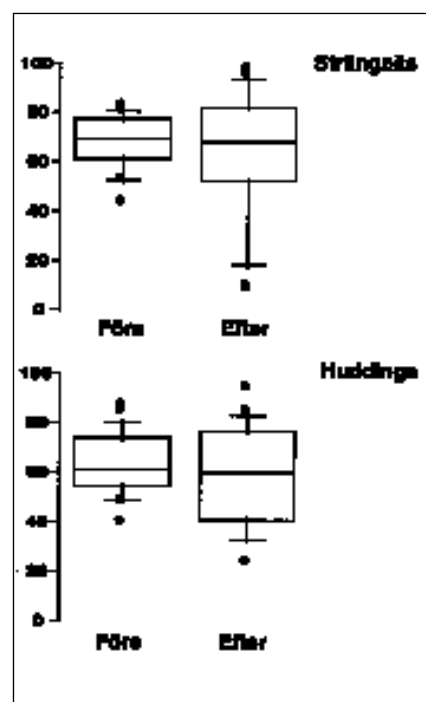
I Strängnäsgruppen opererades 30 patienter inom uppföljningstiden, i Huddingegruppen 22. Medianåldern var 43 respektive 47 år, den yngsta patienten var 22, den äldsta 73 år. Strängnäs-patienterna hade i medeltal haft ont i ryggen i 8 år och varit sjukskrivna eller sjukpensionerade i 4 (median 3) år. Huddingepatienterna hade i medeltal haft ont i ryggen i 10 år och varit sjukskrivna eller sjukpensionerade i 3 (median 2,5) år. 14 Strängnäs-patienter och

11 Huddingepatienter var tidigare ryggopererade. 2 patienter i Strängnäsgruppen har inte kunnat utvärderas 2 år efter operation och räknas därför som bortfall.

Smärtskattningen i båda grupperna är något (men inte signifikant) lägre efter operationen. För skattning av funktion har vi använt Spangforts »Disability rating index», DRI, som består av 12 VAS-skalar för självskattning av olika funktioner. Grupperna ligger ganska lika preoperativt. En viss spridning sker med några patienter som blir förbättrade och några försämrade. Skillnaden är inte signifikant för någon av grupperna (Figur 1).

På direkt skriftlig fråga om förbättring/försämring uppger sig 8 Strängnäs-patienter och 4 Huddingepatienter mycket förbättrade. 9 Strängnäs-patienter och 6 Huddingepatienter anger sig försämrade.

Återgången i arbete tycks delvis återspegla andra faktorer än smärtskattning. 4 Strängnäs- och 2 Huddingepati-



Figur 1. Funktionsskattning enligt DRI.

ANNONS

enter arbetade i någon omfattning före operationen (25–100 procent). Vid uppföljningen arbetade 3 Strängnäspatienter och 8 Huddingepatienter i någon omfattning.

Fakturerad kostnad per opererad patient på Huddinge har i genomsnitt varit 85 000 kr/patient. Fakturerad kostnad per opererad patient på Strängnäs har i genomsnitt varit 660 000 kr/patient.

I summorna ingår slutenvårdskostnad, öppenvårdskostnad vid återbesök och delvis kostnader för röntgenundersökningar. Andra vårdkostnader, t ex sjukgymnastik, sjukresor, läkemedel, hembesök av sjukvårdspersonal, är inte inberäknade.

Vid en genomgång av Huddinge sjukhus självkostnadsdata för 1992 och 1993 har vi sett att 13 patienter år 1992 är underersatta med 38 000 kr och 9 patienter år 1993 är underersatta med 5 700 kronor. (Detta ökar summan per patient på Huddinge sjukhus till 110 000.) En del av kostnadsskillnaden uppkommer genom att man på Strängnäs kliniken tar andra vårdbeslut som bland annat ger en högre reoperationsfrekvens.

Det så kallade KSP-formuläret har använts i ett flertal studier för att beskriva personlighetsdrag hos olika patientgrupper, däribland exempelvis flera smärtgrupper. I denna undersökning uppvisar patientgruppen normala värden i 9 av 15 KSP-skolor. De moderata avvikelser från normdata som finns tyder på att individerna i gruppen generellt är mera extroverta (det vill säga sociala, självhävdande och monotoniundvikande) samt mera spänningsbenägna och stresskänsliga.

Två tidigare prospektiva studier av hur personlighetsdrag mätt med KSP kan predicera utfall efter kirurgi har visat på substantiella samband. Här har vi funnit att de personer som minskat smärtskattningen mest, mätt med VAS, är de som initialt hade bättre stresstålighet och ork. Vad gäller skattningar av funktion och smärtbeteende finner vi inga samband mellan personlighet och klinisk förändring.

När patienterna själva efter 2 år har fått skatta utfallet, så finner vi ett samband mellan självhävdade preoperativa personlighetsdrag och skattad förbättring.

Behandlingen dyrbar men ej särskilt effektiv

Trots att flera av de undersökta patienterna förbättrats gör risken för försämring och den ibland tunga och långa konvalescensen efter operationen att vi känner en viss tveksamhet till denna typ av behandling. Den diagnostiska arsenalen har utvecklats mycket under de senaste årtiondena men det finns fortfa-

rande inte möjligheter för en så exakt diagnostik av smärttillstånd i ländryggen att man på ett tillfredsställande sätt har kunnat välja ut en lämplig målgrupp för kirurgisk behandling vid denna typ av långvarig inverterad ryggsmärta. I detta skede av studien har resultat framkommit som kan vara av intresse för »myndigheter» såväl som läkare – bland annat att den kirurgiska behandlingen inte var särskilt effektiv och att den var dyrbar. Man kan med fog hävda att kostnaderna var särskilt höga på Strängnäs kliniken. Slutligen är det vår mening att en djupgående multidisciplinär analys rörande dessa patienter med exceptionellt uttalat ryggglidande är väl motiverad.

Urban Lindgren

professor, enheten för ortopedi, Huddinge sjukhus

Petter Gustavsson

med dr, institutionen för klinisk neurovetenskap, Karolinska sjukhuset

Einar Kristjansson

leg läk, psykiatriska kliniken, Huddinge sjukhus

Jan Persson

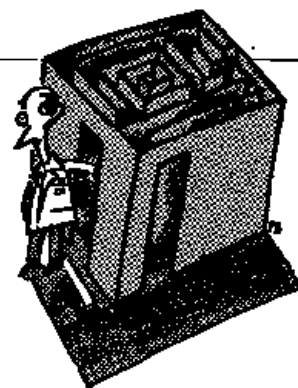
leg läk, smärtenheten, Huddinge sjukhus

Åsa Storkamp

leg läk, rehabiliteringsmedicinska kliniken, Huddinge sjukhus

Gunnar Öhlen

med dr, ortopedkliniken, Huddinge sjukhus.



**enligt
min
erfarenhet**

Läkartidningens serie 1990–1992 i särtryck

När konsensus saknas om hur läkaren bör behandla, spelar den beprövade erfarenheten stor roll. Det 48-sidiga häftet innehåller 32 korta, praktiskt inriktade artiklar med anknytning till vårdens vardag och vänder sig till alla kliniskt verksamma läkare. Förutom diagnostik med terapi speglas goda exempel på prevention, ledningsfrågor och administration.

Pris 55 kr. Vid 11–50 ex 50 kr, vid högre upplagor 47 kr/ex.

Beställ här:

..... exemplar av
Enligt min erfarenhet

Namn

Adress.....

Sändes till Läkartidningen,
Box 5603, 114 86 Stockholm

Märk gärna kuvertet med
»Enligt min erfarenhet»

Beställning per fax:
08-20 76 19