

**HFS**  
**HÄLSOVINSTMÄTNINGSPROJEKT**

**delrapport 1**

Evalill Nilsson, Preben Bendtsen & Margareta Kristenson

för nätverket  
HälsaFrämjande Sjukhus

Linköping mars 2006

# INNEHÅLLSFÖRTECKNING

<b>SAMMANFATTNING</b> .....	<b>3</b>
<b>BAKGRUND</b> .....	<b>4</b>
<b>METOD</b> .....	<b>5</b>
PLANERING .....	5
SJKDOMSGRUPPER .....	5
IN/EXKLUSIONSKRITERIER .....	5
MÄTNINGARNA .....	5
PROCESSANTECKNINGAR .....	9
STATISTISKA METODER .....	10
<b>RESULTAT</b> .....	<b>11</b>
RESULTAT FRÅN DIABETESDELPROJEKTET (DELPROJEKT 3 I ÖVRIG-GRUPPEN) .....	12
RESULTAT FRÅN ÖVRIGA DELPROJEKT .....	17
RESULTAT FRÅN UTVÄRDERINGSENKÄTEN .....	24
<b>DISKUSSION</b> .....	<b>30</b>
METODDISKUSSION .....	32
<b>KONKLUSION</b> .....	<b>34</b>
<b>REFERENSER</b> .....	<b>35</b>
<b>BILAGOR</b> .....	<b>37</b>

## SAMMANFATTNING

### BAKGRUND

Mått på hälsorelaterad livskvalitet (HRQoL) används idag allt oftare för monitorering av en befolknings hälsa, vid kliniska prövningar och som utfallsmått i hälso- och sjukvården. Två viktiga instrument för att mäta HRQoL som ofta används i Sverige idag är SF-36 (Short Form-36) och EQ-5D (EuroQoL). Det svenska WHO-nätverket Hälsofrämjande Sjukhus (HFS) arbetar för en hälsoorientering av den svenska hälso- och sjukvården, där utvecklingsarbetet med hälsovinstmätning inom rutinverksamheten ingår som en viktig del.

### SYFTE

Under år 2004 har HFS påbörjat ett läroprojekt med följande syften: att skapa ökad kunskap om processen vid hälsovinstmätning ur patientens och medarbetarnas perspektiv, att pröva i vad mån kunskap om patienternas hälsorelaterade livskvalitet efter behandling kan användas som stöd för verksamhetsutveckling, samt att jämföra de två HRQoL-instrumenten SF-36 och EQ-5D i dessa avseenden.

### METOD

Varje medlemssjukhus i HFS skulle initialt genomföra två delprojekt om minst 20 patienter vardera, främst från följande fem sjukdomsgrupper: kronisk obstruktiv lungsjukdom (KOL), smärtrehabilitering, hjärtinfarkt, knä/höftkirurgi, och reumatologi, men det fanns även möjlighet att välja andra områden.

Inom varje delprojekt gjorde man mätningar före, och en tid efter, en behandling/intervention, genom att deltagarna fick besvara ett enkätbrev bestående av patientkaraktäristik, HRQoL-instrumenten EQ-5D och SF-36 samt, vid eftermätningen, en utvärderingsenkät (för att få en bild av hur patienterna ställer sig till hälsovinstmätning och de två olika instrumenten).

### RESULTAT OCH DISKUSSION

Detta är projektets första delrapport, med huvudfokus på patientperspektivet. Medarbetarperspektivet och verksamhetsutvecklingspotentialen blir huvudfokus för projektets nästa fas.

18 av 23 medlemssjukhus deltog med ett eller flera delprojekt i den första fasen. Totalt gjordes mätningar på 615 individer vid första mättillfället och 503 vid andra mättillfället. Utvärderingsenkäten visade att över hälften av deltagarna ansåg att hälsovinstmätning är värdefullt, medan endast 5% sade sig inte se något värde i mätningarna. Valet av instrument verkade inte vara avgörande då 72% svarade att det inte spelade någon roll vilket instrument som användes. 20% föredrog SF-36 medan 8% föredrog EQ-5D. De flesta tyckte dessutom att (båda) instrumenten var lätta att använda, och gav god möjlighet till att få beskriva ens hälsotillstånd.

Resultatet av själva mätningarna ger kunskap om utvecklingen av patienternas (egenupplevda) hälsa efter avslutad behandling. Trots små delprojekt kunde man se tydliga skillnader mellan sjukdomsgrupperna. Hjärtinfarktgruppen uppvisade en mycket likartad bild för de ingående delprojekten, med klara förbättringar efter intervention, medan andra, tex smärtgruppen, uppvisade en mer inhomogen bild, där vissa delprojekt uppvisade förbättringar och andra inte, kanske beroende på mer skilda interventionsmodeller på de olika sjukhusen. Även om SF-36 och EQ-5D överlag uppvisade likartade övergripande bilder över patienternas HRQoL, gav SF-36, som förväntat, en mer nyanserad bild än EQ-5D. Generellt såg man att för EQ-5D var det dimensionen *smärtor/besvär*, som dominerade bilden inom varje sjukdomsgrupp, utom möjligen för KOL-patienterna. När det gäller SF-36 hade sjukdomsgrupperna överlag lägst värden för *fysisk rollfunktion*, särskilt vid föremätningen.

## BAKGRUND

Mått på hälsorelaterad livskvalitet (HRQoL, Health Related Quality of Life) används idag allt oftare för monitorering av en befolknings hälsa, vid kliniska prövningar och som utfallsmått inom hälso- och sjukvården [1]. I de nationella kvalitetsregisterna blir det allt vanligare att inkludera patientsubjektiva mått, vilket även stöds från sjukvårdspolitiskt håll. HRQoL mäts med internationellt erkända, valida och reliabla instrument, som kan vara antingen generella (efterfrågar generella besvär) eller sjukdomsspecifika (efterfrågar besvär som följer med en specifik sjukdom). De förra (som ibland även kallas generiska) har fördelen att de medger möjligheter till jämförelser mellan olika sjukdomsgrupper, medan de senare används för att bättre försöka fånga alla aspekter av en specifik sjukdom.

Två viktiga generella instrument som ofta används i Sverige idag är SF-36 (Short Form-36) och EQ-5D (fd EuroQoL). SF-36, som är utvecklat i USA, är troligen det internationellt sett mest använda generiska HRQoL-instrumentet. Det bygger på erfarenheter från ett välkänt, mer omfattande formulär, MOS (Medical Outcome Study) med 40 skalor och över hundra frågor, som ska spegla WHO:s hälsodefinition av hälsa som fysiskt, psykiskt och socialt välbefinnande, och inte enbart frånvaro av sjukdom. Intentionen att skapa ett kortare och mer användarvänligt instrument resulterade i 36 frågor fördelade på åtta skalor (dimensioner), utvalda bla för sin starka korrelation till sjukdom. De ändå relativt många frågorna i SF-36 ansågs nödvändiga för att kunna få med alla viktiga aspekter i varje skala [2-6]. EQ-5D (5 Dimensions), som förut hette EuroQoL (vilket också är namnet på den europeiska forskningssamarbetsgrupp som skapat instrumentet), är tänkt att istället, utifrån fem olika hälsodimensioner, skapa ett summamått på en individs HRQoL, för att ge möjlighet till hälsoekonomiska utvärderingar [7-10]. Grundtanken var att EQ-5D skulle vara ett komplement till andra, mer utförliga HRQoL-instrument.

De båda instrumentens användningsområden har emellertid på senare tid allt mer närmats sig varandra. EQ-5D har börjat ”stå på egna ben” och används nu ofta som enda HRQoL-instrument i både utvärderings- och monitoreringssyfte, och utifrån SF-36 har SF-6D (6 Dimensions) utvecklats för att, precis som EQ-5D, kunna användas för hälsoekonomiska beräkningar. SF-6D kan även appliceras på redan existerande SF-36-dataset [11]. Både SF-36 och EQ-5D har sina förespråkare och ingen konsensus råder om vilket instrument som skulle vara bäst eller om det skiljer i olika patientgrupper och/eller situationer. Båda instrumenten har i ett befolkningsperspektiv visat sig vara lika bra på att skilja frisk från sjuk. De diskriminerar också mellan sjukdomsgrupper, även om det finns vissa skillnader mellan instrumenten [12]. I forskningssammanhang har de genom åren visat sig vara mycket användbara som utfallsmått. Däremot kvarstår frågan om de också är lika bra som utfallsmått inom hälso- och sjukvårdens rutinverksamhet, dvs hur de, gällande både användarvänlighet och informationsinnehåll, kan fungera som kunskapsstöd för verksamhetsutveckling.

Det svenska WHO-nätverket Hälsöfrämjande Sjukhus (HFS) arbetar för en hälsoorientering av den svenska hälso- och sjukvården, där utvecklingsarbetet med hälsovinstmätning inom rutinverksamheten ingår som en viktig del. Under år 2004 har därför HFS påbörjat ett läroprojekt med följande syften:

- att skapa ökad kunskap om processen vid hälsovinstmätning ur patientens och medarbetarnas perspektiv
- att pröva i vad mån kunskap om patienternas hälsorelaterade livskvalitet efter behandling kan användas som stöd för verksamhetsutveckling
- att jämföra de två HRQoL-instrumenten SF-36 och EQ-5D i ovanstående avseenden

## **METOD**

### ***Planering***

Medlemssjukhusen inom HFS har olika grad av erfarenhet från användning av SF-36 och EQ-5D för hälsövinstmätning. Vid ett processledarmöte inom nätverket HFS i maj 2004 beslöts därför att nätverket skulle starta ett gemensamt läroprojekt inom hälsövinstmätning, för att lära av varandra. Man valde att koncentrera sig på ett mindre antal sjukdomsgrupper och att hålla delprojekten (se nedan) små, för att kunna fokusera på själva lärandet. HFS sekretariatet skulle bistå med praktiska göromål, såsom tryckning och distribution av enkäterna, och inscanning och bearbetning av data.

### ***Sjukdomsgrupper***

Varje medlemssjukhus i HFS (23 sjukhus när projektet startade) skulle genomföra två delprojekt om minst 20 patienter vardera. Vid processledarmötet valde man ut följande fem områden, där nätverket ansåg sig ha möjlighet att ganska omgående få igång tillräckligt många delprojekt (med tillräckligt antal deltagare):

- hjärtinfarkt
- smärtrehabilitering
- kronisk obstruktiv lungsjukdom (KOL)
- reumatologi
- knä/höftkirurgi

men det fanns även möjlighet att få välja andra sjukdomsgrupper.

### ***In/exklusionskriterier***

Alla patienter som skulle genomgå en specifik utvald intervention (patientskola, rehabgrupp eller motsvarande) och som accepterade deltagande inkluderades. Dementa och konfusoriska patienter samt de under 16 år och de med svårigheter att förstå svenska språket exkluderades.

### ***Mätningarna***

Inom varje delprojekt gjorde man mätningar före en behandling/intervention och, beroende på interventionens innehåll, en viss tid efter, genom att deltagarna fick besvara ett enkäthäfte (bilaga 1a-e) bestående av (i denna ordning):

1. ett blad med patientkaraktäristik
2. HRQoL-instrumenten EQ-5D och SF-36
3. vid eftermätningen, en utvärderingsenkät

Ett sjukhus var redan igång med datoriserade mätningar för SF-36 sedan tidigare och har därför använt sig av sina egna pekskärmar för att mäta SF-36 och av projektets pappersenkäter för att mäta EQ-5D.

## 1. Patientkaraktäristik

Deltagarna fick själva ange ålder, kön och sjukdomar (dvs även sjukdomar utöver den som behandlingen/interventionen gällde), det senare för att kunna studera effekter av att ha mer än en sjukdom och även olika sjukdomskombinationer (ko/multimorbiditet). Bladet var något olika utformat vid före- och eftermätningarna (bilaga 1a och 1b).

## 2. Instrumenten

**EQ-5D** (bilaga 1c) är ett så kallat preferensbaserat mått, som innehåller två delar. I den första får respondenten ange sin funktionsförmåga för dagen inom var och en av följande fem hälsodimensioner:

- **rörlighet**
- **hygien**
- **huvudsakliga aktiviteter**
- **smärtor/besvär**
- **oro/nedstämdhet**

Varje dimension innehåller tre påståenden, som indikerar om man har inga, måttliga eller svåra problem inom just den dimensionen. Man väljer ett av dessa alternativ för var och en av de fem dimensionerna, vilket medför totalt  $3^5 = 243$  olika svarskombinationsmöjligheter för instrumentet. Till varje kombination har kopplats en sk livskvalitetsvikt, vilket innebär att varje svarskombination har värderats i förhållande till full hälsa (med hjälp av den sk Time Trade-Off metoden). De vikter vi oftast använder i Sverige idag (och även i detta projekt) har tagits fram i England, men det finns även andra viktsystem (tex från Danmark).

Ett sammanfattande index mellan 0 och 1 räknas sedan fram, som kan användas för hälsoekonomiska beräkningar. Egentligen kan indexet även ge negativa värden, men dessa sammanförs oftast till värdet 0. För att få fram ett index krävs att det finns ett svar från var och en av de fem dimensionerna, vilket gör instrumentet mycket känsligt för svarsbortfall. EQ-5D har dock oftast endast ett litet internt bortfall.

Förutom genom indexvärdet kan EQ-5D också redovisas i sina fem dimensioner, vilket ger möjlighet att avgöra inom vilka dimensioner problemen föreligger, något som indexvärdet inte ger svar på. Dimensionerna presenteras på individnivå genom den femsiffriga kombinationen av svarspoängen 1, 2 eller 3 för varje dimension (ex 11112), och på gruppnivå genom den procentuella fördelningen av problem inom varje dimension.

Det saknas ”officiella” svenska normvärden för EQ-5D index, men i en svensk befolkningsenkät från Stockholms län 2001 (3069 män och kvinnor, 20-88 år) har man redovisat ett indexvärde på 0,84 [13], och i en befolkningsenkät från Östergötlands och Kalmars län 2002 (9131 män och kvinnor, 20-74 år) har man funnit precis samma siffra [14]. I båda dessa studier kan man dessutom finna populationssiffror för den procentuella fördelningen av problem inom varje dimension. Mest problem rapporteras inom dimensionen *smärtor/besvär* och minst inom *hygien*.

I den andra delen av instrumentet får respondenten skatta sin allmänna hälsa för dagen på en 20 cm vertikal skala från 1-100 kallad VAS (Visual Analogue Scale), även om det egentligen inte är en äkta VAS (=en skala utan synlig gradering förutom ändvärdena) utan snarare liknar en NRS (Numeric Rating Scale, en numeriskt graderad skala). Skalan har liknats vid en termometer och kallas följaktligen också ofta för EQ-5D termometern.

EQ-5D har uppvisat god validitet (att ett instrument verkligen klarar av att mäta precis det begrepp, ex hälsorelaterad livskvalitet, som det utger sig för att kunna mäta). Reliabiliteten (att instrumentet är tillräckligt noggrant vid mätningarna, så att resultatet blir tillförlitligt och så fritt som möjligt från slumpmässiga variationer) har prövats med den sk test-retestmetoden, vilken har visat på god stabilitet för instrumentet [7].

**SF-36** (bilaga 1d) är ett så kallat profilbaserat mått. Respondenten svarar på 36 frågor om sitt hälsotillstånd de senaste fyra veckorna ("standardversionen", som vi använt i detta projekt) eller den senaste veckan ("akutversionen"), varav 35 sedan, med hjälp av en speciell beräkningsalgoritm, förs samman i åtta olika skalor:

- **PF** fysisk funktion (fråga 3a-j)
- **RP** rollfunktion-fysisk (fråga 4a-d)
- **BP** kroppslig smärta (fråga 7, 8)
- **GH** allmän hälsoupplevelse (fråga 1, 11a-d)
- **VT** vitalitet (fråga 9a, e, g, i)
- **SF** social funktion (fråga 6, 10)
- **RE** rollfunktion-emotionell (fråga 5a-c)
- **MH** psykiskt välbefinnande (fråga 9b, c, d, f, h)

Förkortningarna för skalorna härrör från de engelska beteckningarna. Fråga nr 2, den tidsjämförande frågan, utelämnas från beräkningarna, men redovisas ofta separat.

De olika skalorna åskådliggörs sedan som en profil, och det finns olika presentationsmodeller. Den, åtminstone hittills, vanligaste innebär en transformering av svars-poäng från de olika frågorna till en svarsskala från 0 (sämst tänkbara) till 100 (bäst tänkbara), för var och en av de åtta skalorna, som sedan redovisas i linje- eller stapeldiagram. Man brukar placera skalorna i den redovisade ordningen ovan, med den mest fysiskt orienterade skalan (PF) till vänster och den mest psykosociala/psykologiska (MH) till höger. I mitten återfinns GH och VT, som är de två skalor som har ungefär lika stora både fysiska och psykiska aspekter [15]. Man brukar ändå räkna GH till de fysiskt orienterade skalorna och VT till de psykosociala. Varje skala omfattar olika många frågor, med olika antal svarsalternativ, och vissa skalor är bipolära (GH, VT och MH), vilket gör att man inte direkt kan jämföra de olika skalorna med varandra. Med bipolär menas att det finns ett "neutralt" mellanläge, med avsaknad av negativa aspekter men utan att man känner sig "på topp", vilket skulle ha medfört full poäng på skalan. Detta medför lägre medelvärden på dessa skalor än på de skalor där det räcker att ha avsaknad av negativa aspekter för att få full poäng.

Norm-Based Scoring (NBS), däremot, innebär att man ger alla skalor normvärdet 50 med standarddeviationen 10. Detta medför att man snabbt och enkelt kan avgöra om ett mätvärde skiljer sig från normvärdet, och dessutom kan alla skalor jämföras direkt med varandra. Man måste dock hålla i minnet att variationsvidden (skillnaden mellan högsta och lägsta värde) för skalorna med NBS-metoden inte längre är 0-100, utan istället nu varierar något från skala till skala. Minsta värde når aldrig så lågt som 0 och högsta värde överstiger inte 70 för någon skala.

I denna rapport har vi, för 0-100 presentationsmodellen, använt oss av normvärden ur manualen för SF-36 (8930 män och kvinnor, 15-74 år) från början av 90-talet [4]. Göteborgsgruppen som har licensen för SF-36 i Sverige planerar enligt uppgift en ny

befolkningsstudie 2006, som ska ge uppdaterade normvärden. För NBS presentationsmodellen har vi använt amerikanska normvärden. Svenska värden vore naturligtvis att föredra även här, men vi har denna gång tagit med NBS-graferna mest för att presentera modellen.

Till skillnad från EQ-5D har SF-36 en algoritm för att hantera svarsbortfall, vilket gör instrumentet mindre känsligt för enstaka saknade värden. Enligt manualen brukar det interna bortfallet inte överstiga 2% [4].

Precis som EQ-5D har SF-36 uppvisat god validitet. Eftersom SF-36 har en annan uppbyggnad än EQ-5D kan man när det gäller reliabilitet istället för (eller utöver) att använda test-retestmetoden mäta instrumentets sk interna konsistens (Cronbach  $\alpha$ ). För SF-36 har man funnit att  $\alpha > 0,7$  för alla skalor, vilket brukar vara kravet för att anse mätningarna tillförlitliga på gruppnivå. Flera skalor har dessutom  $\alpha$ -värden upp mot (och ett par även över) 0,9, vilket brukar anses vara nödvändigt för säkra analyser även på individnivå [4].

### **Kända för- och nackdelar med instrumenten**

SF-36 anses ibland vara för långt och krångligt att fylla i. Däremot upplevs frågorna som relevanta även för relativt friska personer att svara på, medan EQ-5D däremot kan upplevas som irrelevant för andra än mycket sjuka personer.

Förutom validitet och reliabilitet är ett instruments responsivitet (känslighet för att mäta förändringar) mycket viktigt för dess användbarhet inom hälsovinstmätning [16]. Det faktum att EQ-5D bara har fem hälsodimensioner, och endast tre svarsnivåer inom varje, minskar dess möjlighet att vara känsligt även för små förändringar. Ibland rekommenderas att man bara bör använda EQ-5D i de fall där man kan förvänta sig någorlunda stora förändringar [17]. Studier har visat att inom samma EQ-5D nivå (måttliga problem) finns en stor variationsvidd i hur man svarat på SF-36, som alltså var mycket mer fingeraderat [18]. EuroQoL-gruppen har tidigare aviserat att de arbetar på en EQ-5D-version med fem svarsnivåer istället för tre, men den senaste informationen är att en sådan version dröjer [10]. Ett sätt att mäta responsiviteten är att undersöka samvariationen med andra (sedan tidigare accepterade) mått på förändringar hos patienten. En aktuell diskussionsfråga är i vilken mån statistiskt signifikanta skillnader i HRQoL också är kliniskt signifikanta, just i betydelsen att de samvarierar med att patientens sjukdomsgrad förbättras [19]. En risk med denna, i övrigt nödvändiga och viktiga, diskussion är att man kopplar förändringar i HRQoL-mått till förändringar i traditionella medicinska mått utan att reflektera tillräckligt över att HRQoL var menat att mäta något utöver det biomedicinska, och sällan korrelerar väl med den typen av mått [20]. Mer kunskap om instrumentens responsivitet är önskvärd för framtiden.

En stor del av resultatet för EQ-5D index beror på en enda dimension, *smärtor/besvär*. Det kan vara en svaghet att en enda dimension på detta sätt mäter i princip alla typer av ohälsa, men å andra sidan så inrymmer just den dimensionen sjukdomar och besvär som ej täcks av övriga dimensioner (och inte ens alltid av frågorna i SF-36), vilket gör att det faktiskt också kan vara en fördel [12]. Det var också mellan dimensionen *smärtor/besvär* och skalan BP i SF-36 som den största variationsvidden i studien ovan kunde ses [18].

Förfarandet vid uträkningen av EQ-5D-index har ibland ifrågasatts eftersom värderingen utförs av friska personer. Studier har nämligen visat att friska personer tenderar att skatta livskvalitet vid olika tänkta sjukdomstillstånd lägre än de patienter gör, som de facto befinner sig i tillstånden [21].



Rollfunktionsskalorna RP och RE har kraftiga tak och golfeffekter (= många hamnar på bästa respektive sämsta värde), som kan göra dem mindre lämpliga vid tidsjämförelser med före- och eftermätningar. Problemet med tak- och golfeffekter har minskat betydligt i den nya SF-36 version 2, och dessutom har fler skalor ett Cronbach  $\alpha$  runt 0,90 (=säkrare för mätning på individnivå) i version 2 [22]. Tyvärr görs svenska normdata för denna nya version av olika skäl inte tillgängliga för närvarande. Utan normdata kan man heller inte välja presentationsmodellen NBS.

Även EQ-5D har kraftiga takeffekter (kraftigare än SF-36), vilket ett byte till fem svarsnivåer skulle kunna tänkas avhjälpa i framtiden. Dock ökar då antalet möjliga svarskombinationer drastiskt, vilket medför andra svårigheter (viktningproceduren).

Det finns även två summamått (framtagna med hjälp av faktoranalys) för SF-36, PCS (Physical Component Scale) och MCS (Mental Component Scale) [15]. Vi har valt att inte använda oss av dessa mått eftersom principen bakom deras uträkningsalgoritm har utsatts för en del kritik [23]. Algoritmen har byggts så att högst värde för PCS fås när man ligger högt på de mer fysiska skalorna och samtidigt lågt på de mer psykosociala. Detta leder till att vid mycket låga värden för de psykosociala skalorna kan ett högt värde för PCS snarare komma att avspegla den dåliga psykiska hälsan än en sann god fysisk hälsa. Det omvända gäller för MCS. Ska man ändå använda dem är det alltså viktigt att använda och tolka dem tillsammans med de åtta originalskalorna, och inte fristående, och då förlorar de, enligt vår mening, en aning i betydelse.

### *3. Utvärdering (enbart vid eftermätningen)*

För att få en bild av vad patienterna tyckte om de två HRQoL-instrumenten, och hur de ställer sig till hälsovinstmätning överlag, ställdes vid eftermätningen följande tre frågor, separat för varje instrument

1. ”Var det lätt eller svårt att förstå frågorna?”
2. ”Var det lätt eller svårt att välja svarsalternativ?”
3. ”Gav enkäten dig möjlighet att beskriva ditt hälsotillstånd på ett fullständigt sätt?”

som följdes av ytterligare två frågor

4. ”Skulle du se något värde i att man inom sjukvården rutinmässigt använder sig av dessa, eller liknande, enkäter i framtiden för att mäta hälsa?”
5. ”Om ja, vilken enkät vore då att föredra?”

Utvärderingen avslutades med en öppen fråga där deltagarna ombads att gärna motivera varför. Tanken var att denna motivering skulle gälla varför man, i förekommande fall, föredrog ett visst instrument (bilaga 1e).

### ***Processanteckningar***

För att identifiera hinder och framgångsfaktorer, både gällande mätningproceduren och tolkning och användning av själva mätresultaten, uppmanades de delprojektansvariga vid varje sjukhus att föra processanteckningar under projektets gång. Dessa kunde föras under fria former, dvs det fanns inga i förväg formulerade frågor som skulle besvaras.

### ***Statistiska metoder***

95%-iga konfidensintervall (=med 95% säkerhet ligger det sanna medelvärdet inom detta intervall) för SF-36 och EQ-5D medelvärden beräknades för varje delprojekt. Konfidensintervallen var överlag vida, som de oftast är när medelvärdet beräknats utifrån ett litet antal mätvärden. Ju smalare ett konfidensintervall är, desto bättre har man varit på att ringa in det sanna medelvärdet. Ofta använder man konfidensintervall för att studera skillnader mellan olika (oberoende) gruppers medelvärden. Om konfidensintervallen inte överlappar varandra kan man med viss säkerhet (95% säkerhet för ett 95%-igt intervall, osv) säga att det finns en sann skillnad mellan medelvärdena.

För att studera skillnader i HRQoL vid före- och eftermätningarna i detta projekt gjordes statistisk signifikansprövning (med signifikansnivån 0,05). Eftersom mätningarna har skett på samma individer, fast vid olika tidpunkter, kan mätningarna anses vara beroende. Parat t-test har därför använts för SF-36, EQ-5D index och VAS, medan Wilcoxon Signed-Rank test har använts för EQ-5D dimensionerna. Det senare testet har också använts för de tre första utvärderingsfrågorna.

För utvärderingen delades deltagarna också in i tre åldersgrupper: 16-39 år, 40-64 år, och 65 år eller äldre.

Förutom för varje delprojekt separat har signifikanstestning gjorts efter sammanläggning av alla delprojekt inom varje sjukdomsgrupp. En försiktig tolkning av dessa sammanlagda värden rekommenderas, eftersom patientgrupperna vid de olika sjukhusen inte är helt jämförbara. Interventionerna de genomgått är inte identiska, mätningarna har skett vid olika tidpunkter efter interventionens slut och patienterna kan eventuellt skilja avsevärt i sjukdomsgrad och övrig sjukdomsbild.

## RESULTAT

Sammanlagt genomfördes 33 delprojekt från 18 olika medlemssjukhus runt om i landet, från Piteå i norr till Trelleborg i söder. Totalt deltog 615 individer vid första mättillfället och 505 vid andra mättillfället. Alla delprojekt lyckades inte få ihop 20 st deltagare. Delprojektens fördelning redovisas i tabell 1.

**Tabell 1.** Fördelning av delprojekt inom varje sjukdomsgrupp.

<i>Delprojekt Hjärtinfarkt</i>		<i>Delprojekt Smärta</i>		<i>Delprojekt KOL</i>		<i>Delprojekt Reuma</i>		<i>Delprojekt Knä/höftkirurgi</i>		<i>Delprojekt Övriga</i>	
<b>1</b>	25/16	<b>1</b>	8/8	<b>1</b>	20/18	<b>1</b>	50/30	<b>1</b>	18/18	<b>1</b>	9/9
<b>2</b>	21/15	<b>2</b>	7/5	<b>2</b>	16/12	<b>2</b>	11/11	<b>2</b>	22/22	<b>2</b>	33/32
<b>3</b>	24/22	<b>3</b>	19/16	<b>3</b>	6/5	<b>3</b>	20/18	<b>3</b>	15/13	<b>3</b>	17/17
<b>4</b>	19/18	<b>4</b>	10/10	<b>4</b>	15/13	<b>4</b>	28/26	<b>4</b>	16/10	<b>4</b>	23/19
<b>5</b>	26/22			<b>5</b>	13/9			<b>5</b>	20/17	<b>5</b>	7/6
<b>6</b>	22/19			<b>6</b>	13/13						
<b>7</b>	22/18										
<b>8</b>	25/18										
<b>9</b>	15/--										
<b>S:a</b>	199/148	<b>S:a</b>	44/39	<b>S:a</b>	83/70	<b>S:a</b>	109/85	<b>S:a</b>	91/80	<b>S:a</b>	*

Siffrorna (n/n) anger antalet deltagare inom varje delprojekt vid före- respektive eftermätningen. \* Ingen summering pga skilda typer av delprojekt.

Förutom från de fem förbestämda sjukdomsgrupperna genomfördes alltså fem ytterligare delprojekt: två medarbetargrupper (1,2), en diabetesgrupp (3) som ordinerats Fysisk aktivitet på Recept (FaR), en hjärtarytmigrupp (4) och en hjärtsviktsgrupp (5). Ett sjukhus hade tre hjärtinfarktgrupper (delprojekt 6, 7 och 8), med något olika interventioner (se bilaga 2). Delprojekt 6 är det av dessa tre som liknar övriga delprojekt inom denna sjukdomsgrupp, och redovisas därför tillsammans med delprojekt 1-5 nedan (figur 8).

Flera av delprojekten uppvisar enstaka bortfall vid eftermätningen, men endast ett (en hjärtinfarktgrupp, delprojekt 9) hade så stort bortfall att eftermätningen blev meningslös. Det exakta antalet deltagare vid varje mättillfälle går att utläsa ur tabell 1. I figur 8-12 redovisas endast resultat för de deltagare som genomförde både före- och eftermätningarna, och vars identitetsmärkning av enkäterna var fullständigt korrekt.

Fördelningen åldersmässigt (16-39/40-64/65- år) vid första mättillfället var 54/249/191 och vid andra mättillfället 37/173/147, och fördelningen könsmässigt (kvinnor/män) var 254/234 respektive 173/177. Tyvärr saknas en del uppgifter om kön och ålder (ca 20% vid första mätningen och ca 30% vid andra), ofta för hela delprojekt. Våra komorbiditetsdata (data om övriga sjukdomar) visade sig tyvärr vara bristfälliga och har därför ej använts som planerat i detta skede.

Varje sjukhus har fått sina egna resultat sig tillsända, här redovisas endast det anonyma, sammanställda resultatet för sjukdomsgrupperna. En beskrivning av de olika delprojektens interventioner finns i bilaga 2.

### **Resultat från diabetesdelprojektet (Delprojekt 3 i Övrig-gruppen)**

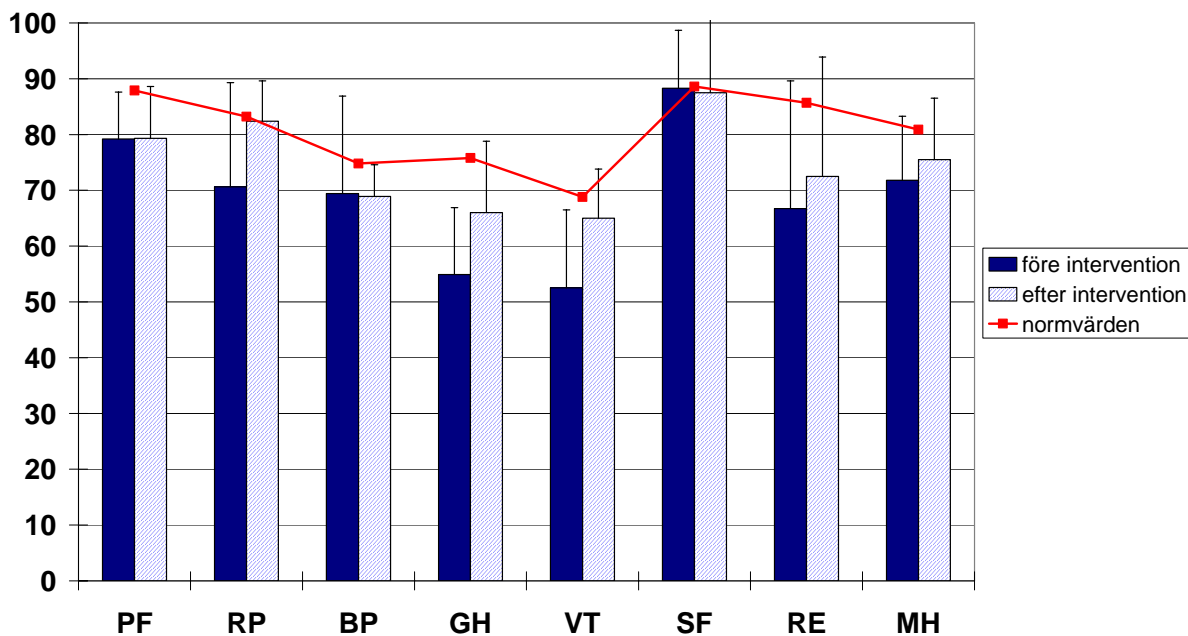
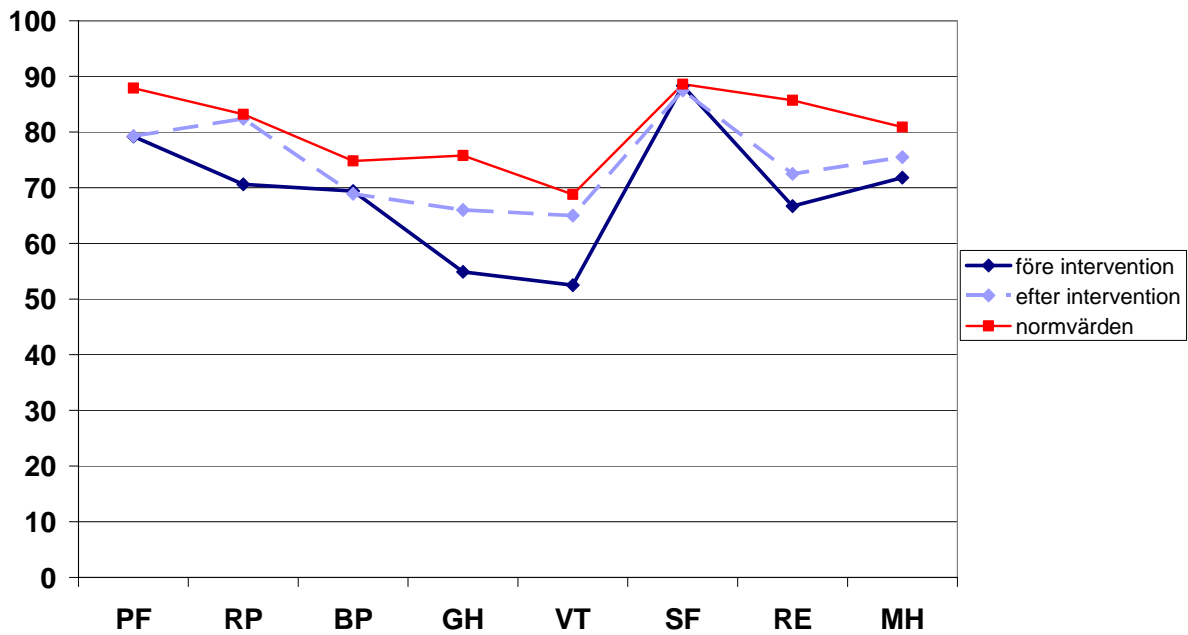
Diabetespatienterna som ordinerats FaR används här som exempel på de olika mått och presentationssätt som de enskilda sjukhusen/delprojekten fått sig tillsända (figur 1-7). Tex kan man för SF-36 jämföra 0-100 presentationsmodellen (figur 1) med NBS-modellen (figur 2), där den senare underlättar inbördes jämförelser, men samtidigt har snävare variationsvidd (0-70), vilket man måste beakta vid tolkningen, speciellt om man är van vid 0-100 modellen. Man kan även jämföra presentationssätten med linje- och stapeldiagram. Tanken bakom att använda linjediagram är att man vill skapa en grafisk bild, en HRQoL-profil, medan de som föredrar stapeldiagram hävdar att linjediagrammen snarast är förvirrande, eftersom sådana är tänkta att avspegla ett förlopp (oftast över tid), vilket det ju inte är frågan om här.

SF-36 uppvisade signifikanta förbättringar för allmän hälsa GH ( $p=0,002$ ) och vitalitet VT ( $p=0,048$ ), medan inga signifikanta skillnader sågs för vare sig EQ-5D index (figur 3), VAS (figur 4) eller de separata dimensionerna inom EQ-5D instrumentet, även om vissa trender kunde anas.

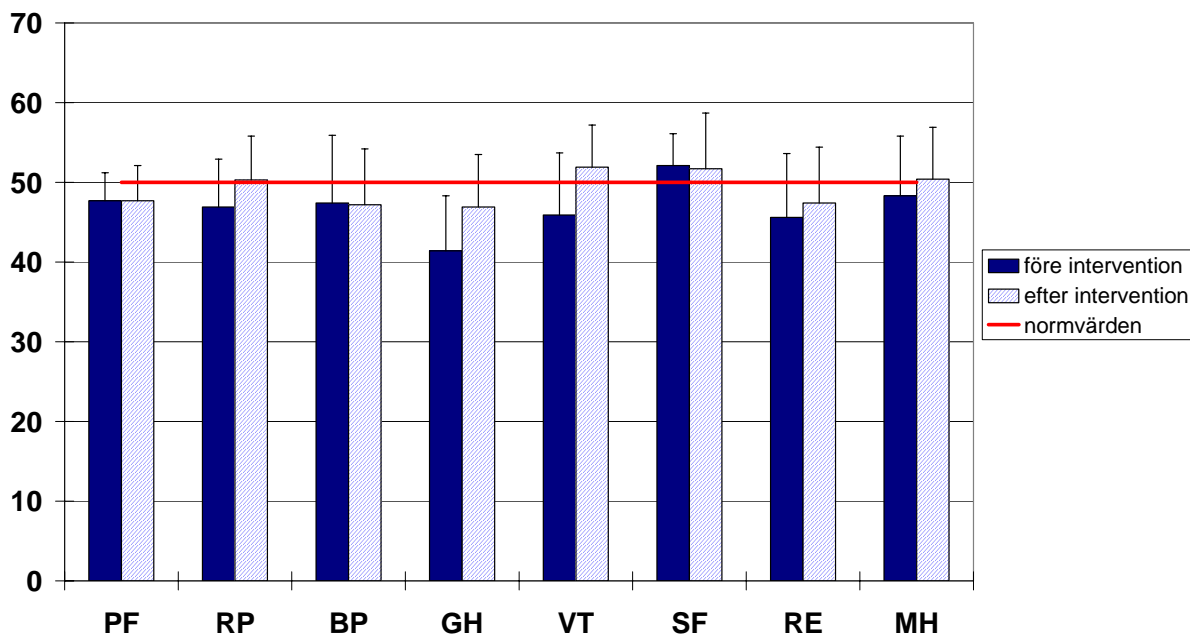
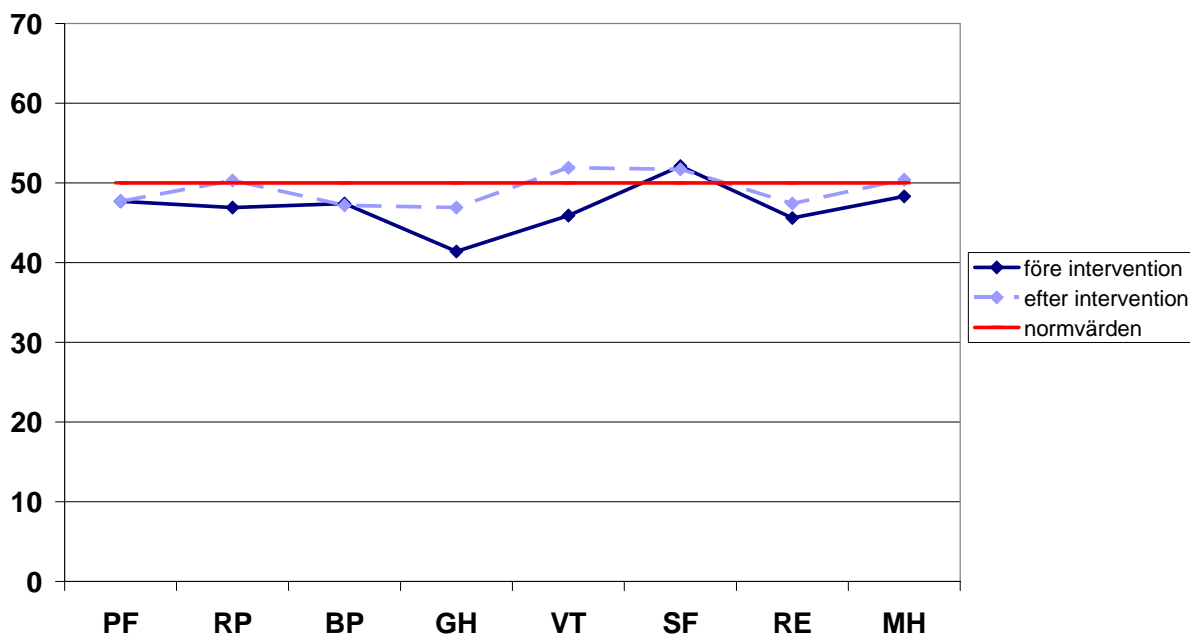
Det faktum att man kan se signifikanta förbättringar trots att alla konfidensintervall överlappar varandra beror på att vi använt parade signifikanstester, eftersom mätningarna är beroende (se Statistiska metoder). Parade tester ger en skarpare analys, eftersom de jämför skillnader i före- och eftermätningarna utan att behöva ta hänsyn till (spridnings)skillnader mellan individerna (något som däremot avspeglas i konfidensintervallen).

Den procentuella fördelningen av svaren (inga, måttliga eller svåra problem) inom var och en av dimensionerna i EQ-5D (figur 6 och 7), liksom medelvärdena för dessa (figur 5, se vidare under Metoddiskussion), ger information om i vilken/vilka av de fem dimensionerna problemen föreligger, här i *smärtor/besvär* och *oro/nedstämdhet*. Jämför man med SF-36 så låg där problemtyngden i GH, VT och RE (emotionell rollfunktion) skalorna. Både för EQ-5D dimensionerna och för SF-36 var det för de dimensioner/skalor med initialt sämst utgångsläge som de största förbättringarna/förbättringstendenserna sågs.

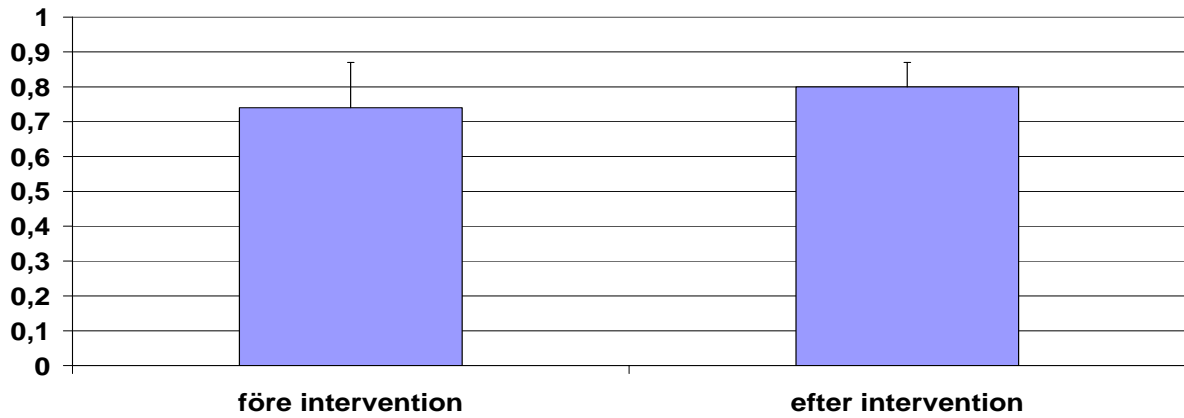
När man jämför resultaten för SF-36 med resultaten för EQ-5D dimensionerna måste man hålla i minnet att man därmed jämför graderade skalor med frekvenssiffror, vilket naturligtvis inte är helt rättvisande.



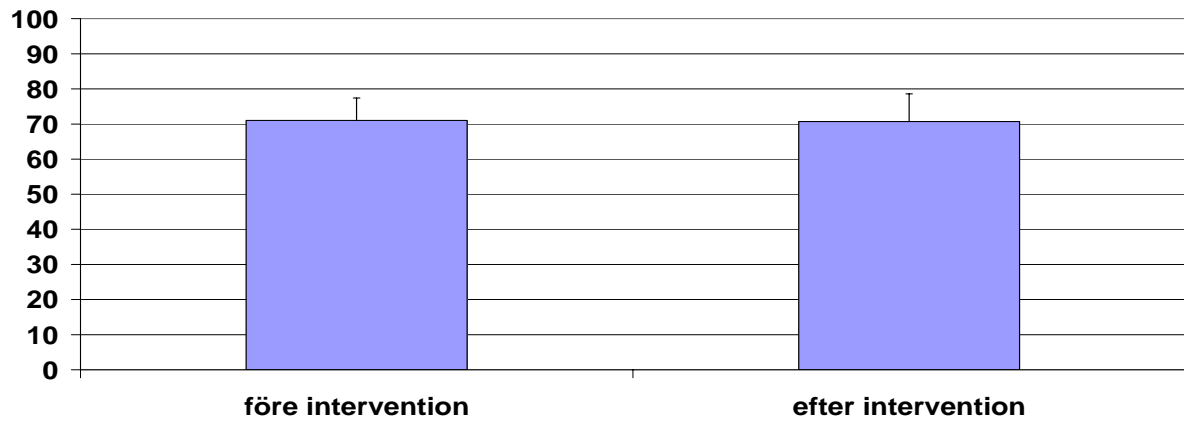
**Figur 1a och b.** Diabetiker, före och efter FaR (n=17). SF-36 (0-100 modellen), som linje- och stapeldiagram. 95% konfidensintervall för normvärdena är alla <1 och syns därför ej i figuren. 95% konfidensintervall för mätmedelvärdena presenteras endast i stapeldiagrammet.



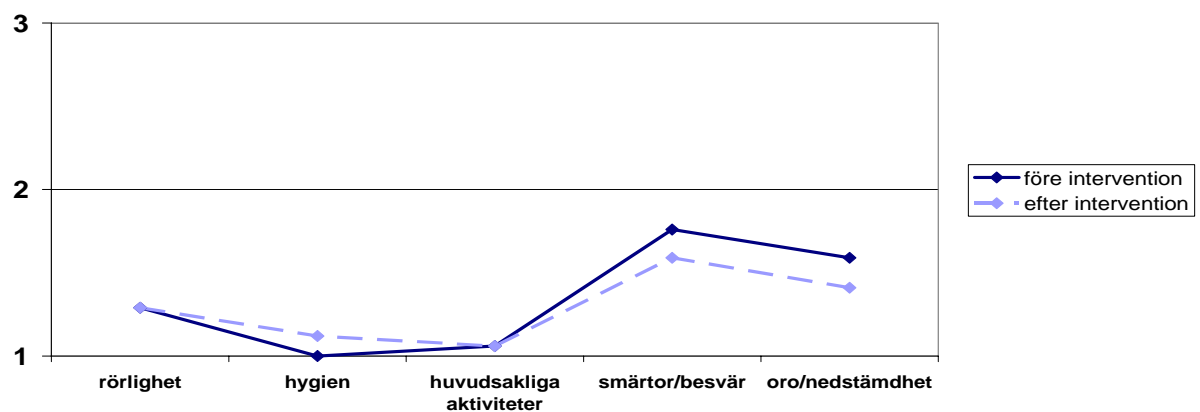
**Figur 2a och b.** Diabetiker, före och efter FaR (n=17). SF-36 NBS-modellen (Norm-Based Scoring), som linje- och stapeldiagram. Normvärdet är alltid 50 och standardavvikelsen 10 för alla skalor. 95% konfidensintervall för mätmedelvärdena presenteras endast i stapeldiagrammet. Observera att y-axeln i dessa diagram endast går till 70, vilket är taket för NBS-skalan.



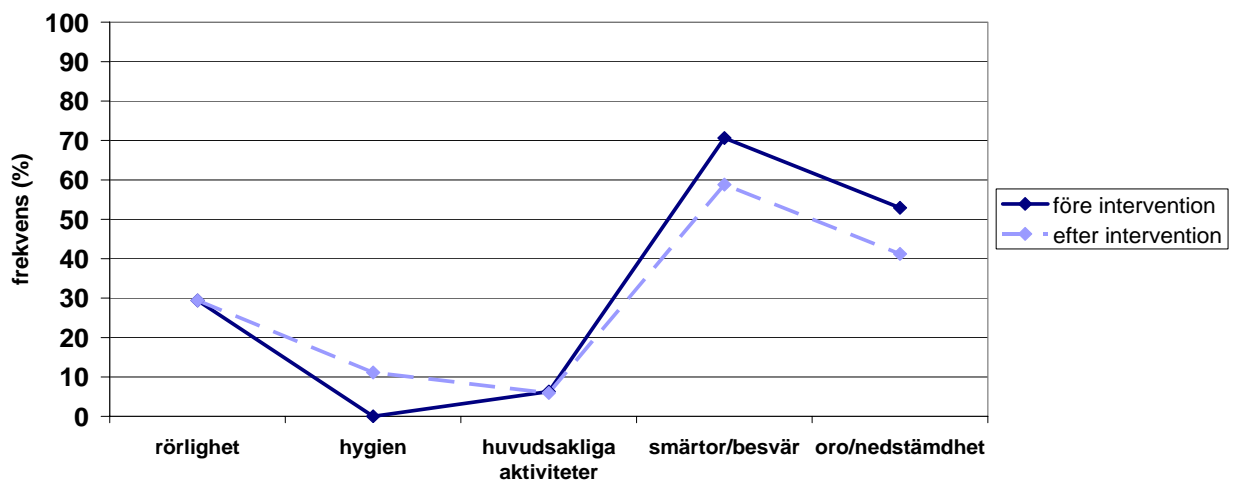
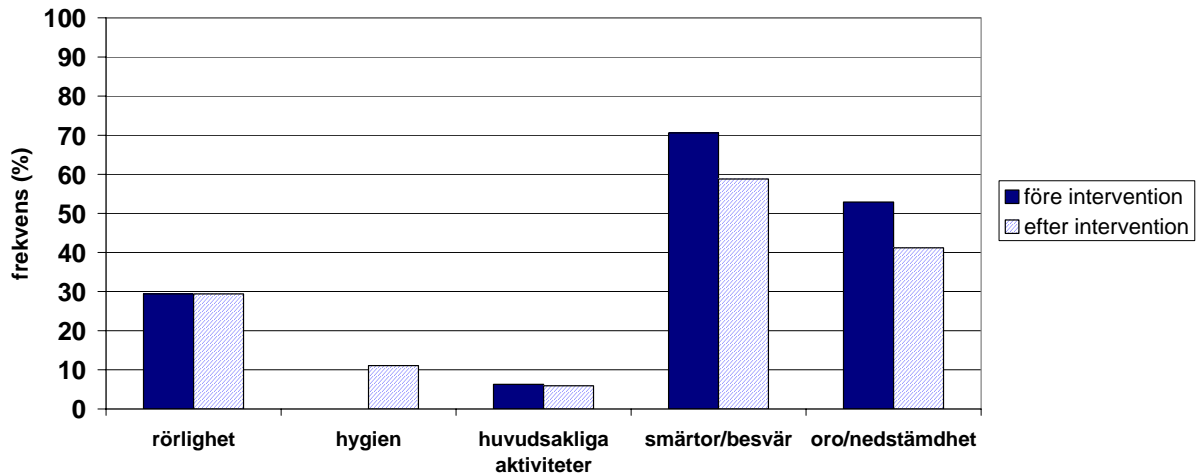
**Figur 3.** Diabetiker, före och efter FaR (n=17). EQ-5D index samt 95% konfidensintervall för medelvärdena.



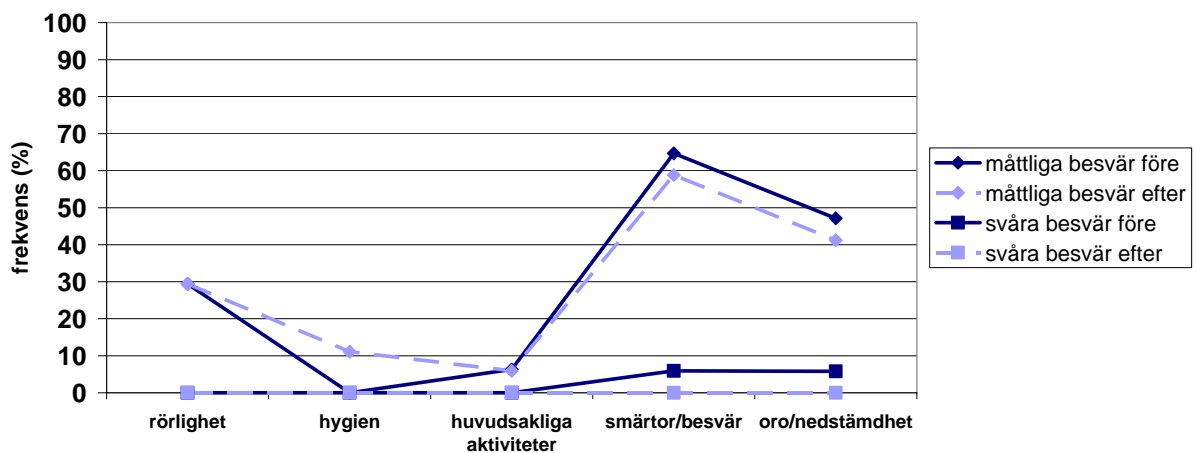
**Figur 4.** Diabetiker, före och efter FaR (n=17). EQ-5D VAS samt 95% konfidensintervall för medelvärdena.



**Figur 5.** Diabetiker, före och efter FaR (n=17). EQ-5D dimensionerna (medelvärden), som linjediagram.



**Figur 6a och b.** Diabetiker, före och efter FaR (n=17). EQ-5D dimensionerna. Andel som anger problem (måttliga eller svåra) inom varje dimension, som stapel- och linjediagram.



**Figur 7.** Diabetiker, före och efter FaR (n=17). EQ-5D dimensionerna. Andel som anger problem inom varje dimension, uppdelat på måttliga och svåra problem, före och efter intervention.



## **Resultat från övriga delprojekt**

### *Generellt*

På delprojektsnivå är flera grupper väldigt små och därför har vi slagit samman de olika delprojekten inom varje sjukdomsgrupp för de statistiska beräkningarna. Vi vill påpeka att man måste vara försiktig vid tolkningen av dessa sammanslagna värden och medveten om de eventuella problem som redovisas i avsnittet Statistiska metoder. Enbart deltagare som genomfört både före- och eftermätningarna ingick i de statistiska analyserna.

HRQoL-bilden kunde skilja betydligt mellan de olika sjukdomsgrupperna (se vidare nedan). Även mellan de olika delprojekten inom en sjukdomsgrupp kunde man ana skillnader, och också mellan de båda instrumenten, men det fanns även likheter. För EQ-5D kunde man se att det var den fjärde dimensionen, *smärtor/besvär*, som dominerade bilden inom varje sjukdomsgrupp, utom möjligen för KOL-patienterna. När det gäller SF-36 hade alla sjukdomsgrupper överlag lägst värden för RP-skalan (som mäter hur kroppsliga besvär inverkar på förmågan att utföra arbete eller andra regelbundna aktiviteter), särskilt vid föremätningen.

### *Hjärtinfarktgruppen*

Man kan se att hjärtinfarktgruppen (figur 8) var väldigt homogen. Både vid före- och eftermätningar låg de olika delprojekten mycket väl samlade, och de fick alla en förbättring av HRQoL efter interventionen, trots att de redan innan interventionen skattade sin HRQoL högst bland de olika sjukdomsgrupperna. Gällande SF-36 uppvisade dock RP, åtminstone initialt, mycket låga värden jämfört med de övriga skalorna. Statistiskt sågs för den sammanslagna gruppen signifikanta förbättringar för alla skalor i SF-36 ( $p < 0,001$ ). Även EQ-5D index och VAS visade signifikanta förbättringar ( $p < 0,001$ ) medan det bland EQ-5D dimensionerna bara var *huvudsakliga aktiviteter*, *smärtor/besvär* och *oro/nedstämdhet* som uppvisade signifikans ( $p < 0,01$ ). Å andra sidan var de två övriga dimensionerna relativt problemfria (=uppvisade låga värden) redan från början.

Förutom hjärtinfarkt fanns ytterligare två hjärtsjukdomsgrupper, nämligen en arytmi-grupp (övrigt delprojekt 4) och en hjärtsviktsgrupp (övrigt delprojekt 5). En tydlig skillnad mot infarktpatienterna var att dessa båda patientgrupper initialt inte låg lika lågt på RP-skalan. Inga signifikanta skillnader mellan före- och eftermätningen sågs för något av instrumenten för arytmi-gruppen. Detta gäller även för den lilla gruppen hjärtsviktspatienter, där man emellertid kunde se tydliga förbättringstrender (bilaga 3).

I bilaga 3 visas även en jämförelse mellan hjärtinfarktdelprojekt 6, 7 och 8, samt arytmi-gruppen, eftersom dessa fyra är patientgrupper vid samma sjukhus hjärtklinik. Skillnaden mellan delprojekt 6 och de två övriga hjärtinfarktgrupperna var att de senare mäter före och efter ett ingrepp, medan delprojekt 6 enbart mäter efter ingreppet, men före och efter klinikens hjärtskola.

### *Smärtgruppen*

Till skillnad mot den enhetliga hjärtinfarktgruppen var smärtgruppen (figur 9) mycket mer inhomogen. Smärtpatienterna skattade sin HRQoL lägst bland de olika sjukdomsgrupperna, och varierade mycket mer gällande förbättringar av interventionen, tex fanns även tendenser till försämringar för vissa delprojekt. Inga signifikanta skillnader sågs överhuvudtaget för den sammanslagna gruppen, vare sig för SF-36 eller EQ-5D.

### *KOL-gruppen*

Inom KOL-gruppen (figur 10) kunde man, precis som för smärtgruppen, se tendenser till förbättringar för vissa delprojekt och försämringar för andra. Den sammanslagna KOL-gruppen uppvisade signifikanta förbättringar för GH ( $p=0,018$ ) och för EQ-5D dimensionen *rörlighet* ( $p=0,033$ ).

### *Reumagruppen*

Även inom reumagruppen (figur 11) fanns variation angående eventuella förbättringar efter intervention för de olika delprojekten. För den sammanslagna reumagruppen sågs signifikanta förbättringar för fyra skalor i SF-36, PF ( $p=0,029$ ), VT ( $p=0,002$ ) samt RP och BP ( $p<0,001$ ). EQ-5D uppvisade en signifikant förbättring för VAS ( $p<0,001$ ) men ej för indexet. Bland EQ-5D dimensionerna sågs signifikanta förbättringar för *hygien* ( $p=0,018$ ), *huvudsakliga aktiviteter* ( $p=0,006$ ) och *oro/nedstämdhet* ( $p=0,019$ ). Om man tar bort delprojekt 3 (som utmärker sig med stora förbättringar) ur mätningarna, blir inte längre PF signifikant, och bland EQ-5D dimensionerna kvarstår endast signifikans för *huvudsakliga aktiviteter*.

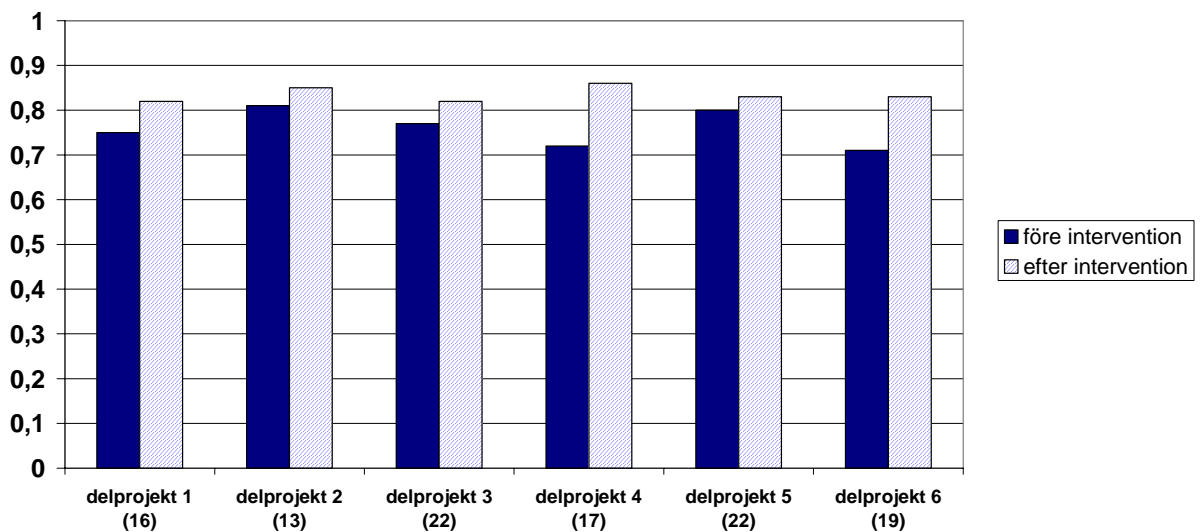
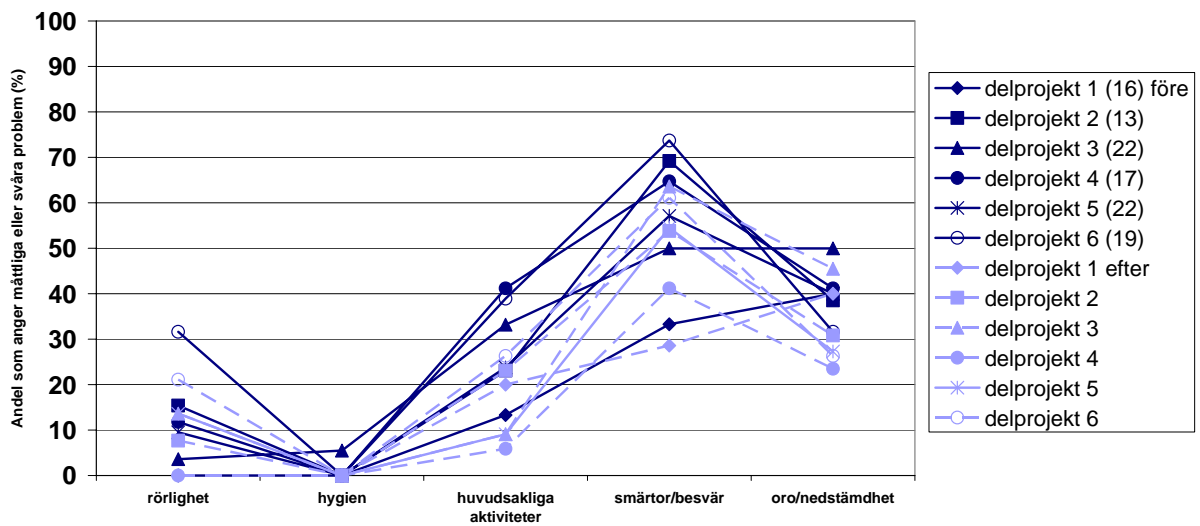
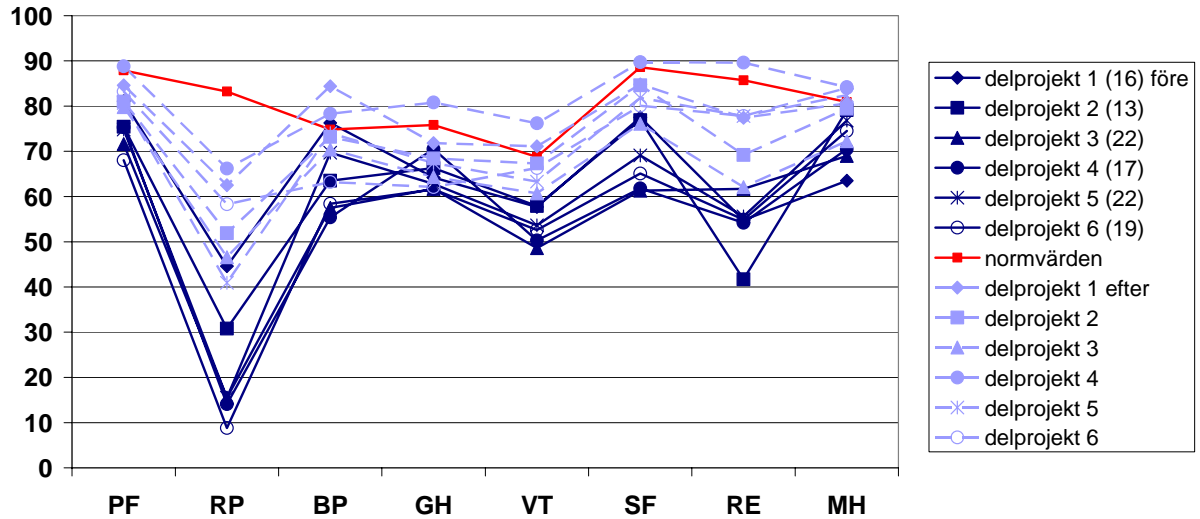
### *Knä/höftkirurgigruppen*

Knä/höftkirurgigruppen (figur 12) kan delas i två, där de heldragna och streckade linjerna representerar planerad knä/höftplastik (samt i ett fall reparation av skadade korsband i knäet), medan pricklinjerna representerar akut operation efter höftfraktur. I det senare fallet innebär föremätningen att patienten, vid ankomst till sjukhuset (dvs innan operation), skattat sin HRQoL som den var före skadan. För den sammanslagna gruppen planerade operationer (delprojekt 1, 2, 4)\* sågs signifikanta förbättringar för SF-36 för PF och BP ( $p<0,001$ ) samt RE ( $p=0,026$ ). För EQ-5D uppvisade index ( $p=0,011$ ) och VAS ( $p=0,002$ ) signifikanta skillnader, likaså dimensionerna *rörlighet* ( $p=0,012$ ) och *smärtor/besvär* ( $p<0,001$ ). Däremot fann man inga skillnader alls för den mindre gruppen akuta operationer (delprojekt 5), där föremätningen alltså var en skattning av patientens HRQoL innan skada. Knä/höftkirurgigruppen var den grupp, vars delprojekt uppvisade störst förbättringar för EQ-5D index.

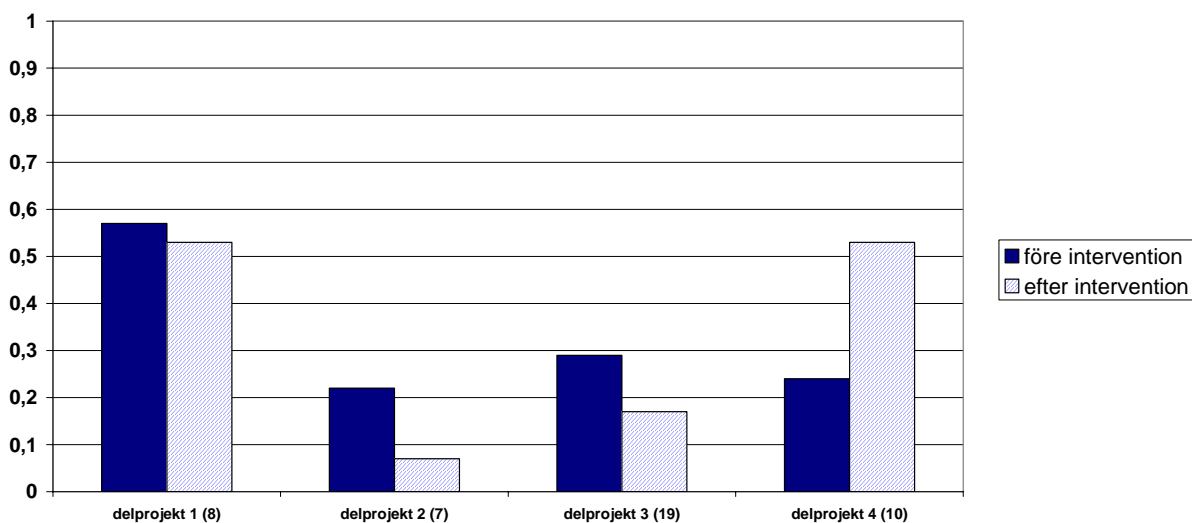
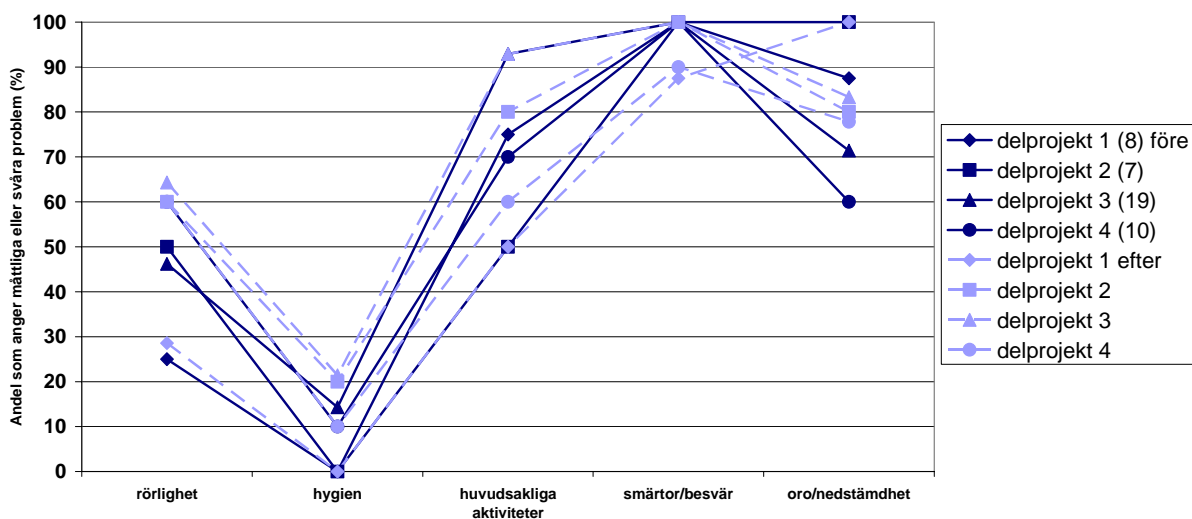
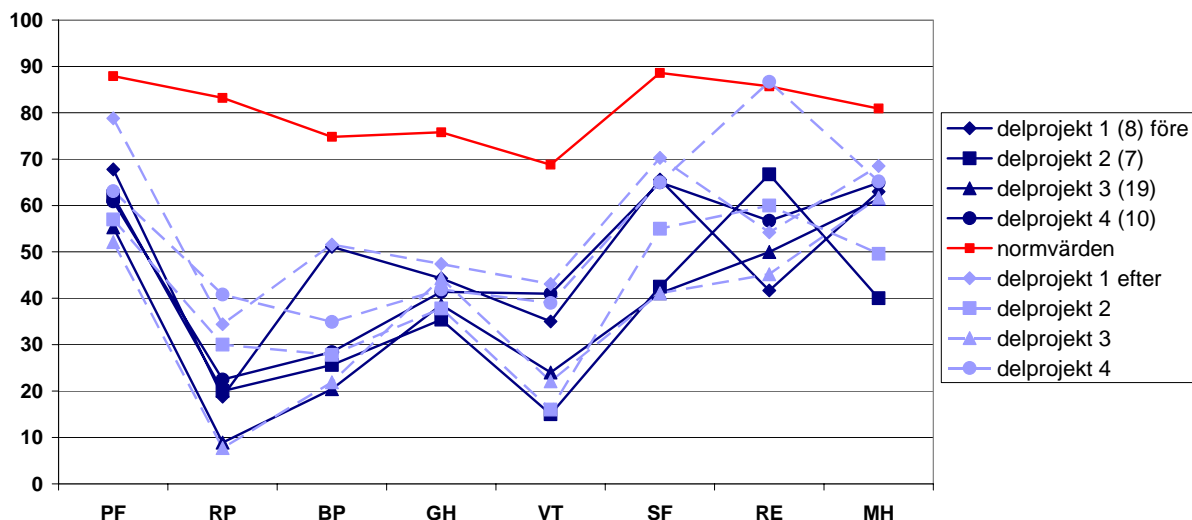
\*Delprojekt 3 hade problem med id-märkningen av sina enkäter och kunde därför inte ingå i parade signifikanstester. Detta medför också att i figur 12 visas resultaten för alla ingående deltagare både vid före- och eftermätningen, vilket var 15 respektive 13. Oparad signifikanstestning visar dock signifikanta förbättringar för skalorna BP, VT och SF i SF-36.

### *Delprojekt medarbetargrupper*

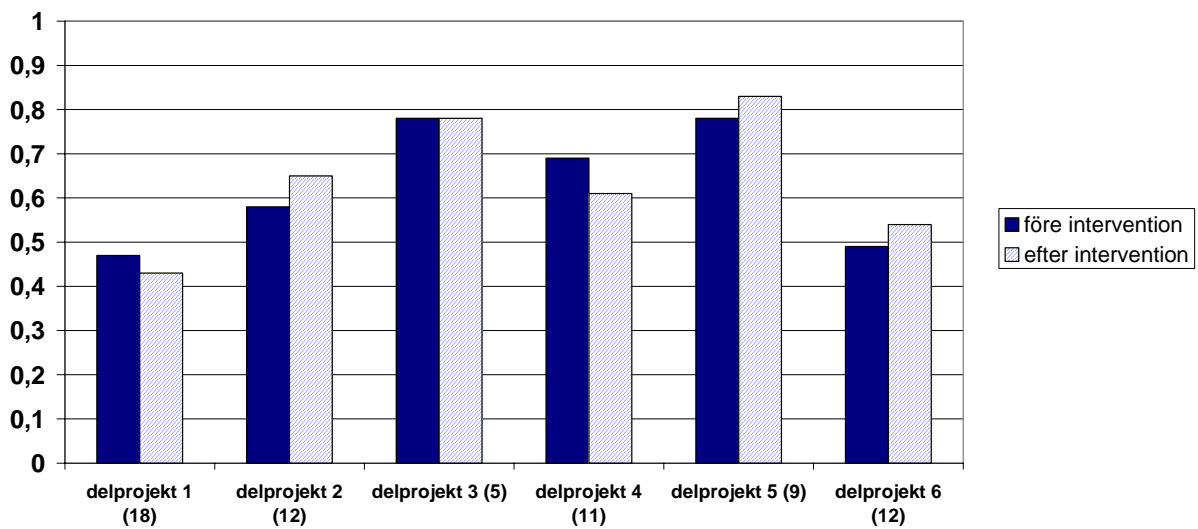
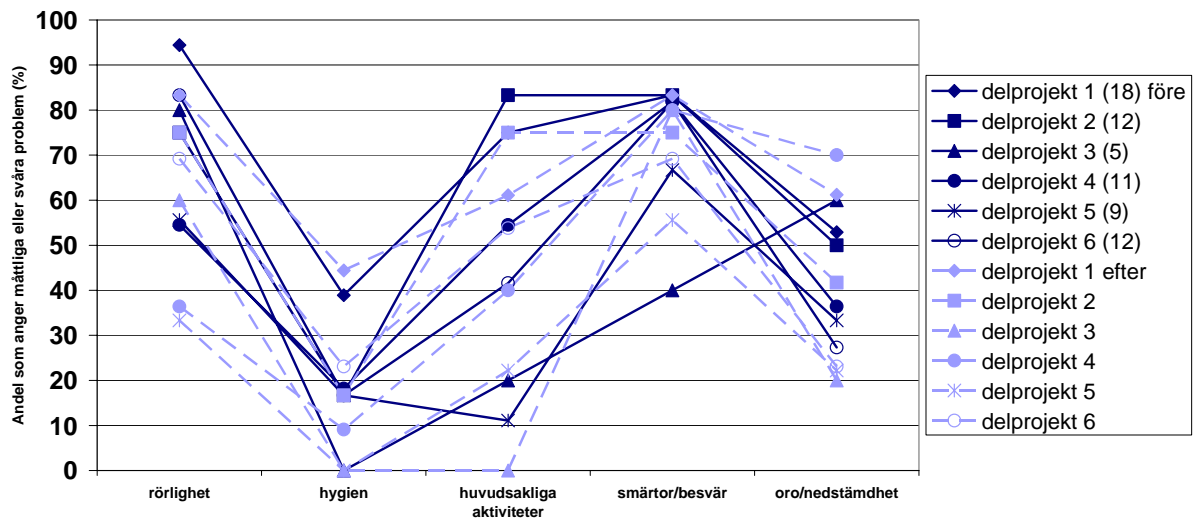
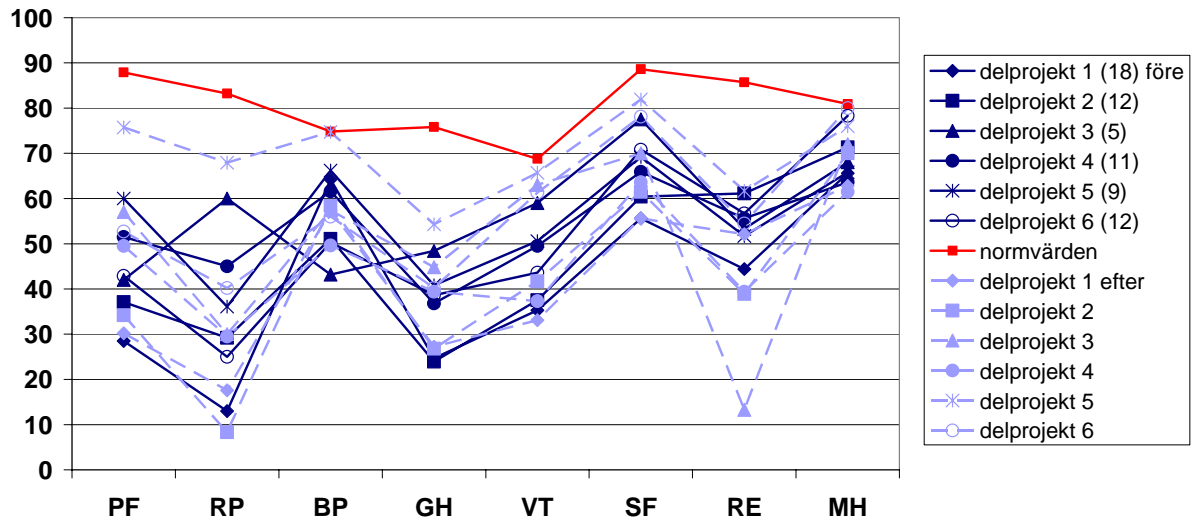
Förutom de övriga patientdelprojekt som redan redovisats ovan genomfördes också två delprojekt bland sjukhusens medarbetare, en viktskola (övrigt delprojekt 1) och ett arbetslivsprojekt (övrigt delprojekt 2). Man kan tydligt se att deltagarna i dessa delprojekt ligger runt, och i vissa fall även över, normvärdena (den sk *Healthy worker-effekten*). Inga signifikanta skillnader sågs mellan före- och eftermätningar, men viktskolan hade väldigt få deltagare och arbetslivsprojektgruppens enkäter saknade individuell id-märkning, varför parade tester ej kunde utföras för denna grupp. Oparad signifikanstestning för delprojekt 2 visar en signifikant försämring för EQ-5D dimensionen *huvudsakliga aktiviteter*. Se bilaga 3 för diagram.



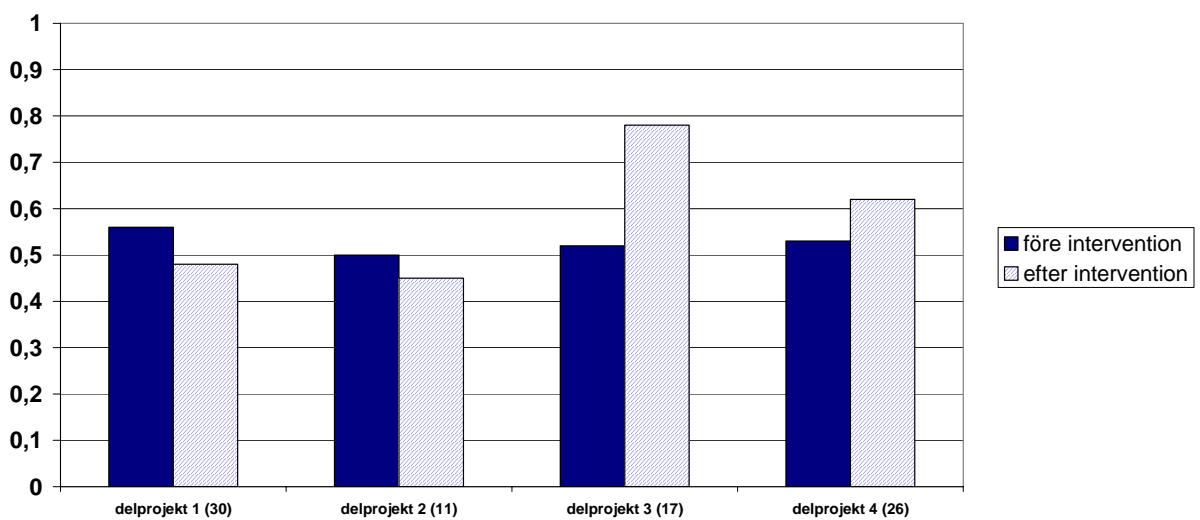
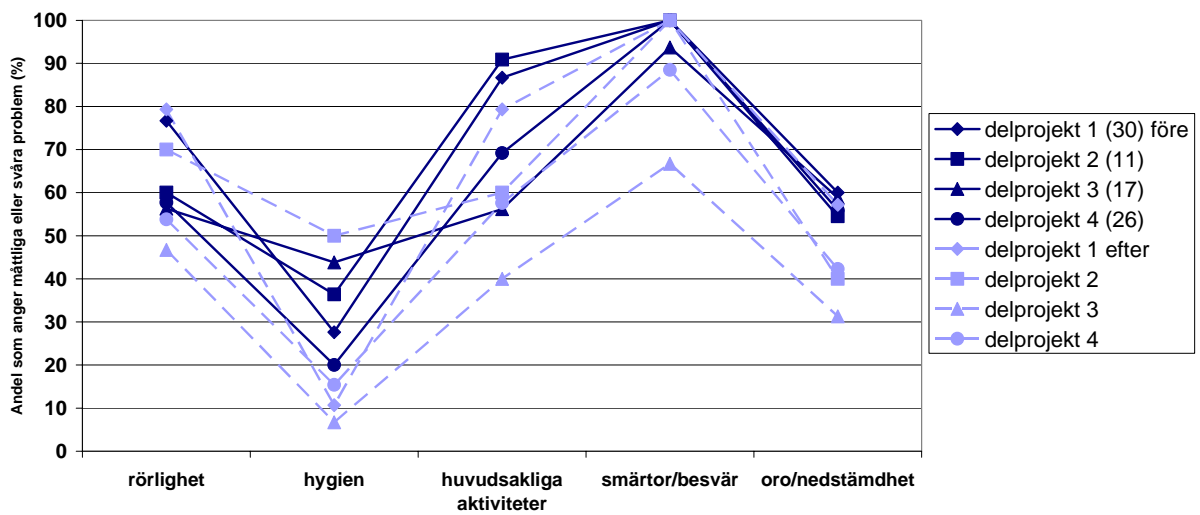
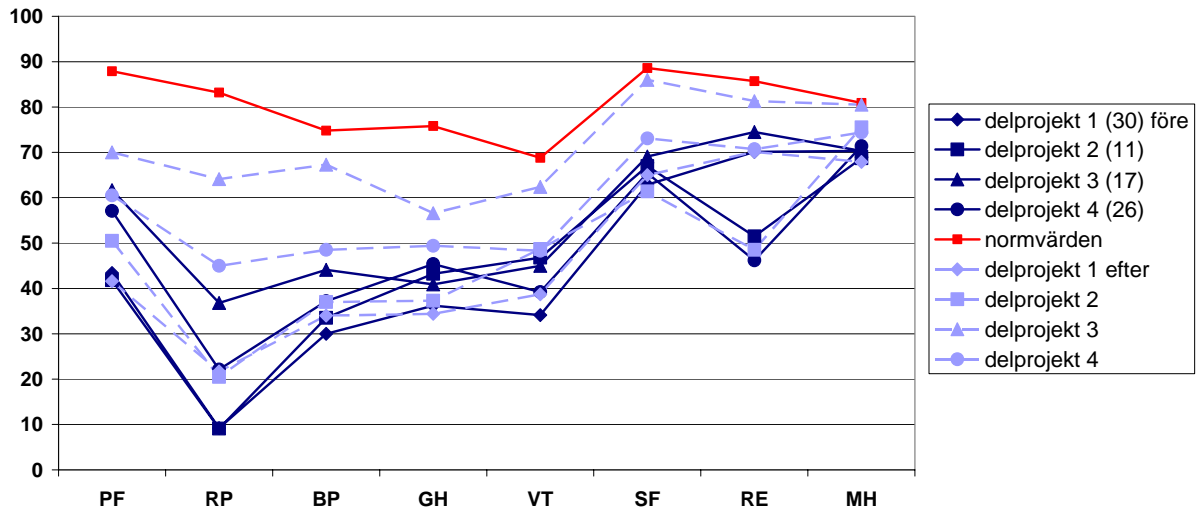
**Figur 8a, b och c.** Hjärtinfarktdelprojekten före och efter intervention: SF-36 (överst), EQ-5D dimensioner (mitten) och EQ-5D index (nederst). Siffror inom parentes anger deltagarantal.



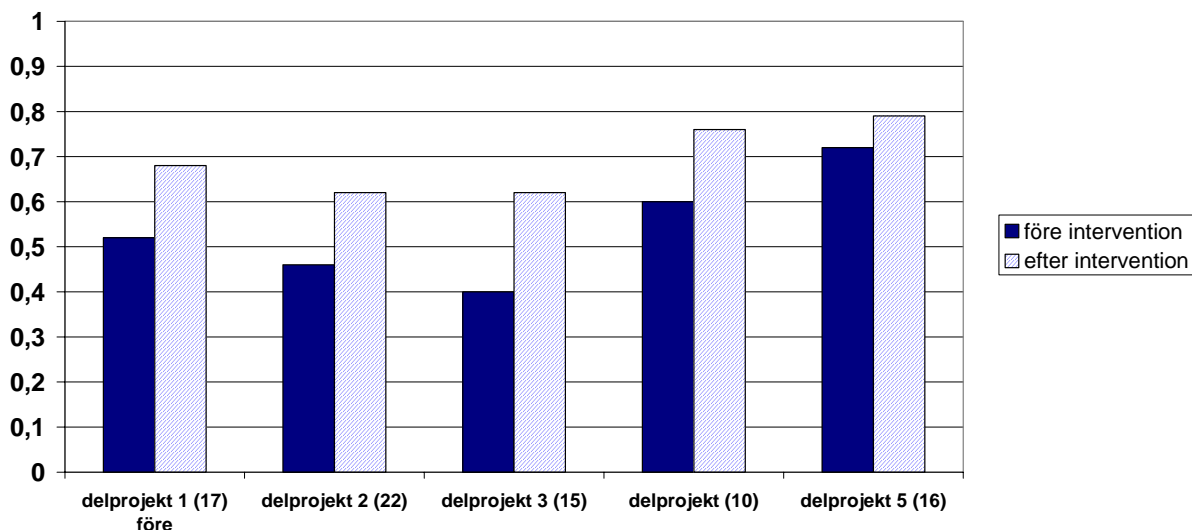
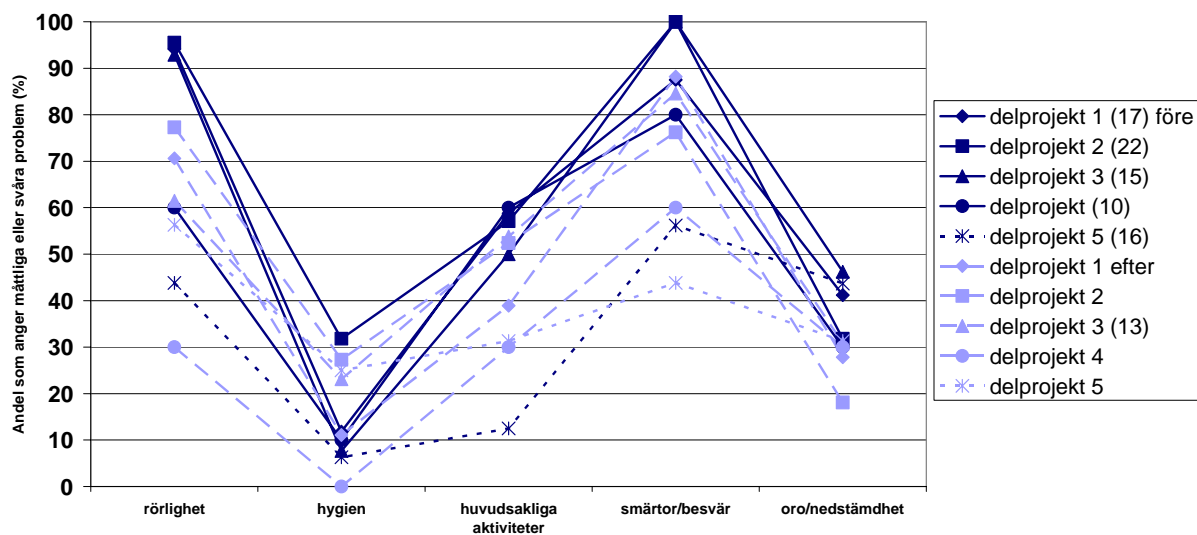
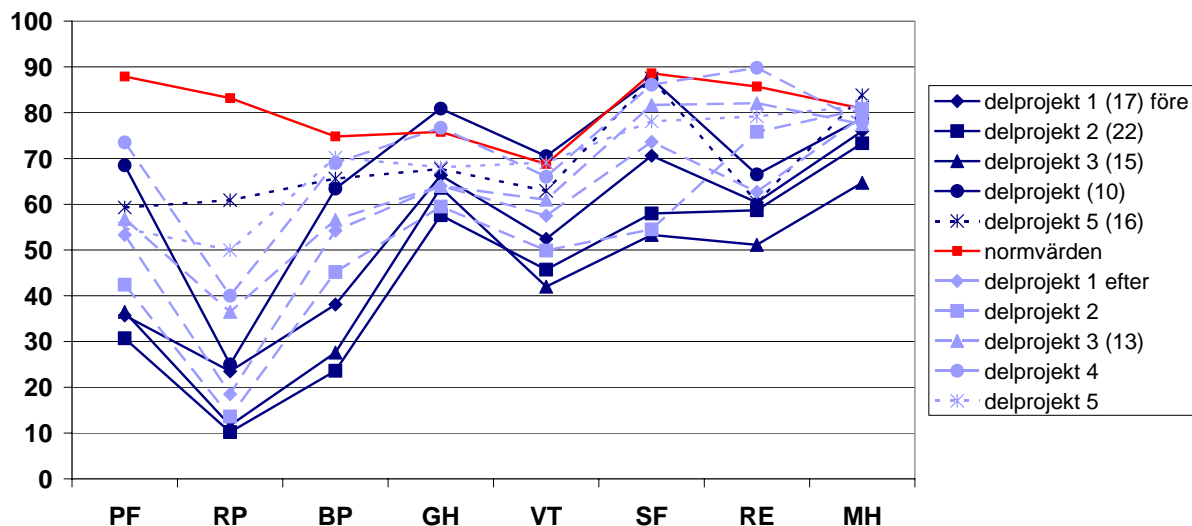
Figur 9a, b och c. Smärtdelprojekten före och efter intervention: SF-36 (överst), EQ-5D dimensioner (mitten) och EQ-5D index (nederst). Siffror inom parentes anger deltagarantal.



Figur 10a, b och c. KOL-delprojekten före och efter intervention: SF-36 (överst), EQ-5D dimensioner (mitten) och EQ-5D index (nederst). Siffror inom parentes anger deltagarantal.



Figur 11a, b och c. Reumadelprojekten före och efter intervention: SF-36 (överst), EQ-5D dimensioner (mitten) och EQ-5D index (nederst). Siffror inom parentes anger deltagarantal.

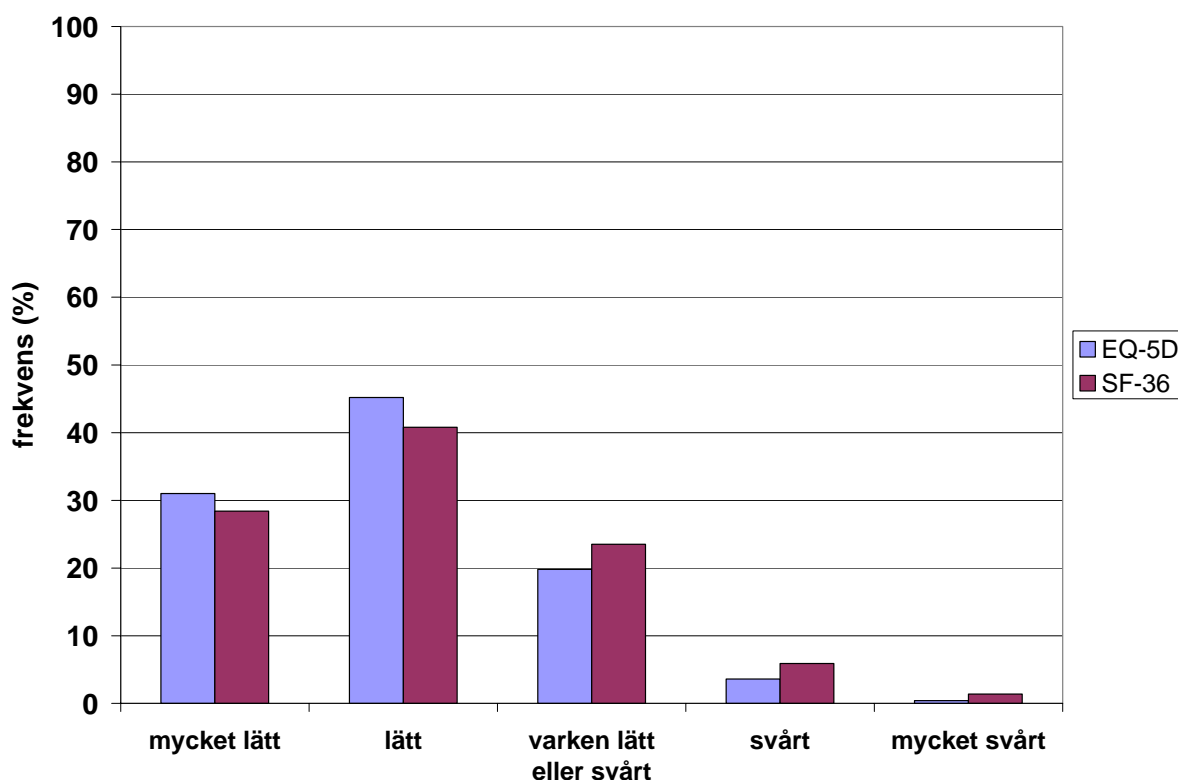


**Figur 12a, b och c.** Knä/höftkirurgidelprojekten före och efter intervention: SF-36 (överst), EQ-5D dimensioner (mitten) och EQ-5D index (nederst). Delprojekt 1-4 visar planerade operationer, medan delprojekt 5 visar akuta operationer, där före-värdet anger självskattad HRQoL innan skada. Siffror inom parentes anger deltagarantal. (Se texten gällande delprojekt 3)

### Resultat från utvärderingsenkäten

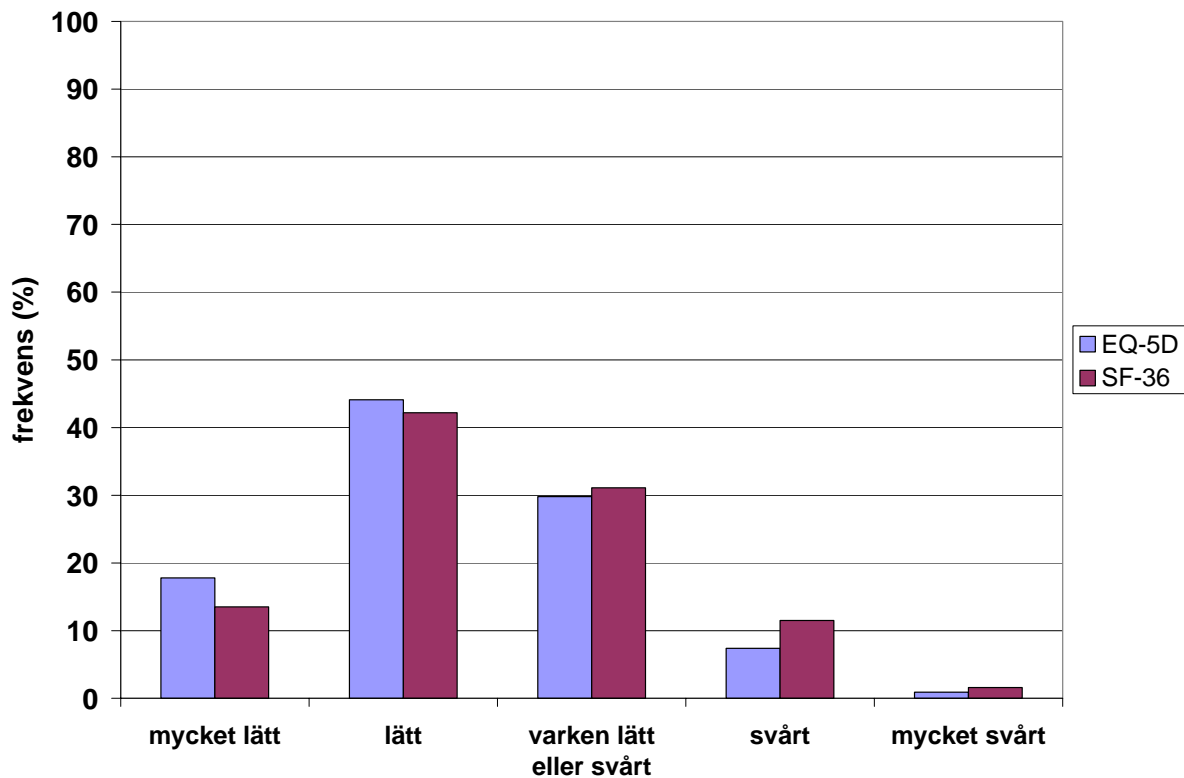
505 deltagare har genomfört andra mättillfället och därmed fått utvärderingsenkäten. Majoriteten av dessa har valt att svara på frågorna (något fler har valt att svara på frågorna om SF-36 än om EQ-5D). Huvudfyndet var att över hälften av patienterna anser att hälsovinstmätning är värdefullt och valet av instrument verkade inte vara avgörande. De flesta tyckte också att instrumenten var lätta att använda, och gav god möjlighet till att få beskriva ens hälsotillstånd. EQ-5D upplevdes vara något lättare, både att förstå och att välja svar från, än SF-36, medan det däremot inte sågs någon signifikant skillnad mellan instrumenten när det gällde hur de ger möjlighet till att få beskriva sitt hälsotillstånd (figur 13-15). Flertalet visade sig dock ha valt samma svarsalternativ för båda instrumenten, dvs tyckte man tex att det ena instrumentet var lätt att svara på, så tyckte man oftast även att det andra var lätt att svara på.

För båda instrumenten gällde att män tyckte att de var lättare (oftare använt svarsalternativen mycket lätt och lätt) att förstå och svara på än vad kvinnor tyckte, och för SF-36, att yngre tyckte det var lättare än vad äldre tyckte. Män var klart nöjdare (oftare använt svarsalternativen mycket bra och bra) med hur båda instrumenten gav möjligheter till att få beskriva sitt hälsotillstånd än vad kvinnorna var, medan, för SF-36, en tendens till att äldre var något mer missnöjda (oftare använt svarsalternativen ofullständigt och mycket ofullständigt) än yngre kunde anas. Jämför man de olika sjukdomsgrupperna, så var hjärtpatienterna den grupp som tyckte det var lättast att förstå frågorna och välja svar-salternativ, gällande både SF-36 och EQ-5D, förutom för EQ-5D när det gällde att välja svarsalternativ (där sågs inga skillnader mellan sjukdomsgrupperna). Mest nöjd med hur instrumenten gav möjlighet till att få beskriva sitt hälsotillstånd var diabetesgruppen (gäller båda instrumenten). Inte heller för utvärderingen har vi kunnat använda våra komorbiditetsdata för subgruppsanalyser i detta skede.

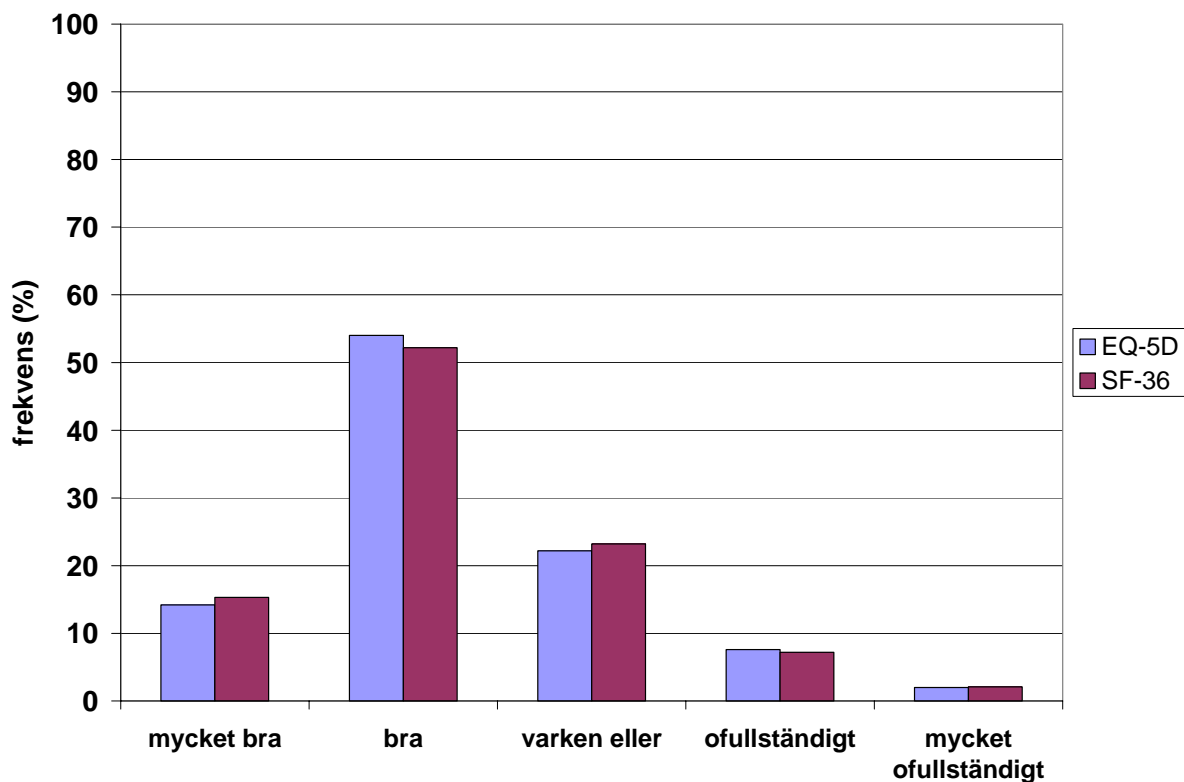


**Figur 13.** Fördelning bland dem som besvarat frågan "Var det lätt eller svårt att förstå frågorna?"; EQ-5D (n=449) jämfört med SF-36 (n=490).



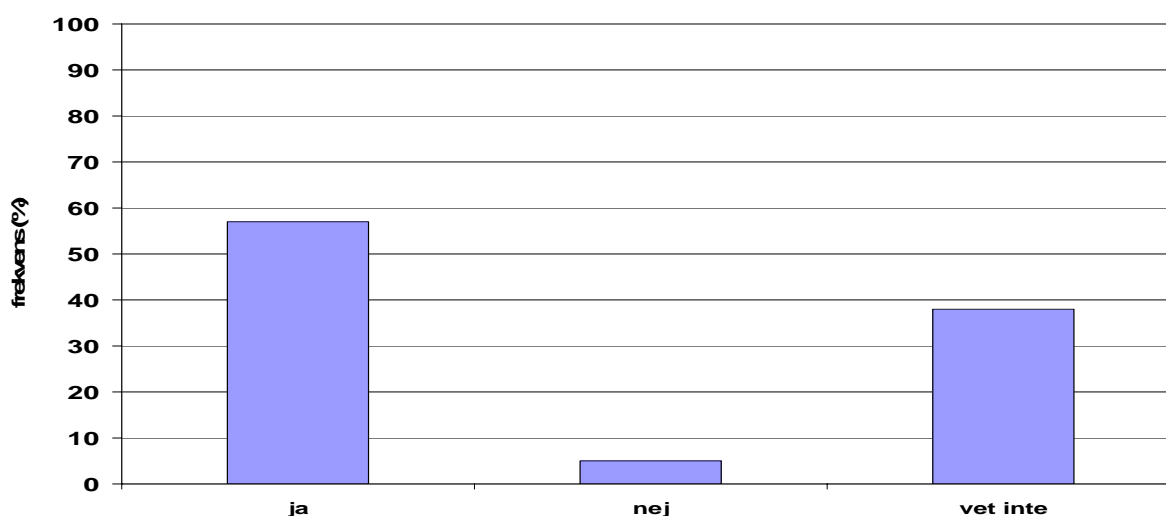


**Figur 14.** Fördelning bland dem som besvarat frågan "Var det lätt eller svårt att välja svarsalternativ?"; EQ-5D (n=449) jämfört med SF-36 (n=488).

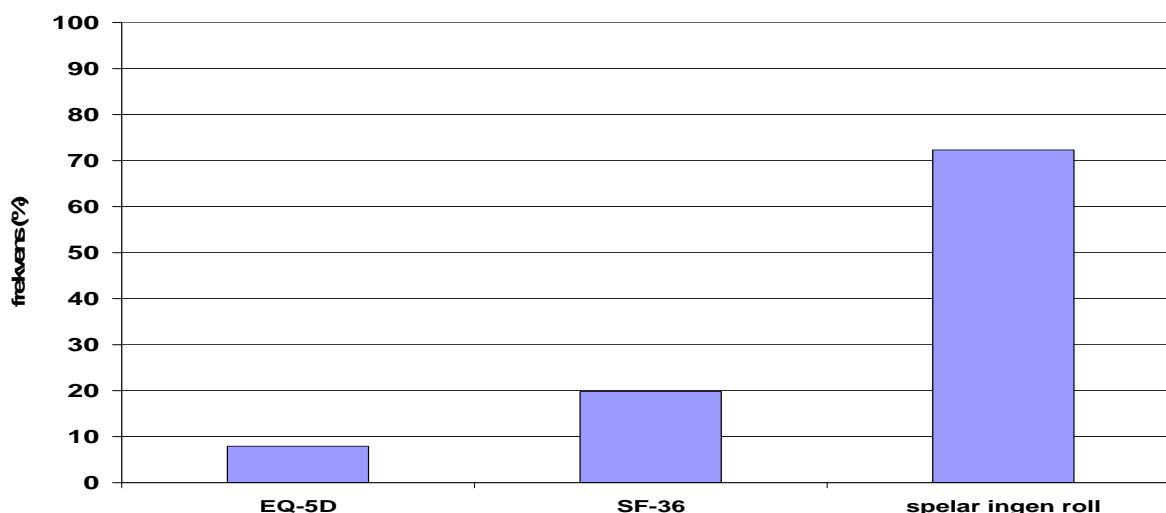


**Figur 15.** Fördelning bland dem som besvarat frågan "Gav enkäten dig möjlighet att beskriva ditt hälsotillstånd på ett fullständigt sätt?"; EQ-5D (n=446) jämfört med SF-36 (n=487).

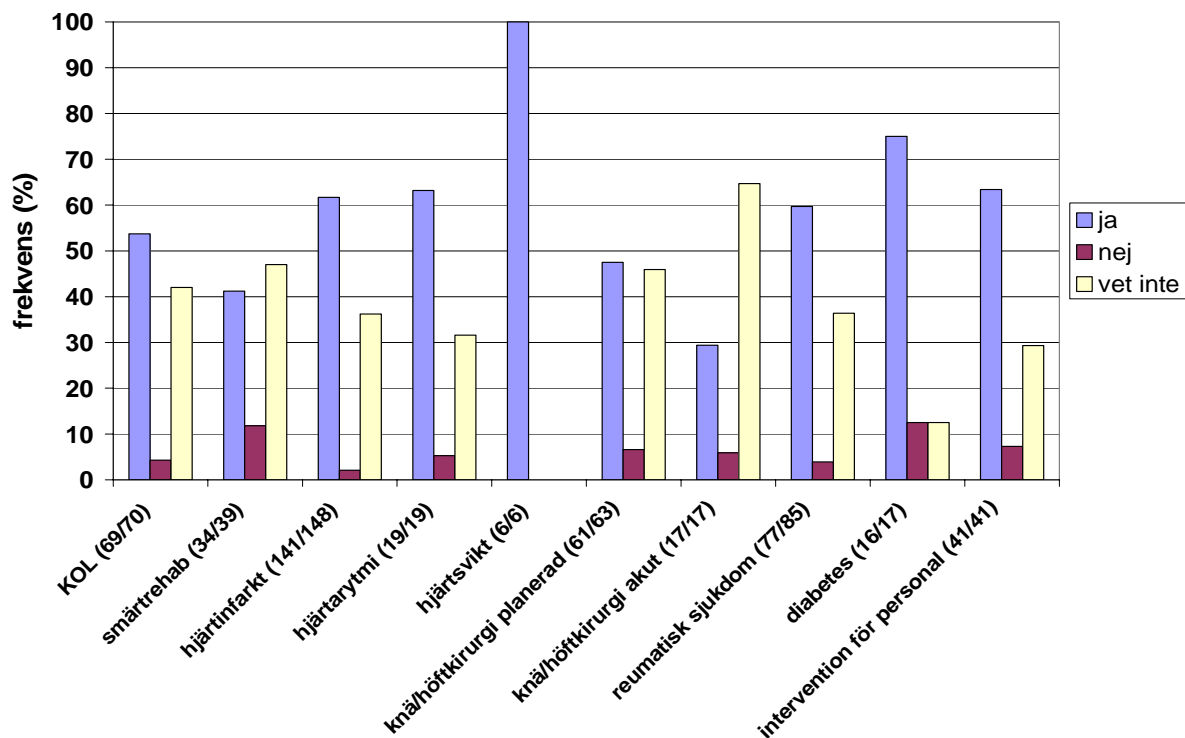
Lite drygt hälften svarade *ja* på frågan om de ser något värde i hälsovinstmätningar. Endast 5% svarade *nej*. Fler män än kvinnor tar ställning, och de svarar ja i större utsträckning. Likaså är det fler yngre än äldre som tar ställning, och även de svarar ja i större utsträckning. Bland sjukdomsgrupperna var hjärtinfarkt- och reumagrupperna mest positiva, fast mest positiva av alla var hjärtsvikts- och diabetesgrupperna (obs små grupper). Tre fjärdedelar har inte velat ta ställning till om de föredrar något av de båda instrumenten, men av dem som har framfört en åsikt är det dubbelt så många som föredrar SF-36 framför EQ-5D. Män och kvinnor har likartade siffror, medan yngre tenderar att ta ställning något oftare än äldre, och de väljer då SF-36. Mest positiv till EQ-5D var gruppen akuta operationer, som dock var liten och hade låg svarsfrekvens på denna fråga. Bland de fem utvalda sjukdomsgrupperna var smärtgruppen mest positiv till EQ-5D, men lika många i den gruppen var även positiva till SF-36. Smärtgruppen var också den av de fem sjukdomsgrupperna där flest valde att ta ställning till instrumenten. I figurerna nedan redovisas resultatet från totalmaterialet för andra mättillfället samt uppdelat på sjukdomsgrupperna (figur 16-19).



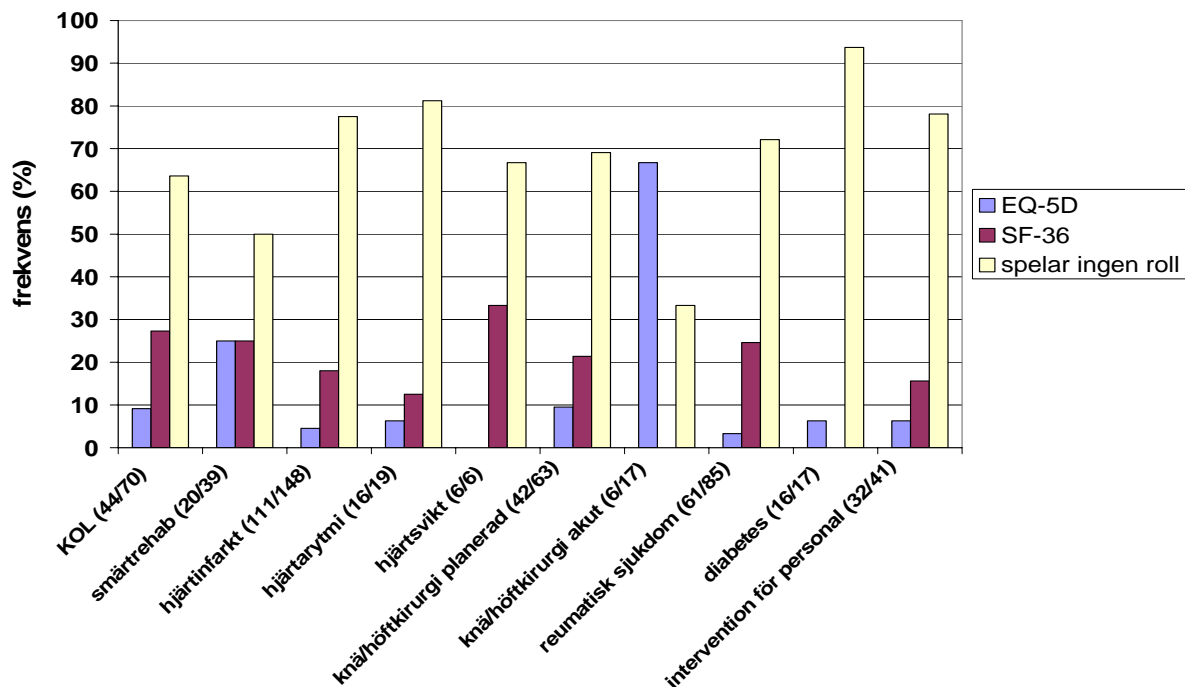
**Figur 16.** Fördelning bland dem som besvarat frågan "Skulle du se något värde i att man inom sjukvården rutinemässigt använder sig av dessa, eller liknande, enkäter i framtiden för att mäta hälsa?" (n=481).



**Figur 17.** Fördelning bland dem som besvarat följdfrågan "Om ja, vilken enkät vore då att föredra?" (n=354, varav 257 som verkligen svarat ja på förra frågan).



**Figur 18.** Fördelning bland dem som besvarat frågan ”Skulle du se något värde i att man inom sjukvården rutinmässigt använder sig av dessa, eller liknande, enkäter i framtiden för att mäta hälsa?”, uppdelat per sjukdomsgrupp/annan grupp. Inom parentes anges hur många som svarat inom varje grupp (antal som svarat/totalantal i gruppen).



**Figur 19.** Fördelning bland dem som besvarat följdfrågan ”Om ja, vilken enkät vore då att föredra?”, uppdelat per sjukdomsgrupp/annan grupp. Inom parentes anges hur många som svarat inom varje grupp (antal som svarat/totalantal i gruppen).

När det gäller svaren på den öppna frågan om varför man föredrar ett visst instrument har 92 (av 505) deltagare valt att svara. De flesta har svarat enligt intentionen, men några har valt att istället svara på varför de tycker hälsovinstmätning som sådan är bra eller dålig. Bland dessa 92 finns alltså även sådana som svarat *nej* på frågan om de ser något värde i hälsovinstmätning och de som svarat *vet inte* på frågan om vilken enkät de föredrog. De som svarar enligt intentionen har valt att antingen tala om varför de föredrar det ena instrumentet eller varför de inte tycker om det andra, varför vi har både positiva och negativa kommentarer om båda instrumenten. Inga systematiska skillnader har kunnat upptäckas, vare sig för kön, ålder eller sjukdomsgrupp.

#### **Positiva kommentarer om SF-36**

- Bättre formulering på frågorna.
- Lite djupare frågor.
- Flera svarsalternativ.
- Något mer nyanserade svar att välja ibland.
- Går lite mer in på hur man mår fysiskt och psykiskt.

#### **Positiva kommentarer om EQ-5D**

- Enklare att förstå och fylla i.
- Mer lättläst och konkret.
- Mer rak på sak – inga frågetecken.
- (VAS-)skalan är bra!
- Snabbare.

#### **Negativa kommentarer om SF-36**

- Man blir lätt förvirrad av SF-36. För många svarsalternativ. En del frågor går in i varandra.
- På SF-36 kan man lätt misstolka frågor och svarsalternativen. När är man sjuk, tex? Vad är god hälsa?

#### **Negativa kommentarer om EQ-5D**

- Det finns svarsalternativ där jag skulle vilja svara mitt emellan två alternativ.
- Man känner sig som knäpatient lite för frisk för frågorna.

#### **Kommentarer från dem som svarat: *spelar ingen roll***

- Båda var lika lätta att fylla i.
- Jag förstod båda.
- Snarlika.
- Frågeställningarna ligger så nära varandra.

Dessa kommentarer gäller alltså såväl den praktiska biten, att förstå och fylla i, som själva strukturen på instrumenten.

Några övriga kommentarer:

- Enkätfrågor/svar alltid ofullständiga.
- Beroende på dagsform som varierar.
- enkät + prover (med prover ett bättre resultat)

Som tidigare nämnts valde några alltså att istället ange vad de tycker om hälsovinstmätning i sig:

#### **Positiva kommentarer om hälsovinstmätning**

- Någon bryr sig.
- För att få upplysning om hur folk mår.
- Insikt i patientens upplevelse.
- För att få fram bredare kunskaper om olika sjukdomar.
- Kan få hjälp om det behövs mera.
- Lättare upptäcka psykisk sjukdom.
- Tidsbesparande vid läkarbesök.
- För att tidigare hitta ohälsa.
- Man blir ”tvungen” att själv fundera över sin situation.

#### **Negativa kommentarer om hälsovinstmätning**

- Enkäter är bara papper. Människan bakom pappret ska man istället ägna sig åt.
- Jag blir inte bättre av att svara på frågor.
- Jag föredrar vård framför enkäter.
- För mycket frågor.

#### **Processanteckningarna**

När det gäller processanteckningarna har upplägget inte fungerat (endast två sjukhus har lämnat in någon form av anteckningar) och därför kan vi inte uttala oss om processen ur medarbetarperspektivet i nuläget, vare sig om själva mätningproceduren eller huruvida medarbetarna tyckte mätresultaten var lätta att förstå och om de kunde använda resultaten som ett lärande för att utveckla sin verksamhet.

## DISKUSSION

Detta är den första delrapporten från ett läroprojekt i sin första fas, med det tredelade syftet att skapa ökad kunskap om processen vid hälsovinstmätning ur patientens och medarbetarnas perspektiv, samt att pröva i vad mån kunskap om patienternas hälsorelaterade livskvalitet efter behandling kan användas som stöd för verksamhetsutveckling, och dessutom jämföra två olika HRQoL-instrument, SF-36 och EQ-5D, i dessa avseenden.

Ur patientperspektivet fann vi att över hälften av patienterna ansåg att hälsovinstmätning är värdefullt, och att valet av instrument inte var avgörande. De flesta tyckte dessutom att (båda) instrumenten var lätta att använda, och gav god möjlighet till att få beskriva ens hälsotillstånd. Däremot har det varit svårt att få in processanteckningar, och eftersom dessa är nyckeln till att se projektet ur medarbetarnas perspektiv har detta mål endast delvis uppfyllts. Man har ute på sjukhusen fått prova på att mäta, men kunskapen om medarbetarnas syn på mätningarna, genomförbarhet och användbarhet för kliniskt förbättringsarbete saknas ännu. Erfarenheter från England visar bland annat på vikten av att resultaten måste gå lätt och snabbt att tolka för användarna och att de presenteras på ett konstruktivt, positivt sätt, med fokus på goda exempel istället för att utpeka någon som ”sämst i klassen” [24]. Rent praktiskt har datorstöd, i form av användning av pekskärm och beräkningsprogram som direkt visar resultatet av en individs mätning, visat sig vara framgångsrikt [20].

När man tittar närmare på själva mätresultaten, ser man i exemplet med diabetespatienterna att problemtygden för EQ-5D låg i dimensionerna *smärtor/besvär* och *oro/nedstämdhet*, och för SF-36 i GH (allmän hälsa)-, VT (vitalitet)- och RE (emotionell rollfunktion)-skalorna. Det relativt höga värdet för BP (=patienterna har inga smärtproblem) pekar på att problemen härrör sig till besvär snarare än till smärtor, vilket tydligt illustrerar hur bred och inkluderande frågan om *smärtor/besvär* är. *Oro/nedstämdhet* verkade inte ge något större utslag i MH-skalan (men möjligen i GH, VT och/eller RE). Kanske detta beror på att EQ-5D bara har tre svarsnivåer för varje dimension? Om man inte känner sig helt problemfri, så återstår att välja mellan måttliga eller svåra problem. Av de nio (av 17) som anger problem i dimensionen *oro/nedstämdhet*, har åtta angett måttliga problem. Kanske har de flesta av dem egentligen lättare problem, vilket inte ger lika stort utslag i ett mer nyanserat instrument som SF-36? Eller så handlade kanske problemen som gav utslag i *oro/nedstämdhet* egentligen om sådana psykiska aspekter som tex VT-skalan är bättre på att fånga, som psykisk ork och energi? Medan GH uppvisade en signifikant förbättring efter interventionen för diabetespatienterna var EQ-5D VAS oförändrad, vilket indikerar att dessa båda mått på allmän hälsa inte mäter riktigt samma sak och därmed inte heller är utbytbara, trots höga korrelationskoefficienter i denna och andra studier (runt 0,7). Tittar man på frågorna som ingår i GH-skalan så finner man, utöver den generella självskattningen av den nuvarande hälsan, också antaganden om hur ens hälsa kommer att utvecklas i framtiden, och jämförelser med andra människors hälsa, vilket kanske kan förklara skillnaderna gentemot VAS-skalan. Dessutom ska man hålla i minnet att medan SF-36 mäter en sorts genomsnittlig (enligt respondentens egen skattning) HRQoL de senaste fyra veckorna (eller senaste veckan för akutversionen), så mäter EQ-5D läget precis den dag mätningen sker (om respondenterna svarar enligt instrumentens intentioner).

Våra fynd stödjer tidigare studier som menar att SF-36 och EQ-5D, trots att båda mäter HRQoL, inte representerar exakt samma domäner och därför inte är helt utbytbara [25].

Ett genomgående tema för alla delprojekt var betydelsen av rollfunktionsskalorna i SF-36, som initialt ofta låg lågt eller mycket lågt. Om vi som exempel tar de planerade knä/höftplastikerna, där den fysiska rollfunktionen inte riktigt verkade hänga med

förbättringarna i fysisk funktionsförmåga och smärta, så var uppföljningstiden möjligtvis för kort (mellan 2 och 3 månader). Det skulle vara mycket intressant att få följa dessa grupper över tid, speciellt med avseende på rollfunktionen. Tar det kanske ett tag innan man vågar tro på sin nya/återvunna förmåga? Det skulle också vara intressant att jämföra med studier där man använt SF-36 version 2, för att se att det inte bara är en effekt av att SF-36 har trubbiga rollfunktionsskalor. Kanske har man visst blivit bättre, men inte helt bra? Till skillnad mot i den ursprungliga SF-36, där man bara kan svara ja eller nej, så kan man ju i version 2 gradera sin rollfunktionsinskränkning efter svårighetsgrad, vilket, som tidigare nämnts, är en fördel vid före- och eftermätningar. För korsbandsskadorna (delprojekt 4 i figur 12) sågs däremot en klar förbättring av RP efter operation, dvs de kände sig inte längre lika hindrade i sitt liv pga sitt kroppsliga tillstånd. Beror detta på själva skadans art (korsband kontra ledytebesvär), någon speciellt framgångsrik del av själva interventionen, eller är det helt enkelt en effekt av att patientgruppen med korsbandsskador var mycket yngre?

De förbättringar i flera fysiska skalor/dimensioner som sågs för de planerade operationerna var tydliga och förväntade. När det gäller de akuta operationerna så var ju föremätningarna skattade värden av tillståndet före skadan. Därmed kan det faktum att man inte såg några signifikanta förändringar mellan före- och eftermätningarna för denna grupp tas som belägg för att patienten fått tillbaka/behålla sin normala HRQoL, och att behandlingen således kan anses vara lyckad (tendensen var dessutom mot något förbättrade värden).

När man samlade (och för statistiska ändamål även slog samman) delprojekten i olika sjukdomsgrupper framträdde vid en jämförelse intressanta likheter och skillnader. Hjärtinfarktpatienternas mätresultat var väldigt likartade för olika sjukhus, medan smärt- och KOL-patienternas var de mest heterogena. De låg mer utspridda initialt och uppvisade tendenser till både förbättringar och försämringar. Reflekterar detta behandlingens/interventionens homogenitet/evidens? Man kan tänka sig att vårdprogrammen för hjärtinfarkt och knä/höftkirurgi är mer enhetliga och evidensbaserade än för tex smärtrehabilitering. För de förra finns också sedan länge etablerade, heltäckande kvalitetsregister, medan tex registret för smärtrehabilitering ännu inte kommit lika långt.. För KOL-patienter saknas idag kvalitetsregister. Eller speglar skillnaderna en varierande grad av psykosociala inslag i sjukdomsbilden, vilket kanske kräver mer individualiserade interventioner? Eller försvinner skillnaderna om man har likstora, och större, grupper?

Enligt svaren på utvärderingsenkäten tyckte över hälften av patienterna att hälsovinstmätning inom rutinsjukvården skulle ha ett värde, men samtidigt har många valt att inte ta ställning. Varför är det så? Är de trötta på att fylla i enkäter? Ovana vid att deras hälsa efterfrågas? Eller har vi inte lyckats förmedla tanken bakom hälsovinstmätning på ett begripligt sätt? Finns det andra orsaker? Kvalitativa intervjustudier skulle vara en möjlighet att få ytterligare information.

På delprojektsnivå (även inom samma sjukdomsgrupp) kunde man ibland se en övervikt i preferensen för det ena eller det andra instrumentet. Speglar detta hur instrumenten redan idag används på sina håll, och de positiva och negativa stämningar runt instrumenten som finns där?

Det sjukhus som använt sig av pekskärm för att svara på SF-36 har i särklass flest positiva svar för SF-36. Beror det på att patienterna föredrar SF-36 eller på att man föredrar att svara via pekskärm? Denna grupp var också den klart yngsta av alla grupper, kanske föredrar yngre datoriserade enkäter? Eller föredrar yngre SF-36?

Vi har bara lyckats hitta en enda studie där man på liknande sätt frågat patienterna om vilket

HRQoL-instrument de föredrar. Det är en studie på MS-patienter (multipel skleros), där man jämförde SF-36, EQ-5D och ett sjukdomsspecifikt instrument [26]. Patienterna tillfrågades om vilket instrument som bäst summerade deras livskvalitet. Enligt författarna föredrog 17% det sjukdomsspecifika instrumentet, 5% SF-36, 2% EQ-5D och 76% alla tre instrumenten tillsammans. Det är dock inte i detalj redovisat hur frågan och svarsalternativen var formulerade, vilket gör att man inte kan låta bli att undra om patienterna egentligen svarat att det inte spelar någon roll vilket instrument man använder, snarare än att man vill ha alla tre tillsammans? Studien måste ändå ses som ett stöd för tanken att kombinera sjukdomsspecifika och generiska instrument, vilket också vi instämmer i. Längre fram i projektet tänker vi oss därför att även inkludera olika sjukdomsspecifika instrument.

Många som förespråkar SF-36 är tveksamma till om EQ-5D ger en tillräckligt nyanserad bild för att fungera som kunskapsunderlag för verksamhetsutveckling. Det är ännu för tidigt i projektet för att kunna uttala sig om detta. Kanske är olika instrument lämpliga i olika situationer? Frågan är först vad det egentligen är vi behöver lära (och mäta) för att kunna möta patienternas behov vid planering och utveckling av hälso- och sjukvårdens verksamhet. Vi har idag väldigt många olika mått på patientens sjukdom och hälsa under sjukhusvistelsen, men vi vet mycket mindre om utvecklingen efter utskrivningen, även om naturligtvis enskilda kliniker runt om i landet på egen hand genomför sådana mätningar. Tex har man vid smärtrehabkliniken på ett av våra medlemssjukhus följt hälsovinstprojektets arbete samtidigt som man själv använt sig av SF-36 på en grupp smärtpatienter och följt dem över en längre tid, även efter avslutad behandling, och dessutom använt sig av jämförelsegrupper (bilaga 4). Även här uppvisade patienterna (oväntat) låga värden för RP, i förhållande till både PF och vårdprofessionernas egna bedömningar, vilket var ny kunskap. Man planerar också för fler uppföljningar längre fram [27].

När man mäter över en längre tid måste man också ta fenomenet *response shift* med i beräkningen när man tolkar sina resultat. Med detta menas här respondentens benägenhet till att över tid förändra de värderingar som ligger bakom hur man väljer svarsalternativ vid självskattningar av hälsorelaterad livskvalitet. Dessa förändringar kan bero på omställningar både i respondentens yttre och inre förutsättningar [28].

### **Metoddiskussion**

Eftersom detta är ett läroprojekt i sin första fas har vi medvetet tillåtit delprojekten att ha få deltagare, något som förstås kan göra statistiska analyser vanskliga. Framtida mätningar bör därför inkludera fler deltagare, för säkrare resultat.

Vi har heller inte använt kontrollgrupper, vilket medför att vi inte vet hur patienter utan intervention skulle ha skattat sin hälsorelaterade livskvalitet. Det är dock tveksamt om kontrollgrupper, i betydelsen ingen intervention alls, är görligt ur etisk synvinkel för de flesta patientgrupper. Däremot har vi möjlighet jämföra olika interventioner på motsvarande patientgrupper runt om i landet (alla kliniker arbetar inte på samma sätt), och låta dem fungera som varandras kontroller.

I manualerna till båda instrumenten kan man läsa rekommendationen att de ska placeras först om man tänker inkludera dem i ett större enkät häfte. Endast ett kan dock vara först, och i denna studie placerades EQ-5D före SF-36. Vår tanke är att det inte borde ha någon avgörande betydelse vilket som kommer först, eftersom båda är HRQoL-instrument och dessutom de enda ingående instrumenten i frågeformuläret, som därmed också är väldigt kort.



I EQ-5D manualen kan man också läsa att medelvärden för grupper kan beräknas för varje dimension och redovisas i stapel- eller linjediagram. Egentligen är det förstås statistiskt inkorrekt att beräkna medelvärden på ordinaldata, dvs data där stegen mellan svarsalternativen (här: inga, måttliga eller svåra problem) inte är lika stora, speciellt när fördelningarna är så skeva som de är för EQ-5D dimensionerna. Vi har valt att ändå visa detta presentationssätt (figur 5), eftersom det finns beskrivet i EQ-5D manualen, men vi använder oss framförallt av det mer matematiskt korrekta presentationssätt (procentuell fördelning) som numera förespråkas. För att bättre motsvara presentationen av skalorna i SF-36 har vi valt att använda linjediagram även för EQ-5D dimensionerna.

När det gäller utvärderingsenkäten är bortfallet för frågorna om EQ-5D större än för frågorna om SF-36. Vår gissning är att detta beror på layouten, där de förra återfinns i högerspalten och de senare i vänsterspalten. Skulle det ha varit tvärtom är vår gissning att frågorna om SF-36 hade haft störst bortfall. Vi tar detta som ytterligare en indikation på att man inte ser någon större skillnad mellan instrumenten.

Betydelsen av *case-mix* (patientheterogenitet; en viss sjukdoms/patientgrupp består av individer med olika kön, ålder, socioekonomisk status, sjukdomsgrad, komorbiditet mm, vilket försvårar generaliseringar) har uppmärksammats mycket på senare tid. Tex fann man nyligen i en studie på patienter som vårdats på en intensivvårdsavdelning (IVA) att dessa (som hel grupp) ett halvår efter utskrivning hade lägre HRQoL än en kontrollgrupp från normalbefolkningen. Om man istället delade in IVA-patienterna i grupper efter de sjukdomar (om någon) de hade innan de hamnade på IVA, och jämförde dessa grupper med individer ur kontrollgruppen som hade samma sjukdomar, så blev bilden en annan. Skillnaderna i HRQoL mellan IVA-patienter och individerna i normalbefolkningen minst halverades, och försvann i vissa fall helt [29]. Även om studier visat att man, som förväntat, har sämre HRQoL ju fler sjukdomar man har, är mycket fortfarande oklart. Är antalet sjukdomar viktigast, eller är typen av sjukdom och/eller sjukdomsgrad viktigare? Är vissa kombinationer av sjukdomar ”bättre” och andra ”sämre”? Om man vill mäta förändring i en viss sjukdomsgrupps HRQoL efter en intervention inriktad just på den aktuella sjukdomen kan förändringar i patientens övriga sjukdomar naturligtvis ”störa” mätningen. Vi hade inkluderat en sjukdomslista, men ej mått på enskilda sjukdomars påverkansgrad, något som bör justeras i det fortsatta arbetet. Dessutom hade sjukdomslistan låg svarsfrekvens, med osäker kvalitet på svaren, och bör även av det skälet omarbetas. De ko/multimorbiditetsdata som insamlats hittills behöver bearbetas ytterligare innan de (eventuellt) kan användas.

## Konklusion

Detta är den första delrapporten från nätverket Hälsöfrämjande Sjukhus (HFS) hälsovinstmättningsprojekt. 18 av 23 medlemssjukhus deltog med ett eller flera delprojekt, där man har mätt hälsorelaterad livskvalitet för några olika utvalda sjukdomsgrupper, före och efter en behandling/intervention. Syftet var att studera genomförbarheten av hälsovinstmätning inom rutinsjukvården ur patient- och medarbetarperspektiv, och dess användbarhet för kliniskt förbättringsarbete, samt att jämföra HRQoL-instrumenten EQ-5D och SF-36 i dessa avseenden.

Ur patientperspektiv fann man att över hälften av patienterna ansåg att hälsovinstmätning är värdefullt och bara 5% hade en negativ inställning. De verkade heller inte tycka att valet av instrument (SF-36 eller EQ-5D) spelar någon avgörande roll. Flertalet tyckte också att (båda) instrumenten var lätta att använda och att de gav god möjlighet till att få beskriva ens hälsotillstånd.

Resultatet av själva mätningarna ger kunskap om utvecklingen av patienternas (egenupplevda) hälsa efter avslutad behandling. Trots små delprojekt kunde man se tydliga skillnader mellan sjukdomsgrupperna. Hjärtinfarktgruppen, tex, uppvisade en mycket likartad bild för de ingående delprojekten, med klara förbättringar efter intervention. Andra grupper uppvisade en mer inhomogen bild, där vissa delprojekt uppvisade förbättringar och andra inte, kanske beroende på mer skilda interventionsmodeller på de olika sjukhusen för dessa sjukdomsgrupper.

Även om SF-36 och EQ-5D överlag uppvisade likartade övergripande bilder över patienternas HRQoL, gav SF-36, som förväntat, en mer nyanserad bild än EQ-5D.

Medarbetarperspektivet och verksamhetsutvecklingspotentialen blir huvudfokus för studiens nästa fas. Hur använder vi på bästa sätt kunskap om patienternas hälsorelaterade livskvalitet inom hälso- och sjukvården, inte bara för hälsoekonomiska beräkningar, utan för att bättre möta patienternas behov?

Ovanstående resultat presenterades på ett HFS processledarmöte 21 november, 2005 i Stockholm. Då beslutades också att nätverket fortsätter arbetet genom den interna arbetsgrupp som nyligen bildats för hälsovinstmätning.

Evalill Nilsson, Preben Bendtsen och Margareta Kristenson  
för nätverket HälsöFrämjande Sjukhus (<http://www.natverket-hfs.se>)  
Linköping mars 2006

Tack till Ulrika Öberg, Eksjö/Jönköping för värdefull hjälp med rapportskrivningen.

## REFERENSER

1. Treurniet HF, Essink-Bot M, Mackenbach JP, van der Maas PJ. Health-related quality of life: an indicator of quality of care? *Qual Life Res* (1997) 6: 363-369
2. Ware JE, Sherbourne CD. The MOS 36-item Short Form Health Survey (SF-36): Conceptual framework and item selection. *Med care* (1992) 30(6): 473-83
3. Sullivan M, Karlsson J, Ware JE. The Swedish SF-36 Health Survey I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Soc Sci Med* (1993) 41: 1349-58
4. Sullivan M, Karlsson J, Ware JE. SF-36 Hälsöenkät. Manual och tolkningsguide. 1994 Göteborg.
5. <http://www.sf-36.org/>
6. <http://www.hrql.se/content/hrql/>
7. Brooks R, with the EuroQol group. EuroQol: the current state of play. *Health policy* (1996) 37: 53-72
8. Brooks R, Rabin RE, de Charro Fth (Eds). The measurement and valuation of health status using EQ-5D: a European perspective. (2003) Kluwer Academic Publishers
9. Henriksson M, Carlsson, P. Att mäta hälsorelaterad livskvalitet – en beskrivning av instrumentet EQ-5D. CMT rapport 2002:1 (2002), Linköpings Universitet
10. <http://www.euroqol.org/>
11. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 survey. *J Clin Epidemiol* (1998) 51(11): 1115-28
12. Wenemark M, Kristenson M, Noorlind Brage, H. Hur speglar instrument för hälsorelaterad livskvalitet olika självrapporterade sjukdomar och symptom? Folkhälsövetenskapligt Centrum i Östergötland rapport 2004:5 (2004) (<http://www.lio.se/fhvc>)
13. Burström K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. *Qual Life Res* (2001) 10: 621-35
14. Eriksson E, Nordlund A. Hälsa och hälsorelaterad livskvalitet mätt med EQ-5D och SF-36 i Östergötlands och Kalmar län: Resultat från befolkningsenkäterna. Folkhälsövetenskapligt Centrum i Östergötland rapport 2002:1 (2002) (<http://www.lio.se/fhvc>)
15. Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary scales: A user's manual. Boston, MA. The Health Institute, 1994
16. Sullivan M. Konsten att läsa en artikel om livskvalitetsforskning. *Läkartidningen* (2002) 99 (26-27): 2933-8
17. Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: A structured review of generic self-assessed health instruments. *Qual Life Res* (2005) 14: 1651-68
18. Nordlund A, Ekberg K, Kristenson M & Linqwest Group. EQ-5D in a general population survey – A description of the most commonly reported EQ-5D health states using the SF-36. *Qual Life Res* (2005) 14: 1099-1109
19. Kosinski M, Zhao S, Dedhiya S, Osterhaus J, Ware J.E. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis & Rheumatism* (2000) 43(7): 1478-87

20. Bendtsen P, Leijon M, Sommer A, Kristenson M. Measuring health-related quality of life in patients with chronic obstructive pulmonary disease in a routine hospital setting: Feasibility and perceived value. *Health Qual Life Outcome* (2003) 1:5 (<http://www.hqlo.com/content/1/1/5>)
21. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Qual Life Res* (2003) 12: 599-607
22. Taft C, Karlsson J, Sullivan M. Performance of the Swedish SF-36 version 2.0. *Qual Life Res* (2004) 13: 251-6
23. Taft C, Karlsson J, Sullivan M. DoSF-36 summary component scores accurately summarize subscale scores? *Qual Life Res* (2001) 10: 395-404
24. Vallance-Owen A, Cubbin S, Warren V, Matthews B. Outcome monitoring to facilitate clinical governance; experience from a national programme in the independent sector. *J Pub Health* (2004) 26 (2):187-92
25. Öberg U, Öberg T. Do EuroQoL and SF-36 yield equivalent information, or do they represent different domains of knowledge? Proceedings of 18<sup>th</sup> Plenary Meeting of the Euroqol group. ISBN 91-88042-17-0 (2001) 177-185
26. Moore F, Wolfson C, Alexandrov L, Lapierre Y. Do general and Multiple Sclerosis-specific Quality of Life instruments differ? *Can J Neurol Sci* (2004) 31: 64-71
27. Personlig kommunikation: dr Sten Larsson, läkare i smärtgruppen, Rehabiliteringskliniken vid Ängelholms sjukhus.
28. Sprangers M, Schwartz C. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med* (1999) 48: 1507-15
29. Orwelius L, Nordlund A, Edell-Gustavsson U, Simonsson E, Nordlund E, Kristenson M, Bendtsen P, Sjöberg F. Role of preexisting disease in patient's perception of health-related quality of life after intensive care. *Crit Care Med* (2005) 33(7): 1557-64

## **BILAGOR**

1. Enkätbäftet
  - a. Patientkarakteristik vid föremätningen
  - b. Patientkarakteristik vid eftermätningen
  - c. EQ-5D
  - d. SF-36
  - e. Utvärderingsenkäten (vid eftermätningen)
2. Interventionsbeskrivningar för alla delprojekt
3. Gemensamma diagram för Hjärtinfarkt delprojekt 6,7,8 och Övrigt delprojekt 4, samt diagram för Övrigt delprojekt 1, 2 och 5.
4. SF-36 diagram, Smärtrehabiliteringskliniken, Ängelholms sjukhus