

# Helseøkonomisk evaluering av fysioterapibehandling i primærhelsetjenesten

*En metodestudie*

Inger Holm



MASTEROPPGAVE

15.10.2013

Avdeling for helseledelse og helseøkonomi  
Institutt for helse og samfunn,  
Det medisinske fakultet

UNIVERSITETET I OSLO

© Forfatter

År: 2013

Tittel: Helseøkonomisk evaluering av fysioterapibehandling i primærhelsetjenesten. En metodestudie

Forfatter: Inger Holm

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

# Sammendrag

**Introduksjon:** Pasienter med muskel- og skjelett lidelser utgjør en av de største gruppene pasienter som oppsøker primærhelsetjenesten. Fysioterapi er en viktig del av tilbudet til disse pasientene og hensikten med behandlingen er å dempe smerte og å bedre funksjon og livskvalitet. Myndighetene har som mål å få mest mulig helse for hver krone, og det er derfor viktig at fysioterapeuter kan dokumentere nytten av og kostnadene ved behandlingen som gis. Hovedhensikten med denne metodestudien var å evaluere responsiviteten til de tre generiske helserelaterte livskvalitetsskjemaene EQ-5D, 15 D og SF-6D som brukes i helseøkonomiske cost-utility analyser og hvilke(t) som eventuelt egner seg best i slike analyser etter en behandlingsserie hos fysioterapeut i primærhelsetjenesten.

**Materiale og metode:** Aktuelle pasienter ble rekruttert av 5 fysioterapeuter i primærhelsetjenesten. I utgangspunktet kunne alle nye pasienter som kom til terapeuten bli spurt om å delta i studien, de eneste kriteriene for inklusjon var: Pasienten skulle være i arbeidsfør alder (18-67 år) og ha god norsk språkforståelse både muntlig og skriftlig. Pasientene fylte ut livskvalitetsskjemaene EQ-5D, 15 D og SF-6D før behandling og etter avsluttet behandling. I tillegg fylte pasienten ut to endrings spørsmål (global rating scales) i forhold til plager og funksjon (disse ble brukt som ”anker” i responsivitetsanalysene).

**Resultater:** 27 pasienter, 70 % kvinner ble inkludert i studien. Gjennomsnittsalderen var  $53,8 \pm 16,5$  år og de hadde hatt symptomer i  $78,5 \pm 178$  uker. Behandlingen varte i gjennomsnitt  $12,5 \pm 3,2$  uker og pasientene mottok  $14 \pm 7$  behandlinger. Resultatene viste at ca 90 % av pasientene rapporterte en bedring både når det gjaldt plager og funksjon. To av livskvalitetsskjemaene, EQ-5D og 15D, viste en høysignifikant endring i utility index og en stor effektstørrelse fra før til etter behandling, en effektstørrelse på henholdsvis 0,34 og 0,48. Målt med SF-6D var det ingen endring.

**Konklusjon:** Både EQ-5D og 15D kan brukes kostnadsfritt, er enkle for pasientene å fylle ut og finnes i elektronisk utgave. De har begge god responsivitet, dvs evne til å fange opp endring over tid. I denne tidlige fasen av cost-utility analyser i fysioterapi vil det derfor sannsynligvis være nyttig å inkludere begge skjemaene, da har man mulighet til å teste ut hvilket av skjemaene som egner seg best til ulike pasientgrupper og for ulike formål.

# Summary

**Purpose:** To evaluate the responsiveness of the generic health-related quality of life questionnaires EQ-5D, 15D and SF-6D used for patients with musculoskeletal problems receiving primary care physiotherapy treatment.

**Design:** Prospective cohort study

**Material and methods:** 5 physiotherapists working in outpatient clinics recruited eligible patients to the study. Inclusion criteria were: age between 18 and 67 years and ability to read and write Norwegian independently. The patients filled in the EQ-5D, 15D and SF-6D prior to treatment and after the last treatment session. They also answered two global rating scales about change in complaints and function, these scales were used as an anchor in the responsiveness analysis.

**Results:** 27 patients, 70% women, were included in the study. Their mean age was  $53.8 \pm 16.5$  years, their mean duration of symptoms was  $78.5 \pm 178$  weeks. The treatment lasted for  $12.5 \pm 3.2$  weeks and the number of treatments was  $14 \pm 7$ . The mean direct treatment cost was  $3338 \pm 1310$  NOK. There was a high correlation between the EQ-5D and 15D utility indexes at baseline, while for SF-6D there were only insignificant correlations with the indexes calculated from the two other questionnaires. Ninety % of the patients reported a positive change both in complaints and function. EQ-5D and 15D showed a highly significant change in utility index from baseline to the last treatment, with an effect size of 0.34 and 0.48, respectively.

**Conclusion:** EQ-5D and 15D both showed acceptable responsiveness and high effect sizes. They are both simple to fill in, are free of costs and electronic versions are eligible for both of them. In this early stage of cost-utility analysis in physiotherapy it might be reasonable to include both questionnaires in future protocols to evaluate which is the most favourable to use for different diagnostic groups and in different settings.

# Forord

Skal man gjennomføre prosjekter i klinisk praksis, er man avhengig av velvillige kolleger som kan formidle kontakt med pasientene og administrere den praktiske delen av prosjektet. Jeg er så heldig å ha gode kolleger /venninner som også denne gang har stilt opp når nye problemstillinger skal besvares (det er ikke første gang). Jeg vil derfor få takke Irma Blaker, Gro Greftegreff, Ragnhild Karlsen, Vivi Nyberg og Arnhild Skjølberg for inkludering av pasienter og administrering av spørreskjemaene før og etter behandling. Uten dere hadde jeg ikke hatt data til denne oppgaven. Dokker e topp ;o)

Tusen takk til Tor Iversen som har vært min veileder. Du har hele tiden gitt meg raske tilbakemeldinger, gode innspill og faglige råd. Selv om prosjektet ikke inneholder en fullverdig økonomisk analyse, har jeg lest og lært utrolig mye, og fått en god basis for arbeid med kostnad-nytte analyser i fremtidige studier.

Takk også til min arbeidsgiver, Oslo Universitetssykehus, KKN, Seksjon for forskning og kompetansesenter, som ga meg muligheten til å være masterstudent ved siden av jobben. Det har vært utrolig lærerikt og motiverende. Regner absolutt med jeg kan gi noe tilbake.

Oslo 15. oktober 2013

Inger Holm

# Innholdsfortegnelse

1	Introduksjon .....	1
1.1	Bakgrunn .....	1
1.2	Hensikt / problemstilling .....	2
1.2.1	Hensikt .....	2
1.2.2	Problemstilling .....	2
2	Teori .....	3
2.1	Muskel- skjelett lidelser, prevalens og kostnader.....	3
2.2	Effekt av fysioterapi .....	4
2.3	Helseøkonomiske evalueringer.....	4
2.4	Kostnader.....	5
2.5	Kostnad-nytte (CUA) analyser .....	6
2.6	Hva er et kvalitetsjustert leveår (QALY)?.....	8
2.7	Verdsetting av ulike helsetilstander.....	8
2.8	Preferansebaserte livskvalitetsskjemaer til bruk i kostnad-nytte analyser .....	9
2.8.1	EQ-5D .....	9
2.8.2	15D.....	10
2.8.3	SF-6D (SF-36).....	10
2.9	Hvilke(t) helserelevanter livskvalitetsskjema skal man bruke i studier som evaluerer effekt av fysioterapi?.....	12
2.10	Responsivitet.....	12
2.11	Kostnad-nytte analyser av fysioterapi i primærhelsetjenesten.....	13
3	Materiale og metode.....	15
3.1	Materiale .....	15
3.2	Metode .....	15
3.3	Statistiske analyser .....	16
4	Resultater.....	18
5	Diskusjon.....	25
5.1	Utility index ved baseline .....	25
5.2	Endring i utility index.....	26
5.3	Kan livskvalitetsskjemaene brukes om hverandre?.....	27
5.4	Endring i plager og funksjon .....	29

5.5	Kostnader.....	30
5.6	Utvalget .....	31
5.7	Hvilket skjema egner seg best i fysioterapipraksis?.....	32
5.8	Et lite apropos.....	33
6	Konklusjon .....	35
7	Referanseliste .....	36
8	Vedlegg .....	42





# 1 Introduksjon

## 1.1 Bakgrunn

Muskel- skjelett lidelser er et av de største helseproblemene i den vestlige verden, er den vanligste årsaken til alvorlige, langvarige smerter og nedsatt funksjon<sup>1</sup> og koster samfunnet store summer. I Norge fører muskel- skjelett lidelser til utstrakt bruk av ulike helsetjenester og er den hyppigste årsaken til langvarige sykemeldinger (46 %) <sup>2,3</sup> og uføretrygd (33 %) <sup>3-5</sup>. Grad av plager og konsekvenser for pasientene når det gjelder livskvalitet og aktivitetsnivå (fysisk aktivitet, arbeid osv), varierer svært mye fra individ til individ.

Pasienter med muskel- og skjelett lidelser utgjør en av de største gruppene pasienter som oppsøker primærhelsetjenesten. Fysioterapi er en viktig del av tilbudet til disse pasientene og hovedhensikten med behandlingen er å dempe smerte og å bedre funksjon og livskvalitet. Det finnes høykvalitets vitenskapelig dokumentasjon (randomiserte studier og oversiktsartikler) på at fysioterapi har god effekt på for eksempel artrose<sup>6</sup> og ryggplager<sup>7</sup>. Aktive øvelser og trening er den behandlingsformen som har best dokumentert effekt. For mange pasienter betyr det å øke sitt aktivitetsnivå en betraktelig livsstilsendring, noe som krever oppfølging over lang tid.

Samhandlingsreformen trådte i kraft i januar 2012<sup>8</sup>. Primærhelsetjenesten har fått både nye oppgaver og større ansvar enn tidligere, blant annet skal man forebygge fremfor å reparere, prioritere tidlig framfor sen innsats og oppfølging av personer med kroniske lidelser skal bli bedre. Kommunene har fått overført friske midler, men må allikevel prioritere strengt mellom ulike helsetiltak. Når prioriteringsbeslutninger skal tas, er sykdommens alvorlighetsgrad og nytte- og kostnadseffektivitet viktige faktorer<sup>9</sup>. Prioriteringsdebatten blant fagfolk og politikere handler i hovedsak om hvor mye penger som skal brukes på nye og ofte dyre medisinske behandlingsmetoder, mens behandling av kronikere har lite fokus. Innen fysioterapifaget er det liten tradisjon for å gjøre slike analyser, så i tiden fremover blir det viktig for fysioterapeuter i primærhelsetjenesten å dokumentere den samfunnsøkonomiske nytten av behandlingen pasientene mottar. Fysioterapi koster lite og kan bidra mye i forhold til å holde folk i arbeid og å hjelpe pasientene til en varig endring av livsstil i form av økt aktivitetsnivå og bedre kostvaner.

## **1.2 Hensikt / problemstilling**

### **1.2.1 Hensikt**

Hovedhensikten med denne metodestudien er å evaluere responsiviteten til utility indexene (nytte indeks) utledet fra de helserelaterte livskvalitetsskjemaene EQ-5D, 15 D og SF-6D og i tillegg å gjøre kostnad-nytte analyser etter en behandlingsserie hos fysioterapeut i primærhelsetjenesten. Hovedvekten i oppgaven vil være på skjemaenes anvendbarhet og evne til å fange opp de eventuelle endringene pasientene opplever i løpet av et behandlingsforløp (metodiske egenskaper) og kostnadssiden og selve cost-utility analysen vil derfor ha mindre fokus

### **1.2.2 Problemstilling**

1. Skjer det en endring i utility index, kalkulert ved hjelp av EQ-5D, 15 D og/eller SF-36/6D, etter en behandlingsserie hos fysioterapeut i primærhelsetjenesten?
2. Hva er den gjennomsnittlige kostnaden pr QALY for fysioterapi gitt til pasienter med kroniske muskelskjelett lidelser i primærhelsetjenesten?
3. Hvilket av skjemaene for helserelatert livskvalitet egner seg best til bruk i cost-utility analyser av fysioterapi i primærhelsetjenesten?

## 2 Teori

### 2.1 Muskel- skjelett lidelser, prevalens og kostnader

Muskel- skjelett lidelser er de kroniske lidelsene som har størst utbredelse i befolkningen <sup>3</sup>. I Europa har 20- 30 % av befolkningen til en hver tid en muskel- skjelett relaterte plager <sup>10</sup>. Breivik og medarbeidere publiserte i 2006 artikkelen ”Chronic pain in Europe” <sup>11</sup>, en kartleggingsstudie som inkluderte 46.394 respondenter fra 15 land. Resultatene viste at de mest vanlige smertelokalisasjonene var rygg og kne (58 %) og at den vanligste årsaken til kronisk smerte var kneleddsartrose (34 %). Norge var det landet som hadde den høyeste prevalensen av kronisk smerte (Numeric rating scale >5), hele 30 %, i de 14 andre landene varierte prevalensen fra 12 – 27 %. 79 % av pasientene med kronisk smerte svarte at de hadde opplevd økt smerte som direkte resultat av aktivitet.

I sin rapport fra 2004 slår Nasjonalt ryggnettverk fast at stadig flere får alvorlige kroniske helseplager <sup>3</sup> og at ca 50 % av befolkningen hadde vært plaget med smerter og nedsatt funksjon i muskler og ledd minst 3 måneder sammenhengende i løpet av et år. Dette er skremmende høye tall og Brage og medarbeidere konstaterte i 2010 at det ser ut som denne forekomsten av muskel- og skjelettlidelser holder seg stabil <sup>12</sup>.

Muskel- skjelett lidelser er lidelser som ofte er lite synlige i helseøkonomisk statistikk, men er av de hyppigste årsakene til at pasienter oppsøker både tradisjonelle og alternative helsetjenester <sup>13</sup> og koster samfunnet store beløp, både som direkte utgifter til behandling og rehabilitering, men også indirekte i form av redusert produksjon <sup>10</sup>. I 2004 ble de samlede kostnadene i Norge beregnet til mellom 37 og 44 milliarder kroner <sup>3</sup>. Muskel- skjelett lidelser er en av de hyppigste grunnene til at folk oppsøker fastlegen sin <sup>14</sup>, ca 20 % alle konsultasjoner gjelder muskel- skjelett problematikk <sup>15</sup>. Fastlegene sender ofte disse pasientene videre til annet helsepersonell som fysioterapeuter, fysikalskmedisinere, kiropraktorer osv. Tall fra HELFO viser at det i 2011 var ca 4.240 fysioterapeuter med driftsavtale med kommunene som hadde sendt inn krav om refusjon og at det gjennomsnittlige årlige refusjonsbeløpet pr fysioterapeut var ca. kr 437.000,- <sup>16</sup>, noe som utgjør et betydelig årlig beløp (ca 1,9 milliarder kr pr år). Det er dessuten mange fysioterapeuter som driver uten driftsavtale, da må pasienten betale alle behandlingskostnadene selv.

## 2.2 Effekt av fysioterapi

Fysioterapifaget har ikke lange vitenskapelig forskningstradisjoner, men de siste to ti-årene er det blitt gjennomført en mengde kvalitetsstudier (randomiserte kontrollerte effektstudier (RCTer)). Det har i løpet av samme tidsperiode vært en betydelig endring i behandlingsmetodene fysioterapeuter benytter, fra passive (massasje, elektroterapi osv) til aktive behandlingsformer som øvelser, trening og fysisk aktivitet i dagliglivet, hvor pasienten i større grad må bidra aktivt og ta ansvar for egen helse. Det finnes f. eks god dokumentasjon for at øvelsesprogrammer er effektive både for pasienter med kronisk rygg<sup>7</sup> og for pasienter med kneleddsartrose<sup>17</sup>. Smith og Grimmer-Somers<sup>7</sup> publiserte i 2010 en oversiktsartikkel som evaluerte effekten av øvelser og trening for pasienter med kroniske ryggplager, hvor 15 artikler av moderat kvalitet ble inkludert. De konkluderte med at sammenlignet med kontrollgrupper, reduserte ulike aktive treningsprogrammer både pasientenes smerter og hyppighet av nye ryggepisoder i inntil 6 måneder etter avsluttet behandling. I en systematisk Cochrane oversikt fra 2009<sup>17</sup> konkluderte Fransen et al med at det er god dokumentasjon for at øvelser og trening reduserer smerte og øker funksjon hos pasienter med kneleddsartrose. Effekten av behandlingen var forholdsvis liten, men helt sammenlignbar med effekten av NSAID for samme pasientgruppe. Hagen og medarbeidere publiserte i 2012 en oversikt over systematiske oversikter som hadde evaluert effekt av øvelser og trening på ulike muskel- og skjelettlidelser<sup>15</sup>. Ni oversiktsartikler med til sammen 224 enkeltstudier ble inkludert. Konklusjonen var at øvelser og trening har effekt på smerte og funksjon for de fleste inkluderte diagnosegruppene. Effekten var best hos pasienter med kneleddsartrose og ryggproblemer og effekten økte når antall treningsøkter økte.

## 2.3 Helseøkonomiske evalueringer

Helseøkonomiske evalueringer inkluderer både kostnader og konsekvenser. Disse evalueringene er avgjørende for å kunne ta beslutninger når det gjelder fremtidige prioriteringer som både kan gjelde valg mellom ulike behandlingsalternativer for samme problem eller prioritering mellom ulike grupper av pasienter. Hovedhensikten med en helseøkonomisk evaluering er å identifisere, måle, verdsette og sammenligne kostnader og konsekvenser av ulike behandlinger/helsetjenester<sup>18</sup>. Det finnes tre ulike former for fullverdige helseøkonomiske analyser: Cost-effectiveness analysis (CEA), Cost-utility analysis (CUA) og Cost-benefit analysis (CBA). Alle analysemetodene bruker penger

(monetære enheter) som mål på de ulike kostnadene som er inkludert. Ulikhetene er knyttet til måling og verdsetting av den nytten de ulike tiltakene gir. Enkelt kan man karakterisere de ulike analysene slik <sup>18</sup>:

Cost-effectiveness analysis (CEA) - måler kostnader i *monetære enheter* og konsekvenser i *naturlige enheter* (smerte-reduksjon, økt muskelstyrke osv)

Cost-utility analysis (CUA) - måler kostnader i *monetære enheter* og konsekvenser i *år med god helse* (QALYs)

Cost-benefit analysis (CBA) - måler både kostnader og konsekvenser i *monetære enheter*

## 2.4 Kostnader

Siden ressurser er knappe, vil en ressursanvendelse fortrenge en annen. Alternativkostnader uttrykker hva vi går glipp av i form av alternativ ressursanvendelse ved å bruke ressurser i et bestemt prosjekt. Alternativkostnadene er som oftest ikke observerbare, og det må gjøres tilnærminger. I dette prosjektet gjøres tilnærmingen ved å bruke kostnader slik de observeres i form av takster og pasientenes egenbetaling. Når man skal gjøre kostnad-nytte analyser er det viktig å *identifisere* alle relevante kostnader forbundet med behandlingen/ressursene som blir brukt og bestemme hvilke kostnader man ønsker å inkludere i analysen <sup>18</sup>. Man bør også ta standpunkt til hvilken arena kostnadene skal identifiseres i forhold til; f.eks behandlende enhets kostnader eller kostnader i et samfunnsperspektiv. Når man har identifisert alle kostnadene man vil inkludere i en studie, må man *kvantifisere* kostnadene; f.eks antall konsultasjoner hos lege/fysioterapeut, mengde medisiner, transport (antall km) osv. Når kostnadene er identifisert og kvantifisert skal de *verdsettes*. Dette er ofte den vanskeligste delen, men i eksemplet over kan man forholde seg til eksisterende markedspris/kroneverdien: både konsultasjoner, medisiner og transport kan forholdsvis enkelt verdsettes i kronebeløp.

Kostnadene kan deles inn i direkte og indirekte kostnader <sup>19</sup>.

*Direkte kostnader:* Refusjonstakster/egenandeler ved konsultasjoner hos ulike helsepersonellgrupper (lege, fysioterapeut, psykolog osv), medikamenter, sykehusinnleggelse, røntgenundersøkelser, hjelp i hjemmet, hjelpemidler og rehabiliteringsopphold er eksempler på direkte kostnader knyttet til behandling.

*Indirekte kostnader:* Produksjonstap (fravær fra jobb for pasient og/eller pårørende), reisetid og reiseutgifter i forbindelse med behandling osv.

Det kan ofte være vanskelig å skaffe oversikt over alle kostnader forbundet med behandling og hva som inkluderes, varierer mye mellom ulike studier. Man kan benytte selvrapporteringsskjemaer og be pasienten selv gi informasjon om utgiftene de har hatt. Problemet med selvrapportering er at det er lett for pasientene å glemme enkelte kostnader, eller at de over- eller underrapporterer. I Norge finnes det dessuten kilder hvor mye nyttig informasjon om kostnader kan skaffes til veie: HELFO har oversikt over refusjon til alle behandlere utenfor sykehus og til all refusjon av utgifter til enkeltpersoner (utgifter til behandling, legemidler, transport osv, som pasienter har rett til å få dekket etter Folketryktdoven), Reseptregisteret har oversikt over alle utdelte reseptbelagte legemidler osv. Å få tilgang til denne informasjonen kan imidlertid være en møysommelig og tidkrevende prosess.

## **2.5 Kostnad-nytte (CUA) analyser**

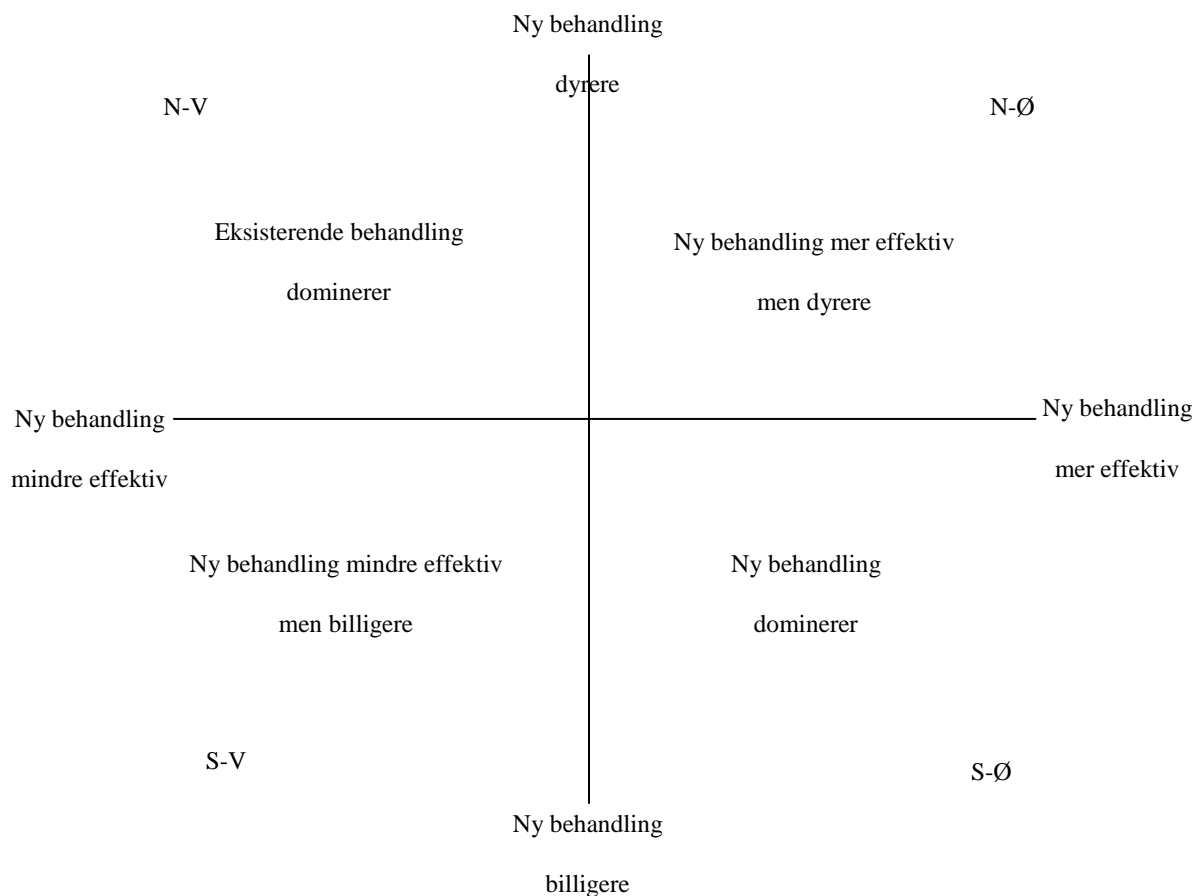
I en kostnad-nytte-analyse studerer man effekten av konkrete helsetiltak og undersøker forholdet mellom gevinstene for pasienter og samfunn på den ene siden og ressursbruk på den andre siden, med andre ord ”hvor mye verdi som skapes pr krone”<sup>20</sup>. I en CUA bruker man måleenheten et kvalitetsjustert leveår, QALY (quality adjusted life year) for å måle helseeffekter av ulike intervensjoner<sup>21</sup>. Man kan med andre ord vurdere kvaliteten av leveårene man ”vinner”, ikke bare antall år<sup>18</sup> og man kan måle helse på tvers av alvorlighetsgrad, ulike intervensjoner og demografiske og kulturelle sub-grupper<sup>22</sup>. Den helserelaterte livskvaliteten er sammensatt av dimensjonene fysisk og sosiale funksjoner og mentalt velvære og er basert på individets egen oppfatning av velvære<sup>22, 23</sup>.

I tillegg må alle kostnader identifiseres. Man kan da sammenligne kostnader og effekt av ulike behandlingsformer (cost per QALY gained or lost) og avgjøre hvilken som er mest kostnadseffektiv, dvs  $\Delta$  kostnader /  $\Delta$  effekt (målt i QALYs). Slike analyser kan brukes som en indikasjon på hvor man får mest helse for pengene, og hvor man bør sette inn ressursene. Hvis man sammenligner to behandlingsformer kan dette uttrykkes i kostnad pr enhet helseforbedring<sup>24</sup>. Man kalkulerer da en inkrementell kostnad-effekt-ratio (IKER). IKER = (kostnad ny behandling – kostnad eksisterende behandlingen) / QALY ny – QALY

eksisterende) = merkostnad / mereffekt (forelesning Olsen 04.03.13). Det er imidlertid viktig å vite at når man sammenligner effekten mellom ulike studier, kan IKER variere mye alt etter hvilken komparator som er valgt, dvs hvilke kostnader man inkluderer og hvilken behandling man sammenligner med.

IKER kan presenteres visuelt ved hjelp av et "incremental cost-effectiveness plane" (Figur 1)<sup>25</sup>. Inkrementell kostnader er vanligvis plassert på nord-sør aksen og behandlingseffekten på øst-vest aksen, begge aksene kan være negative, null eller positive. Dette visualiserer på en enkel måte sannsynligheten for at f.eks en ny behandling er mer kostnadseffektiv enn den eksisterende (S-Ø).

Figur 1



## 2.6 Hva er et kvalitetsjustert leveår (QALY)?

Et kvalitetsjustert leveår (QALY) er et effektmål som har to komponenter, kvantitet (livslengde) og kvalitet (livskvalitet) av et forlenget liv. En QALY er altså produktet av forventet levelengde (overlevelse) og kvaliteten (helsemessig) på de årene man har igjen å leve<sup>26</sup> og gir en indikasjon på effekten av ulike behandlinger. Standard QALYs er ment å utrykke personlig verdsetting av en tenkt helsetilstand vurdert ex ante (før hendelsen skjer)<sup>27</sup>. Verdier på ulike grader av mental, fysisk og sosial helse skal være preferansebaserte og er vanligvis fremkommet ved at friske personer har vektet ulike hypotetiske helsetilstander. QALYs kan dermed gi en indikasjon på hvilke helsegevinster man oppnår hos pasienter i form av kvalitet og livslengde. Skalaen man bruker går fra 0 (indikerer død) til 1 (indikerer perfekt helse). Enkelte instrumenter genererer også negative verdier, at helsetilstander kan oppfattes verre enn død. QALY's blir brukt både når det gjelder prioriteringer i helsevesenet og som effektmål i kliniske studier. 1 QALY = 1 år med perfekt helse. Når det gjelder ressursprioriteringer, sammenligner man kostnad pr vunnet leveår og effekt av de ulike behandlingene for å identifisere de mest kostnadseffektive behandlingalternativene<sup>28</sup>.

## 2.7 Verdsetting av ulike helsetilstander

De generiske livskvalitetsskjemaene er ment å kunne fange opp alle sider ved helse og skal kunne brukes på alle pasientgrupper på tvers av diagnoser. De aller fleste skjemaene er preferansebaserte (basert på svar fra folk flest), dvs at personer velger preferanser på grunnlag av hypotetiske helsetilstander. Det er flere faktorer som må tas hensyn til når man skal verdsette helse; hvordan skal helse beskrives, hvordan skal den verdsettes og hvem skal verdsettingen samles inn fra (generell befolkning, pasienter (som har erfaring med å leve med en gitt helsetilstand), helsepersonell osv)<sup>29</sup>. Det er flere metoder man kan anvende når man skal be personer tallfeste helserelatert livskvalitet og generere tariffer. De tre vanligste er visuell analog skala (VAS), time-trade-off (TTO) og standard gamble (SG)<sup>24</sup>. VAS blir betraktet som den enkleste for respondentene å angi, man ber respondenten om å tallfeste en gitt helsetilstand på en skala. Ved de to andre metodene ber man respondenten velge mellom to alternative helsetilstander. I TTO blir respondenten stilt overfor to valg og må velge mellom et langt liv en i en gitt helsetilstand eller et kortere liv i best tenkelige helsetilstand. Kritikken mot denne metoden er at den gir respondenten to sikre valg, mens realiteten er at utfallet av behandling ofte er usikkert<sup>29</sup>. Dette blir ivaretatt i SG, hvor respondenten blir gitt



et hypotetisk valg mellom et liv i en redusert, men sikker helsetilstand eller et ”lotteri” med to utfall; død eller frisk. Det er med andre ord essensielle forskjeller på disse tre metodene; med VAS ofrer man ingenting, med TTO må man ofre gjenværende levetid for å få perfekt helse og med SG er det stor usikkerhet og man er villig til å ofre livet. En hypotese fra litteraturen er at utility-verdiene som fremkommer ved bruk av SG er høyere enn de som benytter TTO når det gjelder alvorlige tilstander, mens det omvendte kan være tilfelle ved mindre alvorlige helsetilstander <sup>30</sup>.

## **2.8 Preferansebaserte livskvalitetsskjemaer til bruk i kostnad-nytte analyser**

Det finnes minst 10 internasjonale generiske preferansebaserte livskvalitetsskjemaer som man kan kalkulere QALY's på grunnlag av, de mest brukte er, EQ-5D, 15 D og SF 6D. De generiske livskvalitetsskjemaene er ment å fange opp ”alle” sider ved helse og gjøre det mulig å sammenligne helseforbedringer på tvers av diagnosegrupper <sup>24</sup>. De har til hensikt å måle det samme konstrukt, men bruker ulike modeller for helse og ikke minst ulike metoder for å utlede preferanser og ulike skåringssystemer <sup>31</sup>. Preferansene kartlegger grupper av friske personers ”betalingsvilje” for ulike grader av helse verdsatt i leveår og/eller penger. For å gjøre det mulig å bruke livskvalitetsskjemaene i kostnad-nytte analyser, kan det uti fra alle skjemaene kalkuleres en enkelt indexverdi (utility index) som reflekterer respondentens preferanser for ulike helsetilstander <sup>32</sup>. Indexverdien representerer med andre ord en relativ verdi på det å leve med en gitt helsetilstand (fra 0 som indikerer død til 1 som indikerer perfekt helse). Det har i imidlertid vist seg at det bare er dårlig/moderat samsvar mellom verdier kalkulert på grunnlag av de ulike livskvalitetsskjemaene og at indexverdiene fra skjemaene, f.eks EQ-5D og SF-6D, ikke er direkte sammenlignbare <sup>32</sup>, det samme gjelder EQ-5D versus 15D <sup>33, 34</sup>.

### **2.8.1 EQ-5D**

EQ-5D er utviklet av en internasjonal gruppe forskere, the EuroQol Group og er et av de mest brukte helserelaterte livskvalitetsskjemaet i kostnad-nytte analyser. En norsk versjon kan fås ved henvendelse til EuroQol gruppen. Skjemaet er validert for ulike pasientgrupper, er lett å forstå og raskt for pasienten å fylle ut (3-5 minutter). Respondentene angir sine problemer når det gjelder 5 ulike domener; gange, personlig stell, vanlige gjøremål, smerte/ubehag og

angst/depresjon. Tidligere ble respondentene gitt 3 svaralternativer på hvert spørsmål (EQ-5D-3L), nå er dette utvidet til 5 (ingen problemer, litt problemer, middels store problemer, store problemer, ute av stand til) (EQ-5D-5L) og dette kan gi i alt  $5^5$ , 3125 unike helsetilstander. Jansen og medarbeidere<sup>35</sup> sammenlignet de to versjonene av EQ-5D og konkluderte med at EQ-5D-5L hadde bedre måleegenskaper (reliabilitet og validitet), bedre evne til å diskriminere mellom ulike nivåer av helse og at derfor takeffekten (dvs at en test har en øvre begrensning som gjør at den ikke kan oppfatte forskjeller over en viss verdi, alle blir gitt maks verdi) var betraktelig redusert.

Grad av helse blir angitt på en skala fra 0 (død) til 1 (perfekt helse), men verdsetting lavere enn 0 (verre enn død) ned til  $-0,594$  kan forekomme, det finnes heller ingen verdier mellom 0,9 og 1. For EQ-5D-3L finnes preferansebaserte "trade-off-verdier" fra individer fra normalbefolkningen, for EQ-5D-5L etableres slike verdier i pågående studier<sup>36</sup>. Det er også laget en "cross-walk-fil", slik at man kan konvertere EQ-5D-5L til EQ-5D-3L. I tillegg til å angi problemene i de 5 domene, skal pasienten angi "Helsen din i dag" på en VAS-skala, EQ-VAS.

### **2.8.2 15D**

15D er et generisk livskvalitetsskjema utviklet i Finland og inkluderer 15 ulike domener; bevegelighet, syn, hørsel, pust, søvn, spising, tale, vannlating/avføring, vanlige aktiviteter, mental funksjon, ubehag og symptomer, depresjon, stress, livskraft og seksuell aktivitet. Hvert spørsmål har 5 svaralternativer dvs  $5^5$ , 30.518 millioner mulige unike helsetilstander. Skjemaet tar 5-10 minutter å fylle ut og er ment å dekke alle viktige helsesrelaterte domener. 15D gir både en helseprofil på bakgrunn av de 15 domene og en enkel indexskår<sup>23</sup>. Den laveste skåren er 0,11 (verst mulig helse) til 1 (perfekt helse). Det finnes preferansebaserte verdsettelse både fra Finland, Danmark og Hellas<sup>33</sup>. 15D er foreløpig brukt i et begrenset antall studier (primært finske), men et nylig utført søk i Pubmed viser at skjemaet stadig tas i bruk i flere studier, særlig europeiske.

### **2.8.3 SF-6D (SF-36)**

SF-6D er generisk preferansebasert utility-instrument, som består av 11 enkeltspørsmål fra SF-36, et internasjonalt svært mye brukt helsesrelatert livskvalitetsskjema<sup>37</sup>. Spørsmålene har senere blitt konvertert til et system for klassifisering av helse med 6 domener. Grad av helse

blir angitt på en skala fra 0,301 (verst mulig helse) til 1 (perfekt helse). Det finnes tariffer basert både på amerikanske og engelske populasjoner, den nyeste amerikanske er fra 2009. SF-36/SF-6D er lisensbelagt og koster mye i bruk, men det verserer også gratisversjoner og syntax-filer man kan få tilgang til. Skjemaet har forholdsvis lange svaralternativer som en del pasienter har vanskelig for å forstå og problemer med å svare på (Jennifer Davis, personlig meddelelse).

Tabell 1. Utvalgte karakteristika ved tre helserelaterte livskvalitetsskjema, EQ-5D, 15D og SF-6D

	Domener	Antall domener	Antall nivåer	Antall unike helsetilstander	Verdsettingsmetode
<b>EQ-5D-5L</b>	gange, personlig stell, vanlige gjøremål, smerte/ubehag og angst/depresjon	5	5	3125	Time tradeoff (TTO)
<b>15D</b>	bevegelighet, syn, hørsel, pust, søvn, spising, tale, vannlating/avføring, vanlige aktiviteter, mental funksjon, ubehag og symptomer, depresjon, stress, livskraft og seksuell aktivitet	15	5	>30 mill	Visuell analog skala (VAS)
<b>SF-6D (SF-36)</b>	fysisk og sosial funksjon, roller, smerte, mental helse og vitalitet	6	4-6	18000	Standard gamble (SG)

## **2.9 Hvilke(t) helserelatert livskvalitetsskjema skal man bruke i studier som evaluerer effekt av fysioterapi?**

Selvrapportert helserelatert livskvalitet har de siste årene blitt en viktig del av måleinstrumentene som inkluderes i kliniske studier. Skjemaene gir viktig informasjon om pasientens egen vurdering av helse og funksjon og er dessuten nødvendige for å kunne beregne kostnader og nytte av ulike behandlingsformer.

Det er flere faktorer som avgjør hvilke(t) livskvalitetsinstrument man bør velge/inkludere i en studieprotokoll; hvor aktuelle spørsmålene er for den aktuelle pasientpopulasjonen, hvor omfattende og tidkrevende utfyllingen er (antall spørsmål), hvilke instrumenter som er brukt i lignende studier, hvilke kostnader som er knyttet til evt lisenser og hvilke metoder som er brukt for å kalkulere de preferansebaserte verdiene osv. SF-6D (SF-36) er det skjema som er mest brukt internasjonalt, men EQ-5D er nå inkludert i flere studier som evaluerer effekt av fysioterapi<sup>38-41</sup>. 15D har også i det siste blitt inkludert i studier utenfor Finland. Flere studier viser imidlertid at de ulike helserelaterte livskvalitetsskjemaene ikke kan brukes om hverandre, at skjemaene inneholder svært ulike aspekter ved helse, at skårene som genereres har benyttet ulike verdsettingsteknikker, at de ikke er direkte sammenlignbare og gir ulike utility-indexer for samme pasientgruppe<sup>29, 32, 33, 42-44</sup>. Det blir derfor svært vanskelig å sammenligne økonomiske evalueringer fra ulike studier.

Enkelte publiserte studier indikerer at SF-6D (SF-36) beskriver alvorlige helsetilstander bedre enn EQ-5D<sup>42</sup>, mens andre mener at EQ-5D kanskje er mer sensitivt for helseforbedring hos pasienter med relativt dårlig helse ved behandlingsstart og at SF-6D (SF-36) er mer egnet for å fange opp helseforbedringer hos pasienter med relativt små/milde plager<sup>30, 45</sup>.

Jeg har i dette prosjektet valgt å inkludere alle de tre mest brukte livskvalitetsskjemaene, noe som forhåpentligvis kan gi en indikasjon på hvilke(t) som egner seg best for å evaluere effekt og kostnader etter fysioterapibehandling.

## **2.10 Responsivitet**

Å evaluere hvordan pasientens helse endres under og/ eller etter behandling er viktig i klinisk praksis og måleinstrumentenes evne til å fange opp endring blir derfor svært viktig.

Responsivitet er definert som "the ability of an instrument to detect change over time in the

construct to be measured”<sup>46</sup>. Responsivitet har med andre ord fokus på et måleinstruments evne til å fange opp endring og man må ha minst to longitudinelle målinger for å kalkulere et måleinstruments responsivitet. En slik metodestudie må ha et tidsperspektiv hvor man kan forvente at pasientene (i alle fall noen) vil vise en endring til det bedre eller til det verre. Man ønsker med andre ord instrumenter som har evnen til å fange opp en reell endring. For å teste dette, må instrumentet evalueres opp mot en gullstandard, noe som ofte ikke finnes. Et alternativ er da å benytte en ”global rating scale (GRS)”, et endringsspørsmål, som gullstandard for måling av endring. Ved avsluttet behandling blir pasienten forelagt *et* enkelt spørsmål om hvordan de oppfatter en eventuell endring fra før til etter behandling, ofte gradert på en skala fra ”mye verre” til ”mye bedre”. Det er imidlertid viktig at spørsmålet stilles på en slik måte at det eksplisitt innbefatter det samme tema som instrumentet som skal evalueres (for eksempel smerte eller funksjon). Samsvaret mellom endring målt med de to instrumentene blir deretter testet statistisk.

## **2.11 Kostnad-nytte analyser av fysioterapi i primærhelsetjenesten**

Søk i Pubmed med ordene ”Physiotherapy and cost-utility” gav 66 treff, 35 av artiklene var publisert de siste 5 årene. Flere av artiklene var, av ulike grunner, lite aktuelle; de presenterte kun protokoller/design for nystartete studier, evaluerte primært kirurgi og ikke fysioterapi, eller var publikasjoner med ulike vinklinger fra en og samme studie.

Det finnes mange randomiserte studier som evaluerer effekt av ulike behandlingsalternativer for ulike pasientgrupper i fysioterapi, men det finnes få av disse studiene som har inkludert økonomiske analyser for å vurdere kostnadseffektiviteten av de ulike behandlingsformene. Davis og medarbeidere<sup>39</sup> sammenlignet tre ulike treningsprogrammers (A. styrketrening x 1 pr uke, B. styrketrening x 2 pr uke og C. balansetrening x 2 pr uke) effekt på forebygging av fall hos eldre kvinner. De konkluderte med at styrketrening 1 og 2 ganger pr uke ga ”mest helse for pengene” (de totale utgiftene var mindre) når det gjaldt forebygging av fall i den eldre befolkningen.

Pinto og medarbeidere<sup>47</sup> gjorde en økonomisk evaluering i en randomisert studie hvor de sammenlignet: A. øvelsesbehandling/trening, B. manuell terapi, C. en kombinasjon av A og B

og D. "usual care". De konkluderte med at både øvelsesbehandling og manuell terapi var mer kostnadseffektivt enn "usual care" eller de to behandlingsformene i kombinasjon.

## 3 Materiale og metode

### 3.1 Materiale

Aktuelle pasienter ble rekruttert av 5 fysioterapeuter i primærhelsetjenesten. I utgangspunktet kunne alle nye pasienter som kom til terapeuten bli spurt om å delta i studien, de eneste kriteriene for inklusjon var: Pasienten skulle være i arbeidsfør alder (18-67 år) og ha god norsk språkforståelse både muntlig og skriftlig.

### 3.2 Metode

Pasienten fikk ved første konsultasjon en muntlig forespørsel om å delta i studien og fikk ved positiv respons informasjonsskriv og samtykkeskjema (vedlegg 1) med mer detaljert informasjon om studien med hjem, for å kunne ta standpunkt til deltakelse på et bedre grunnlag. Etter utfylling av skjemaene sjekket fysioterapeuten at alle delspørsmålene var fylt ut.

Ved andre konsultasjon fylte fysioterapeuten ut et registreringsskjema for alle pasientene som ønsket å delta i studien. Skjemaet inneholdt demografiske data som kjønn, alder, høyde, vekt, diagnose, aktuelt problem, yrke/arbeidsstatus osv. Pasienten fylte ut flere generiske livskvalitetsskjemaer. De samme skjemaene ble fylt ut både før behandlingsstart og ved behandlingsslutt. Skjemaene var:

EQ-5D - 5 ulike domener; ”gange”, ”personlig stell”, ”vanlige gjøremål”, ”smerte/ubehag” og ”angst/depresjon” med 5 svaralternativer. Pasienten skal i tillegg angi ”Helsen din i dag” på en VAS-skala (se mer detaljert beskrivelse i under pkt 2.7.1). En norsk versjon og syntaxfil for beregning av utility index finner man på [www.euroqol.org](http://www.euroqol.org).

15D - inkluderer 15 ulike domener; ”bevegelighet”, ”syn”, ”hørsel”, ”pust”, ”søvn”, ”spising”, ”tale”, ”vannlating/avføring”, ”vanlige aktiviteter”, ”mental funksjon”, ”ubehag og symptomer”, ”depresjon”, ”stress”, ”livskraft” og ”seksuell aktivitet”. (se mer detaljert beskrivelse i under pkt 2.7.2). Skjemaet får ved henvendelse til [harri.sintonen@helsinki.fi](mailto:harri.sintonen@helsinki.fi)

SF-6D (SF-36) – består av 11 enkeltspørsmål fra SF-36<sup>37</sup> som blir konvertert til et system for klassifisering av 6 helsedomener; ”fysisk funksjon”, ”rollebegrensninger”, ”sosial fungering”,

”smerte”, ”mental helse” og ”vitalitet” (se mer detaljert beskrivelse i under pkt 2.7.3). Skjemaet er lisensbelagt og norsk versjon og syntaxfiler kan kjøpes via [qualitymetric.com](http://qualitymetric.com).

For å registrere pasientens subjektive vurdering av endring svarte pasientene på to endringsspørsmål, hvordan de vurderte 1. sine plager og 2. sin funksjon nå i forhold til før behandlingen startet (”global rating scale”) på en skala med 7 kategorier fra ”mye verre/mye dårligere” til ”mye bedre”, (GRS). Disse to spørsmålene ble brukt som gullstandard når livskvalitetsskjemaenes responsivitet skulle estimeres.

Etter avsluttet behandling anga fysioterapeuten hvilke behandlingsformer som var gitt, hvilke takster som var benyttet og hvor mye behandlingen hadde kostet totalt (refusjonsbeløp + egenandeler).

### 3.3 Statistiske analyser

Demografiske data, global skala for plager og funksjon, poengskår og utility index for EQ-5D, 15D og SF-6D (SF-36) og kostnader ble kalkulert som gjennomsnitt  $\pm$  standard deviasjon (SD), endring som gjennomsnitt og 95 % konfidensinterval. For kategoriske data ble frekvens og min/maks kalkulert.

Utility indexene for de ulike livskvalitetsskjemaene ble kalkulert ved hjelp av syntaxfiler mottatt fra de ulike organisasjonene som administrerer skjemaene internasjonalt, EQ-5D fra [www.euroqol.org](http://www.euroqol.org), 15D ble sendt til Finland og kalkulert av Harri Sintonen og utility indexen for SF36 ble kalkulert ved hjelp av syntaxfil tilsendt fra University of Sheffield (<http://www.sheffield.ac.uk/scharr/>).

For å analysere endringer i frekvens av svaralternativer for hvert enkeltomene i de ulike livskvalitetsskjemaene fra før til etter behandling, ble Wilcoxon Signed Rank Test benyttet.

For å analysere responsivitet ble de to endringsspørsmålene brukt som gullstandard. Samsvaret mellom de ulike livskvalitetsskjemaene og endringsspørsmålene ble først kalkulert med Spearman’s rho korrelasjon, dernest ble gjennomsnittlig endring i utility index delt inn etter de ulike kategoriene i endringsspørsmålene. Til slutt ble effektstørrelsen (ES) beregnet ( $\text{Eta squared} = t^2 / (t^2 + (N-1))$ ), hvor t-verdien er et mål på om forskjellen i skår mellom de to målingene er reell eller skyldes en tilfeldighet og N = antall observasjoner. Eta squared er en



av de vanligste statistiske metodene for beregning effektstørrelser, hvor ,01= liten effekt, ,06= moderat effekt og ,14= stor effekt).<sup>48</sup>

Samsvaret (agreement) mellom utility indexene kalkulert fra de ulike livskvalitetsskjemaene (fylt ut før behandlingsstart) ble evaluert ved å konstruere Bland Altman plot<sup>49</sup>. Plottet identifiserer forholdet mellom differansen mellom to målinger (measurement error) og det beste estimatet for den sanne verdien, gjennomsnittet av de to målingene<sup>32</sup>. Plottet viser også alle ekstreme uteliggere (outlying observations)<sup>49, 50</sup>.

Korrelasjonen mellom endringsskårene for EQ-5D, 15D og SF-6D ble analysert ved Spearman rank korrelasjon koeffisient (r).

For å beregne kostnader pr vunnet QALY, ble de totale behandlingsutgiftene dividert med endring i QALYs (fra før til etter behandling). Det ble antatt at effekten ville vedvare 1 år. Forholdet mellom endring i utility index og kostnader ble også fremstilt visuelt ved hjelp av et scatter plot.

## 4 Resultater

I alt 27 pasienter ble rekruttert til studien i løpet av en 2-måneders periode. Det var 19 kvinner (70 %) og 8 menn (30 %), gjennomsnittsalderen var  $53,8 \pm 16,5$  år og de hadde en gjennomsnittlig kroppsmasse index (KMI) på  $26,5 \pm 5,7$ . De fleste pasientene hadde kroniske muskel- skjellett lidelser, 42 % av pasientene hadde nakke/skulder problemer, 23 % hadde artrose i hofta eller kne og problemet hadde vart gjennomsnittlig i  $78,5 \pm 178$  (3-780) uker.

14 (52 %) av pasientene var i arbeid, 2 var studenter og de resterende gikk enten på attføring, uføretrygd eller var pensjonister. Av dem som var i arbeid var 7 (50 %) sykemeldt ved behandlingsstart, to av disse var tilbake i arbeid ved behandlingsslutt.

Alle pasientene hadde fått individuell øvelsesbehandling, medisinsk treningsterapi og/eller gruppetrening, 1 pasient hadde i tillegg fått varme/kuldepakninger og 3 hadde fått elektroterapi. 78 % av pasientene drev egen trening i tillegg til behandlingen.

Behandlingen varte i  $12,5 \pm 3,2$  uker og pasientene mottok i gjennomsnitt 14 ( $\pm 7$ ) behandlinger. 16 pasienter (59 %) hadde krav på full refusjon av behandlingsutgiftene. De gjennomsnittlig totale behandlingsutgiftene var kr  $3338 \pm 1310$ , gjennomsnittlig refusjon fra trygden kr.  $2664 \pm 1658$  og egenandel kr  $674 \pm 1018$ .

Det var en signifikant positiv endring fra før til etter behandling i fordeling av svar for alle de 5 domeneene i QE-5D, Gange ( $p=0,003$ ), Personlig stell ( $p=0,025$ ), Vanlige gjøremål ( $p=0,005$ ), Smerter/ubehag ( $p=0,002$ ) og Angst depresjon ( $p=0,05$ ). For 15D var det en endring i 7 av de 15 domeneene, Beveglighet ( $p=0,020$ ), Pust ( $p=0,003$ ), Vanlige aktiviteter ( $p=0,008$ ), Ubehag og symptomer ( $p=0,001$ ), Stress ( $p=0,005$ ), Livskraft ( $p=0,008$ ) og Seksuell aktivitet ( $p=0,008$ ).

For SF-36 var det en signifikant endring for ca 2/3 av domeneene fra før til etter behandling, men de blir ikke presentert her, da bare noen av dem inngår i SF-6D.

Tabell 2. Helsereelatert livskvalitet uttrykt som utility index før og etter fysioterapibehandling (n=27)

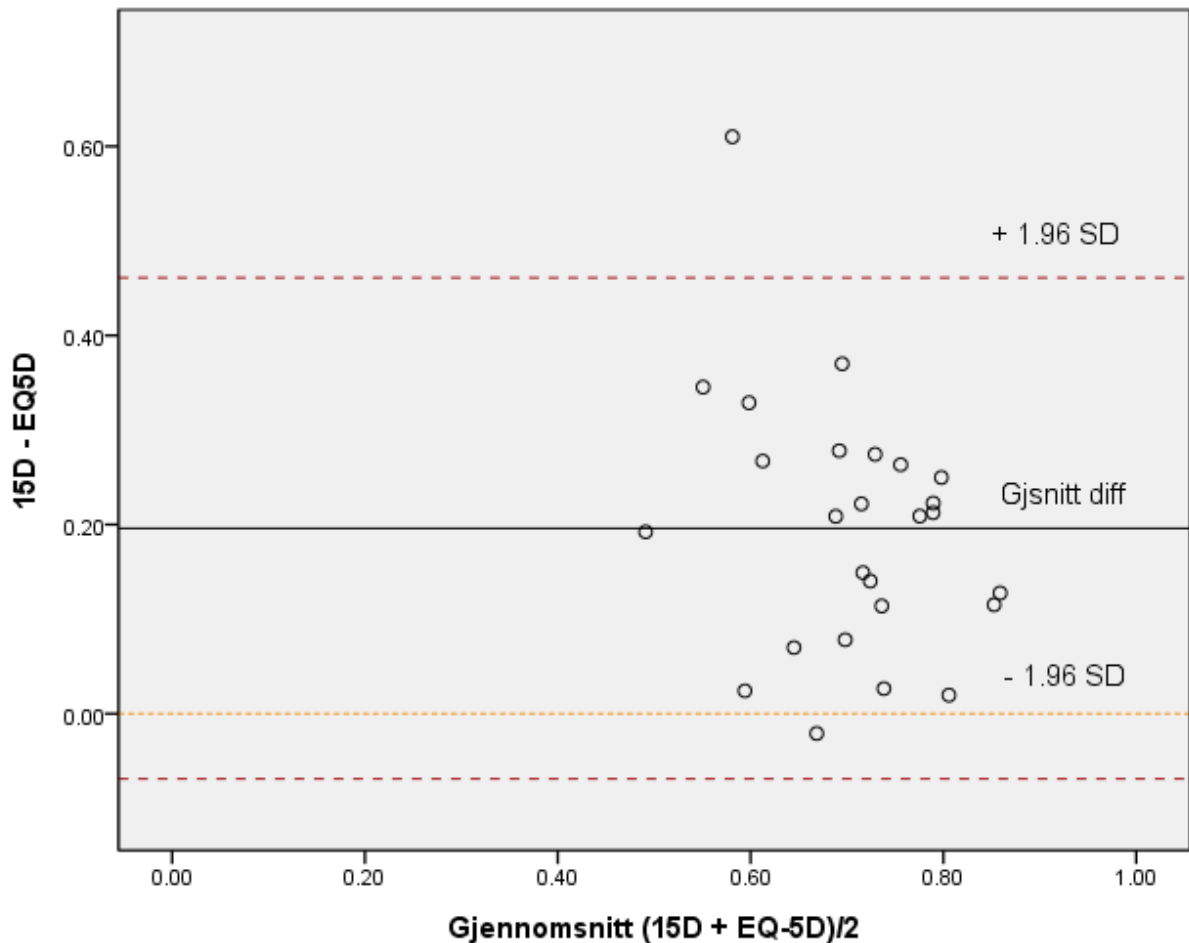
	Før behandling	Etter behandling	Differanse (95 % KI)	p-verdi	ES
<b>EQ-5D</b>	0,605 ±0,130	0,733 ±0,102	0,128 (0,059 – 0,196)	0,001	0,34
<b>15D</b>	0,804 ±0,095	0,866 ±0,092	0,061 (0,037 – 0,085)	< 0,001	0,48
<b>SF-6D</b>	0,629 ±0,073	0,639 ±0,207	0,010 (-0,069 – 0,089)	0,792	0,002

ES = Effekt størrelse, KI = Konfidensintervall

Tabell 2 viser utility index før og etter behandling for de tre helsereelaterte livskvalitetsskjemaene. Det var en høysignifikant positiv endring fra før til etter behandling både for EQ-5D og 15D, høyest for EQ-5D. Endringen for SF-6D var minimal og ikke-signifikant. Effektstørrelsen var stor (>0.14) både for 15D og EQ-5D, henholdsvis 0,48 og 0,34, mens effektstørrelsen for SF-6D var svært liten (<0.01).

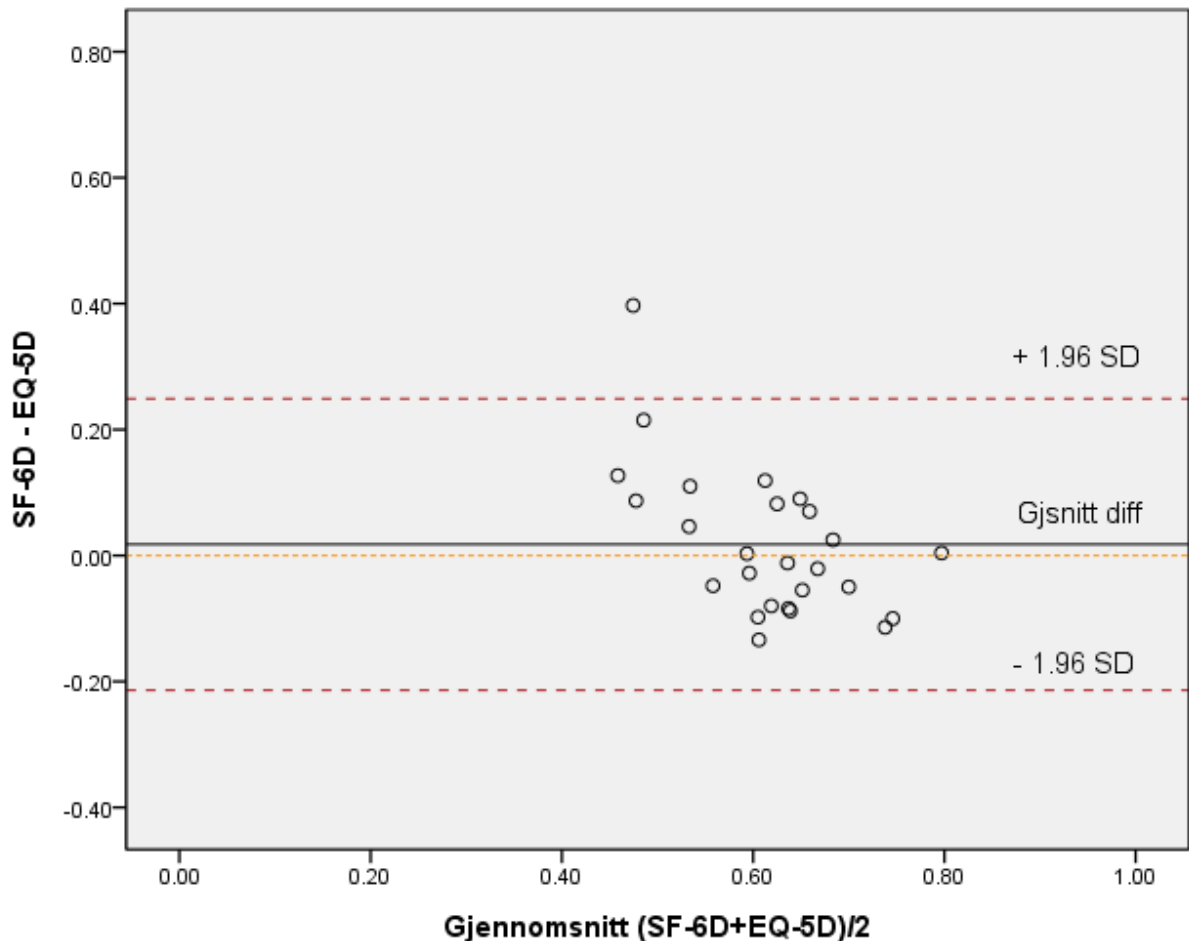
Korrelasjonen mellom de ulike utility indexene ved baseline var EQ-5D – 15D  $r = 0,31$ , EQ-5D – SF-6D  $r = 0,43$  og 15D – SF-6D  $r = 0,53$ .

Figur 2. Bland-Altman plott av 15D og EQ-5D utility index ved behandlingsstart. Heltrukket linje representerer gjennomsnittlig differanse, rosa stiplede linjer Limits of Agreement (+1,96 SD og -1,96 SD av differansen) og oransje linje representerer differanse=0.



Figur 2 viser Bland-Altman plott av differansene mellom 15D og EQ-5D utility index ved behandlingsstart. 95 % Limits of Agreement (LOA) var -0,069 og 0,461, og gjennomsnittlig differanse mellom 15D og EQ-5D ved baseline var  $0,196 \pm 0,135$ , noe som indikerer dårlig enighet mellom skjemaene. Alle pasientene, bortsett fra en, hadde en høyere utility-index målt med 15D enn EQ-5D, gjennomsnittlig skår var henholdsvis  $0,805 \pm 0,095$  og  $0,603 \pm 0,132$  (Tabell 2). Kun en pasient lå utenfor LOA (<5 %).

Figur 3. Bland-Altman plott av SF-6D og EQ-5D utility index ved behandlingsstart. Heltrukket linje representerer gjennomsnittlig differanse, rosa stiplede linjer Limits of Agreement (+1,96 SD og -1,96 SD av differansen) og oransje linje representerer differanse=0.



Figur 3 viser Bland-Altman plott av differansene mellom SF-6D og EQ-5D utility index ved behandlingsstart. 95 % LOA var -0,153 og 0,249, og gjennomsnittlig differanse mellom SF-6D og EQ-5D ved baseline var  $0,018 \pm 118$ , med andre ord ganske nær null (indikert ved oransje stiplet linje), noe som indikerer god enighet. 14 pasienter scoret høyere og 11 skåret lavere på SF-6D enn på EQ-5D, mens to pasienter hadde identisk utility index målt med begge instrumentene. Kun en pasient lå utenfor LOA (<5 %).

Tabell 3 Korrelasjon mellom endring i utility index for de tre livskvalitetsskjemaene EQ-5D, 15D og SF-6D

	EQ VAS	EQ-5D	15D	SF-6D
EQ VAS		0,418	0,452	0,161
EQ-5D	0,418		0,719	0,106
15D	0,452	0,719		0,198
SF-6D	0,161	0,106	0,198	

Tabell 3 viser korrelasjonen mellom endringen i utility index for de tre livskvalitetsskjemaene. Vi ser at SF-6D har lav korrelasjon med begge de andre skjemaene, mens korrelasjonen mellom EQ-5D og 15D er god,  $r = 0,72$ .

Korrelasjonen mellom endring i utility-index og den globale skalaen for vurderingen av plager og funksjon, var henholdsvis  $r = 0,44$  og  $r = 0,43$  for EQ-5D,  $r = 0,44$  og  $r = 0,58$  for 15D og  $r = 0,53$  og  $r = 0,45$  for SF-6D.

Figur 4. Pasientenes angivelse av endring i plager og funksjon etter behandlingsslutt

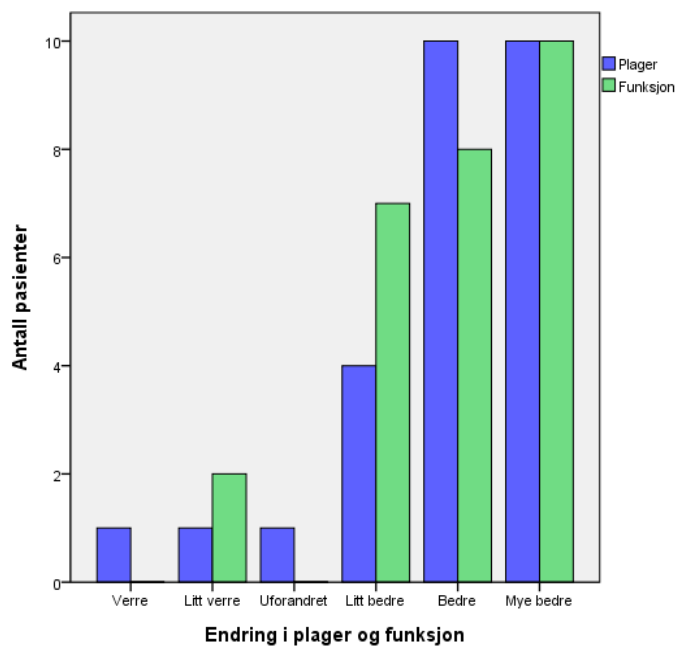


Fig 4 viser pasientenes egen vurdering av plager og funksjon nå (siste behandlingsdag) i forhold til før behandlingen startet. Når den globale 7 punkt skåren ble dikotomisert i ”uforandret/verre” og ”bedre”, oppga 24 (89 %) og 25 (92.6 %) av pasientene bedring i henholdsvis plager og funksjon.

Tabell 4. Gjennomsnittlig endring for de ulike helserelaterte livskvalitetsskjemaene gruppert i forhold til endring i funksjon angitt på en skala fra 1 til 7 (ingen pasienter angav å være uforandret, dårligere eller mye dårligere).

	Litt dårligere (n= 2)	Litt bedre (n= 7)	Bedre (n= 8)	Mye bedre (n= 10)
<b>EQ-5D</b>	-0,057 ±0,039	0,086 ±0,117	0,103 ±0,112	0,211 ±0,210
<b>15D</b>	0,015 ±0,041	0,029 ±0,044	0,039 ±0,046	0,109 ±0,057
<b>SF-6D</b>	0,001 ±0,120	-0,068 ±0,213	0,030 ±0,089	0,051 ±0,267

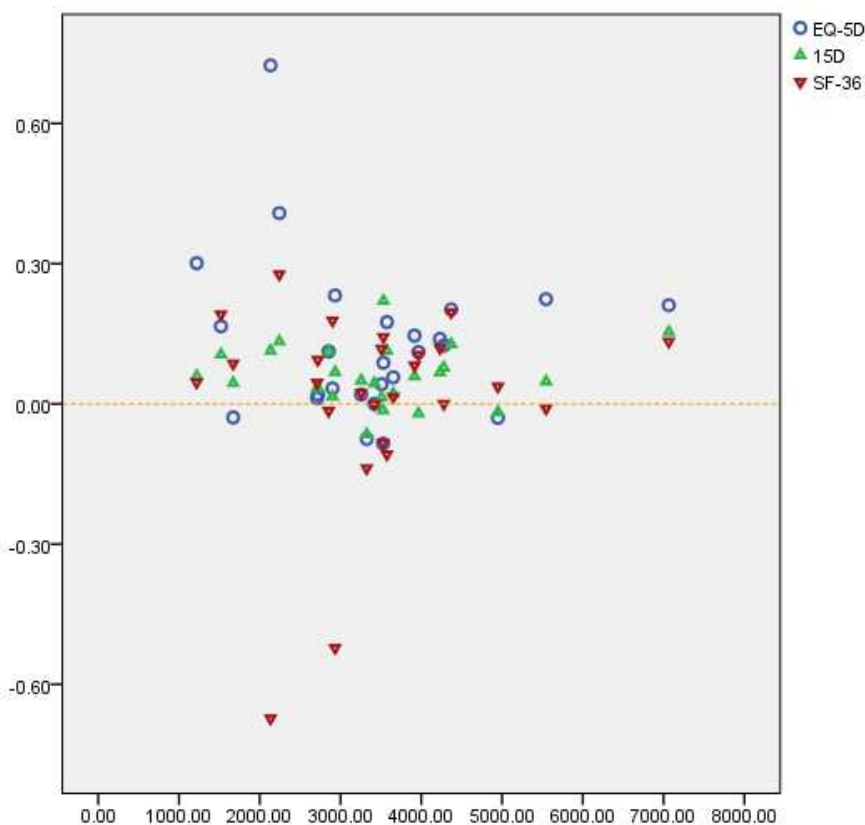
Tabell 4 viser endring i utility index målt med de tre skjemaene gruppert i forhold til endringsspørsmålet ”Sammenliknet med før du startet behandlingen, hvordan vurderer du funksjonen din nå?”. Tabellen viser en gradvis økende positiv endring i utility index etter som skalaen går mot bedre funksjon for alle tre skjemaene, størst endring målt ved EQ-5D og minst endring målt ved SF-6D, dvs at EQ-5D hadde best evne til å skille mellom gruppene som viste små og større endringer.

Tabell 5. Gjennomsnittlig endring for de ulike helserelaterte livskvalitetsskjemaene gruppert i forhold til endring i plager angitt på en skala fra 1 til 7 (kun 1 pasienter angav å være uforandret, litt verre eller verre og er derfor ikke tatt med i tabellen).

	Litt bedre (n= 4)	Bedre (n= 10)	Mye bedre (n= 10)
<b>EQ-5D</b>	0,075 ±0,138	0,109 ±0,111	0,210 ±0,210
<b>15D</b>	0,026 ±0,061	0,052 ±0,033	0,097 ±0,073
<b>SF-6D</b>	-0,144 ±0,268	0,019 ±0,067	0,065 ±0,270

Tabell 5 viser endring i utility index målt med de tre skjemaene gruppert i forhold til endringsspørsmålet ”Sammenliknet med før du startet behandlingen, hvordan vurderer du plagene dine nå?”. Tabellen viser en gradvis økende positiv endring i utility index etter som skalaen går mot bedre funksjon for alle tre skjemaene, størst endring målt ved EQ-5D og minst endring målt ved SF-6D, dvs at EQ-5D hadde best evne til å skille mellom gruppene som viste små og større endringer.

Figur 5. Scatterplot av behandlingsutgifter (x-aksen) i forhold til endring i livskvalitet (y-aksen). Stiplet linje markerer ingen effekt



Figur 5 viser forholdet mellom behandlingskostnader (x-aksen) og nytte (endring i utility index) (y-aksen) kalkulert på basis av de tre helse relaterte livskvalitetsskjemaene. Plottet viser at de fleste av pasientene har hatt nytte av behandlingen (ligger over den stiplede linjen) og at behandlingskostnadene for de fleste ligger mellom 2000,- og 4500,- kroner.

Hvis man kalkulerer kostnader pr vunnet QALY basert på de tre livskvalitetsskjemaene (gjennomsnittlig kostnad / gjennomsnittlig endring i utility index) varierer disse ganske mye, fra henholdsvis kr. 26087,- for EQ-5D, kr. 54721,- for 15D til kr. 333800,- for SF-6D.



## 5 Diskusjon

Hensikten med dette prosjektet var å prøve ut de tre mest brukte helserelaterte livskvalitetsskjemaene, EQ-5D, 15D og SF-6D, som evalueringsverktøy i klinisk fysioterapi praksis, evaluere om skjemaene kunne fange opp endring og eventuelt beregne ”kost-nytte” for behandlingen som pasientene mottok. Resultatene viste at ca 90 % av pasientene rapporterte en bedring både når det gjaldt plager og funksjon (Figur 4). To av livskvalitetsskjemaene, EQ-5D og 15D, viste en høysignifikant endring i utility index og en stor effektstørrelse fra før til etter behandling (Tabell 2). Målt med SF-6D var det ingen endring. Kostnadene for behandlingsserien var i gjennomsnitt ca kr 3500,-.

### 5.1 Utility index ved baseline

Tabell 2 viser utility indexene for de tre livskvalitetsskjemaene før og etter behandling. Vi ser at før behandling var det godt samsvar mellom EQ-5D og SF-6D hvor de gjennomsnittlige utility indexene var henholdsvis 0,605 og 0,629. Dette gjenspeiler seg også i Figur 3, differansen målt med de to skjemaene hadde et gjennomsnitt ved behandlingsstart på 0,018. Kun 1 pasient lå utenfor LOA. Figuren viser ingen klar tendens når det gjelder hvilket av de to skjemaene som estimerte de høyeste, respektive laveste indexene, det var omtrent like mange pasienter som skåret høyest på EQ-5D (under oransje stiplet linje) som på SF-6D (over oransje stiplet linje). Når det gjaldt 15D, lå utility indexen ved behandlingsstart høyere enn for begge de to andre skjemaene (Tabell 2). Dette gjenspeiles også i Figur 2, den gjennomsnittlige differansen mellom 15D og EQ-5D (heltrukket linje) ligger betraktelig over 0 og alle observasjoner, bortsett fra en, viste høyere differanser for 15D enn EQ-5D.

Forholdsvis mange studier som inkluderer pasienter med muskelskjelettplager har benyttet EQ-5D og/eller SF-6D for beregning av kost-nytte. I studier som har brukt EQ-5D, varierer utility indexen før intervensjon fra 0,46 til 0,60 hos pasienter med alvorlige kroniske lidelser (RA <sup>51</sup> og artrose <sup>52</sup>), til mellom 0,67 og 0,81 hos pasienter med milde til moderate muskelskjelett lidelser som nakke- og ryggproblemer og patellofemoralt smertesyndrom <sup>38, 53, 54</sup>. Gjennomsnittlig index før behandling var i denne studien  $0,605 \pm 130$ , noe som stemmer godt overens med tall fra litteraturen som inkluderer de samme kategoriene pasienter.

Sach og medarbeidere <sup>55</sup> inkluderte både EQ-5D og SF-6D i en studie hvor de evaluerte effekten av ulike behandlingsformer for pasienter med kroniske knesmerter. De fant ved baseline en utility index på 0,56 og 0,65 målt med henholdsvis EQ-5D og SF-6D. Johnsen og medarbeidere <sup>42</sup> fant store forskjeller i utility index målt med EQ-5D og SF-6D på pasienter med kroniske ryggplager (smerter > 6 år), hvor indexene var henholdsvis 0,29 og 0,55. Grunnen til at begge disse indexene ligger betraktelig lavere enn verdiene målt i dette prosjektet (Tabell 2), kan være at pasientene i Johnsens studie hadde hatt sine plager over en enda lenger tidsperiode. Vi ser dessuten at verdiene målt med EQ-5D ligger betraktelig lavere enn SF-6D, noe som kan tyde på at det er dårligere samsvar mellom instrumentene jo alvorligere pasientene rapporterer sitt problem å være.

Moock og medarbeidere <sup>43</sup> sammenlignet flere livskvalitetsskjemaer fylt ut av 106 rehabiliteringspasienter med muskelskjelett lidelser og fant ved baseline en utility index på 0,63 for EQ-5D, 0,84 for 15D og 0,62 for SF-6D. Dette er helt i tråd med indexene ved baseline i dette prosjektet (Tabell 2), hvor 15D ga betraktelig høyere verdier enn de to andre instrumentene. Dette gjenspeiler seg også i en studie av Hirvonen og medarbeidere <sup>56</sup>, de sammenlignet artrosepasienter (preoperativt) med friske kontroller og fant en utility index på henholdsvis 0,778 og 0,883 hos de to gruppene målt med 15D.

## 5.2 Endring i utility index

En måte å avgjøre om en endring er klinisk relevant, er å beregne ”minste viktige endring” (minimal important change=MIC). Det finnes studier som har kalkulert dette for de tre livskvalitetsskjemaene som er inkludert i dette prosjektet, minste viktige endring i utility index er henholdsvis  $\pm 0,07$ ,  $\pm 0,03$  og  $\pm 0,04$  for EQ-5D <sup>57</sup>, 15D <sup>56</sup> og SF-6D <sup>57</sup>. Differansene oppgitt i Tabell 2 indikerer, uti fra verdiene nevnt over, at det har skjedd en klinisk relevant endring fra før til etter behandling, kalkulert både ved hjelp av EQ-5D og 15D. For SF-6D var differansen ubetydelig og langt under veiledende MIC. Dette viser med tydelighet at resultatene man får i en studie er avhengig av hvilke(t) måleinstrument man velger å inkludere.

Når det gjelder fysioterapistudier finnes det to studier som har inkludert EQ-5D som evalueringsverktøy. Cuesta-Vargas og medarbeidere <sup>58</sup> evaluerte effekten av et multimodalt fysioterapiprogram (øvelser/trening + informasjon) på helse relatert livskvalitet. Alle

pasientene hadde kroniske muskelskjelett plager og de mottok behandling to ganger i uka i 8 uker. Utility indexen ved baseline var 0,70, og ved behandlingsslutt 0,88, en endring på 0,18. Endringen er noe større enn i dette prosjektet, men begge studiene indikerer at et individuelt tilpasset treningsprogram over minst 8 uker bedrer pasientenes helserelaterte livskvalitet og at endringen er klinisk relevant.

Tan og medarbeidere<sup>54</sup> evaluerte i en RCT effekten av et standardisert øvelsesprogram gitt til unge pasienter med patellofemoralt smertesyndrom. Pasientene var hos fysioterapeut i alt 9 ganger i løpet av 6 uker og hadde i tillegg et egentreningsprogram. Kontrollgruppen fikk råd og veiledning av lege. Ved baseline var utility indexen (EQ-5D) i de to gruppene henholdsvis 0,819 og 0,807, etter 52 uker 0,897 og 0,881, en endring på henholdsvis 0,078 og 0,074. Resultatene viste at øvelser reduserte smertene betraktelig etter 6 og 13 uker, men at begge gruppene hadde en positiv og klinisk relevant endring (endring i utility index >0,07) etter 1 år.

To studier som evaluerer trening /rehabilitering har brukt 15D som helserelatert livskvalitetsmål<sup>59, 60</sup>. Nakkepasientene som var inkludert i Salo sin studie hadde en forholdsvis høy baseline score (0,904) målt med 15D, resultatene etter 12 måneder viste allikevel en signifikant endring hos intervensjonsgruppene (0,022), men altså ikke klinisk relevant hvis man skal se det i forhold til størrelsen på ”minste viktige endring” publisert av Hirvonen<sup>56</sup>, som fant en MIC på 0,04. I en finsk studie som evaluerte effekten av et multidisiplinært rehabiliteringsprogram for artrosepasienter som hadde fått satt inn kneprotese, viste pasientene en signifikant endring etter 12 måneder (0,034), og man så at dimensjonene bevegelse, vanlige aktiviteter, ubehag og symptomer fremdeles var forskjellig fra normalbefolkningen etter behandlingsslutt. Disse to studiene indikerer at 15D både kan fange opp endringer over tid og diskriminere mellom grupper, selv om pasientene har relativt god helserelatert livskvalitet ved baseline/måletidspunktet.

### **5.3 Kan livskvalitetsskjemaene brukes om hverandre?**

De forskjellige livskvalitetsskjemaene er ulike både i innhold og omfang og det er store metodiske forskjeller i måten preferansene er fremkommet på<sup>32</sup>. De har allikevel det til felles, at de til slutt verdsetter den aktuelle helsetilstanden pasientene rapporterer med en enkel utility index. Med andre ord en relativ verdi som samfunnet, basert på preferanser, setter på det å leve med en gitt helsetilstand, fra 0 som indikerer død til 1 som indikerer perfekt helse.

Hensikten med å kalkulere den samme, enkle indexen for alle skjemaene, var at de skulle kunne brukes om hverandre og at man skulle kunne sammenligne resultater på tvers av både diagnosegrupper og studier. Indexen inngår dessuten som en viktig faktor (sammen med totale kostnader) når man kalkulerer antall QALY's man eventuelt har vunnet ved å introdusere en ny behandling. Slike beregninger ligger ofte til grunn når myndighetene skal prioritere blant ulike diagnosegrupper og behandlingsalternativer. Det er imidlertid publisert en rekke studier som viser dårlig samsvar mellom utility indexene kalkulert på grunnlag av de ulike skjemaene<sup>32, 33, 42, 61</sup>, noe som vil få stor betydning for resultatet av de helseøkonomiske analysene som foretas.

Det finnes foreløpig kun et fåtall studier som har sammenlignet de to livskvalitetsskjemaene EQ-5D og 15D, dvs om skjemaene estimerer tilnærmet lik grad av helserelatert livskvalitet på de samme pasientene. Saarni og medarbeidere<sup>61</sup> publiserte i 2006 en studie hvor de undersøkte 29 ulike kroniske sykdommers innvirkning på pasientenes livskvalitet ved hjelp av både EQ-5D og 15D. De fant at EQ-5D var mindre sensitivt enn 15D for å påvise signifikante sammenhenger mellom diagnose og helserelatert livskvalitet. EQ-5D viste seg dessuten å være mindre sensitivt når det gjaldt å oppdage deviasjoner fra perfekt helse, 47 % av respondentene hadde perfekt helse målt med EQ-5D, mens kun 15 % hadde perfekt helse målt med 15D. Resultatene viste at ved bruk av EQ-5D hadde pasienter med artrose størst årlig livskvalitetstap, mens ved bruk av 15D var det største årlige livskvalitetstap ved depresjon, inkontinens og artrose. Lunde<sup>33</sup> sammenlignet utility indexene generert fra EQ-5D og 15D fylt ut av pasienter minst 6 måneder etter slag. Undersøkelsen ble sendt pr post og av dem som svarte (58.5 %), var det 82 % som hadde fylt ut hele 15D, det tilsvarende tallet for EQ-5D var 85 %. Den høye andelen pasienter som hadde fylt ut skjemaene komplett indikerte at begge er lett forståelige og enkle å svare på. 15D gav gjennomgående høyere skår enn EQ-5D, noe som gir 15D mindre rom for store forbedringer<sup>33</sup>. Studien konkluderer med at de to skjemaene ikke kan brukes om hverandre. Saarni<sup>61</sup> fant altså at 15D var mest sensitivt når det gjaldt å oppdage deviasjoner fra perfekt helse, mens Lunde<sup>33</sup> fant at 15D gjennomgående ga høyere skår og at EQ-5D derfor sannsynligvis vil kunne oppdage større helseforbedringer enn 15D. Resultatene fra vårt prosjekt stemmer godt overens med resultatene fra Lundes studie, pasientene fikk ved baseline en betraktelig høyere utility index kalkulert med 15D versus EQ-5D (Tabell 2).

Hovedkritikken mot EQ-5D har vært at det, pga det lave antallet spørsmål og svaralternativer har liten diskriminerende egenskap sammenlignet med 15D og at det derfor også er mindre sensitivt for endring<sup>34</sup>. Dette var også årsaken til at EuroQol-gruppen i 2005 nedsatte en gruppe som hadde som mandat å gjøre skjemaet mer sensitivt for mindre endringer i helse/livskvalitet og å redusere takeffektene. Det ble bestemt å ikke inkludere nye domener, men å utvide antall svarkategorier fra 3 til 5<sup>62</sup>. Metodestudier som evaluerer den nye versjonen foregår og noen er allerede publisert<sup>36,62</sup>. Det er også laget en ”cross-walk-fil”, slik at man kan konvertere EQ-5D-5L (ny versjon) til EQ-5D-3L (gammel versjon).

Whitehurst og medarbeidere<sup>32</sup> publiserte i 2011 en systematisk oversikt over studier som sammenlignet EQ-5D og SF-6D. Studien inkluderte i alt 56 studier. Resultatene viste systematiske forskjeller mellom skjemaene, særlig i begge ender av skalaen. I den nedre enden ga SF-6D høyere gjennomsnittverdier, i den øvre enden ga EQ-5D høyere gjennomsnittverdier. Denne tendensen ser man ikke i dette prosjektet (Figur 3), men det skyldes nok primært at pasientene som var inkludert verken skåret i øvre eller nedre del av skalaen, noe som heller ikke var forventet av pasienter med muskelskjelett lidelser.

Whitehurst konkluderte med at gjennomsnittsverdier fremkommet på grunnlag av EQ-5D og SF-6D ikke var direkte sammenlignbare.

Johnsen og medarbeidere<sup>42</sup> sammenlignet, i en norsk studie, de to instrumentene på kroniske ryggpasienter, hvor også et sykdomsspesifikt instrument ble inkludert. De konkluderte med at det var store forskjeller i måleegenskaper og de ikke kan brukes om hverandre.

Også i dette prosjektet var samsvaret mellom de ulike utility idexene varierende.

Korrelasjonen mellom SF-6D og de to andre skjemaene var lav, mens korrelasjonen mellom EQ-5D og 15D var god (Tabell 3). Ser man imidlertid på indexene for EQ-5D og 15D og samsvaret mellom dem ved baseline (Tabell 2, Figur 2), er verdiene svært forskjellige, noe som indikerer at disse heller ikke er direkte sammenlignbare.

## **5.4 Endring i plager og funksjon**

Figur 4 viser pasientenes vurdering av endring når det gjaldt plager og funksjon, henholdsvis 89 % og 92.5 % oppga en bedring sammenlignet med da behandlingen startet. Selv om disse resultatene kan virke vel ”optimistiske”, stemmer det godt overens med en tidligere studie, som inkluderte 180 pasienter med kroniske muskelskjelett lidelser som mottok behandling i

privat fysioterapi praksis<sup>63</sup>. Studien viste en positiv endring når det gjaldt plager hos 84 % av pasientene.

Det globale aspektet ved GRS (endringsspørsmålet) er viktig, spørsmålet er ikke rettet mot en spesifikk dimensjon, men lar pasienten selv bestemme hvilken endring (innen for eksempel plager) de synes er mest viktig og relevant<sup>64</sup>. Mange stiller seg imidlertid skeptiske til om pasientene klarer å huske eksakt hvordan tilstanden deres var da de startet behandlingen (recall bias), noe som både kan føre til over- og underestimering av den reelle endringen. I dette prosjektet måtte det eventuelt vært overestimering. Det er imidlertid godt samsvar mellom endringene på GRS og graden av endring i utility indexene (Tabell 4 og 5), noe som skulle tilsi at den positive endringen målt med GRS er reell.

## 5.5 Kostnader

Helseøkonomiske analyser for å evaluere både kostnader og nytte av eksisterende og nye behandlingsformer er viktig. Det sentrale spørsmålet er om den ekstra effekten man oppnår av en ny (eller alternativ) behandling er verdt de eventuelle ekstra kostnadene det medfører. Som nevnt innledningsvis, har hovedfokus i prioriteringsdebatten vært på nye, dyre behandlingsformer, og lite på de ulike alternative tilbudene som gis til mange kronikerne (ofte over lang tid), blant annet er det gjort få studier på effekt av ulike fysioterapi intervensjoner som inkluderer helseøkonomiske evalueringer.

For å kunne gjøre en fullverdig økonomisk analyse må man sammenligne to eller flere alternative behandlinger, noe som ikke ble gjort i dette prosjektet. Etter å ha samtykket til å fylle ut de helserelaterte livskvalitetsskjemaene, ble pasientene inkludert fortløpende og fylte ut skjemaene før og etter behandling. Med kun en gruppe pasienter som mottar individuelt tilpasset behandling, man kan kun gjøre en delvis økonomisk evaluering<sup>25</sup>. Etter siste behandling regnet fysioterapeuten ut de totale behandlingstkostnadene (refusjon + egenandeler kalkulert etter fastsatte takster). Kun direkte kostnader ble tatt med, indirekte kostnader som transport, bruk av andre helsetjenester, medikamenter, fravær fra arbeid osv. ble ikke inkludert, da vi ønsket å belaste klinikerne minimalt i forhold til å skaffe supplerende opplysninger. Alle de involverte fysioterapeutene hadde driftsavtale med kommunen. Driftstilskuddet er et fast årlig beløp som er ment å dekke alle driftsutgifter som husleie, vedlikehold av treningsfasiliteter osv. Hver av de fem fysioterapeutene bidro med maksimalt

8 pasienter. Dette er et forsvinnende lite antall i forhold til det totale antall terapeuten behandler i løpet av et år og det ble derfor vanskelig å beregne hva ”driftutgiftene pr pasient” ville utgjøre.

For å regne ut pris pr vunnet QALY, ble de gjennomsnittlige direkte kostnadene dividert på differansen i utility index fra før til etter behandling (kostnader/differanse i utility index). Tabell 2 viser at denne gjennomsnittlige differansen varierte mye etter hvilket skjema som ble benyttet for å kalkulere indexen, fra en økning på 0,128 (ES= 0,34) for EQ-5D til en endring på 0,010 (ES= 0,002) for SF-6D. Så selv om man i dette prosjektet ikke har kunnet sammenligne kostnader og endring i utility index for to ulike behandlingsgrupper, men kun sammenligner utility index før og etter behandling på samme individer, ser man tydelig at kostnadene avhenger mye av hvilket skjema som ligger til grunn for utregning av utility indexene. Kostnadene pr vunnet QALY var henholdsvis kr. 26078,- for EQ-5D, kr. 54721,- for 15D til kr. 333800,- for SF-6D, det vil si at kostnadene går opp når endringen i utility index blir mindre, når nytten av behandlingene er så å si lik null. Dette viser med klarhet at skjemaets responsivitet har svært mye å si i forhold til kostnader pr QALY og det er viktig å velge ”riktig” skjema i forhold til den pasientgruppen man har med å gjøre. Når det gjelder pasienter med kroniske muskelskjelettplager som var inkludert i dette prosjektet, ser man en stor endring i utility indeks både for EQ-5D og 15D. Om man velger det ene eller det andre av disse skjemaene ser ut til å ha mindre betydning, selv om kostnadene blir høyere når utility indexen generert fra 15D brukes i nevneren. Velger man derimot SF-36, vil effekten bli minimal og kostnadene høye (kostnadene vil dessuten øke ytterligere pga at skjemaet er lisensbelagt og koster penger å bruke).

Man kan, utifra resultatene i dette lille prosjektet, kanskje driste seg til å så tvil om hvor meningsfylt det er å gjøre kostnad-nytte analyser når resultatene fra de ulike skjemaene varierer så mye. En annen faktor som gjør disse beregningene usikre er hvilke kostnader man inkluderer i analysene, noe som kan gjøre store utslag på resultatet.

## **5.6 Utvalget**

Pasientene ble fortløpende rekruttert da de oppsøkte fysioterapeut. De fleste pasientene hadde kroniske plager, 42 % hadde nakke-skulder problemer og 23 % hadde artrose i hofte- eller kneledd. Spekteret av pasienter stemmer godt overens med andre studier gjort i privat praksis

<sup>65-67</sup>, hvor pasienter med rygg, nakke-skulder og undereks problematikk utgjør de største gruppene. Utvalget i studien er lite, men i og med at hensikten med prosjektet var å evaluere anvendbarheten av og responsiviteten til skjemaene og ikke effekten av behandlingen, bør resultatene kunne gi nyttig informasjon i metodeevalueringsammenheng.

## 5.7 Hvilket skjema egner seg best i fysioterapi praksis?

For at et evalueringsverktøy skal være egnet for klinisk praksis må det ha gode psykometriske egenskaper (reproduserbarhet, validitet og responsivitet), være lett å administrere, enkelt for pasientene å forstå og ikke være forbundet med ekstra kostnader.

Tabell 5

	Finnes elektronisk	Akseptable psykometriske egenskaper	Kostnader ved bruk	Anvendelse i studier
<b>EQ-5D</b>	JA	JA	NEI	Primært Europa
<b>15D</b>	JA	JA	NEI	Primært Skandinavia
<b>SF-6D</b>	JA	JA	JA	Hele verden

Tabell 5 illustrerer viktige egenskaper ved de ulike helse-relaterte livskvalitetsskjemaene. Alle skjemaene finnes nå i elektronisk utgave, noe som kan lette arbeidet både med å administrere skjemaene i det daglige og å evaluere resultatene, ikke minst hvis de skal benyttes i daglig klinisk arbeid. Alle de tre skjemaene har dokumentert akseptable psykometriske egenskaper og det finnes også retningslinjer for klinisk relevante endringer. SF-6D er det eneste skjema hvor man må betale for bruk. Man søker [qualitymetric.com](http://qualitymetric.com) om å få bruke skjemaet, det er en lang omstendelig prosess og man får til slutt en lisens på x antall pasienter til en gitt pris. Dette er ikke forenelig med bruk i daglig praksis. Både EQ-5D og 15D er gratis og lette å anskaffe.

Gjør man søk i Pubmed viser antall treff at EQ-5D oftere er benyttet for evaluering av kostnad-nytte både i ortopediske- og fysioterapistudier enn 15D (forhold ca 8:1). Dette kan nok skyldes flere forhold, for det første er EQ-5D enklest og raskest å fylle ut. EQ-5D



innholder spørsmål om funksjon og smerte som er de viktigste endepunktene innen disse fagområdene, og man unngår å spørre pasientene om funksjoner de kan oppfatte som irrelevante i forhold til hva som feiler dem og i forhold til behandlingen de har mottatt.

Ser man på endring i forhold til hvert enkeltspørsmål, fant man i dette prosjektet signifikant endring i alle 5 domenene for EQ-5D og i 7 domener for 15D. De domenene hvor man ikke fant endring var syn, hørsel, søvn, spising, tale, vannlating/avføring, mental funksjon og depresjon, domener som er mindre aktuelle når man evaluerer effekt av ulike fysioterapitiltak. 15D inkluderer allikevel så mange aktuelle domener hvor man ser en endring fra før til etter behandling, at det absolutt kan være et aktuelt skjema å inkludere både i studier og klinisk arbeid.

På grunnlag av resultater fra dette prosjektet og tidligere publiserte metode- og effektstudier er sannsynligvis EQ-5D det livskvalitetsskjemaet som egner seg best i fysioterapi praksis i primærhelsetjenester. Det inneholder aktuelle domener, er enkelt å fylle ut, finnes i elektronisk versjon, har gode måleegenskaper, er mye brukt og gratis tilgjengelig i norsk versjon fra [www.euroqol.org](http://www.euroqol.org) og viser god evne til å fange opp endring over tid. 15D har stort sett de samme egenskapene, men er mer omfattende, inneholder en del domener som muligens er lite aktuelle i fysioterapi praksis og tar litt lenger tid å fylle ut. Dette lille metodeprosjektet viser at begge skjemaene har god responsivitet og at det derfor blir opptil hver enkelt hvilket av skjemaene man velger å bruke. SF-6D anbefales ikke, det er svært omfattende og tar mye tid, har dårlig responsivitet for endring etter fysioterapi behandling og er dyrt i bruk.

## 5.8 Et lite apropos

Det har i det siste vært en pågående diskusjon om bruk av QALY som mål på helseeffekt i økonomisk evaluering av helsetiltak. Spørsmålene om hva QALY egentlig står for, fordeler og ulemper osv., er sentrale spørsmål i denne debatten. I sitt møte i april 2013 hadde Nasjonalt råd for prioriteringer i helsevesenet bedt Helsedirektoratet om en orientering om avveiningene de hadde gjort i forbindelse med å anbefale QALY som mål på helseeffekt. Rådet mente informasjonen de fikk fra Direktoratet ikke var tilstrekkelig og at de trengte ytterligere informasjon for å kunne diskutere temaet videre (Dagens Medisin 9.april 2013).

I mars 2013 sendte dessuten The European Consortium of Health Outcomes, WHO, ut nye guidelines for økonomisk evaluering av helsetiltak (Vedlegg 1), hvor det står at QALY ikke

lenger skal brukes når helseøkonomiske prioriteringer skal foretas, da dette kan føre til feilaktige avgjørelser og prioriteringer. Hvor vidt dette er en allmenn oppfatning og om anbefalingene følges, gjenstår å se. Så lenge det ikke finnes noe godt alternativ, vil man sannsynligvis fortsette å bruke QALY's.

## 6 Konklusjon

Etter et individuelt tilpasset behandlings opplegg i primærhelsetjenesten rapporterte 90 % av pasientene bedring både når det gjaldt plager og funksjon. To av de helserelevante livskvalitetsskjemaene, EQ-5D og 15D, viste en høysignifikant endring i utility index og en stor effektstørrelse fra før til etter behandling. Begge skjemaene er enkle og gratis å skaffe til veie, enkle for pasientene å fylle ut og de finnes nå i elektronisk utgave. De har begge gode evner til å fange opp endring over tid. I denne tidlige fasen av cost-utility analyser i fysioterapi vil det derfor sannsynligvis være nyttig å inkludere begge skjemaene, da har man mulighet til å teste ut hvilket av skjemaene som egner seg best til ulike pasientgrupper og for ulike formål.

De direkte kostnadene for en behandlingsserie hos fysioterapeut i primærhelsetjenesten var ca 3500,- og kostnaden pr QALY var kr. 26087,- kalkulert på bakgrunn av endring i utility index for EQ-5D og kr. 54721,- for 15D.

## 7 Referanseliste

- (1) Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003;81(9):646-656.
- (2) Holth HS, Werpen HK, Zwart JA, Hagen K. Physical inactivity is associated with chronic musculoskeletal complaints 11 years later: results from the Nord-Trøndelag Health Study. *BMC Musculoskelet Disord* 2008;9:159.
- (3) Nasjonalt Ryggnettverk. Plager flest - koster mest - muskel-skjelettlidelser i Norge. Available at: URL: [http://www.formi.no/images/uploads/pdf/rapport\\_sept\\_04.pdf](http://www.formi.no/images/uploads/pdf/rapport_sept_04.pdf).
- (4) Ihlebaek C, Brage S, Natvig B, Bruusgaard D. [Occurrence of musculoskeletal disorders in Norway]  
10. *Tidsskr Nor Laegeforen* 2010 December 2;130(23):2365-2368.
- (5) Ihlebaek C, Laerum E. [Hits most, costs most and gets least]  
12. *Tidsskr Nor Laegeforen* 2010 November 4;130(21):2106.
- (6) Jamtvedt G, Dahm KT, Christie A et al. Physical therapy interventions for patients with osteoarthritis of the knee: an overview of systematic reviews. *Phys Ther* 2008 January;88(1):123-136.
- (7) Smith C, Grimmer-Somers K. The treatment effect of exercise programmes for chronic low back pain. *J Eval Clin Pract* 2010 April 29.
- (8) Helse- og omsorgsdepartementet. Samhandlingsreformen i kortversjon. Available at: URL: <http://www.regjeringen.no/nb/dep/hod/kampanjer/samhandling/om-samhandlingsreformen/samhandlingsreformen-i-kortversjon.html?id=650137#>.
- (9) Sosial- og helsedepartementet. Prioritering på ny, Lønning 2. Gjennomgang av retningslinjer for prioriteringer innen norsk helsetjeneste. Available at: URL: <http://www.regjeringen.no/Rpub/NOU/19971997/018/PDFA/NOU199719970018000/DDDPDFA.pdf>.
- (10) Woolf AD, Erwin J, March L. The need to address the burden of musculoskeletal conditions  
10. *Best Pract Res Clin Rheumatol* 2012 April;26(2):183-224.
- (11) Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment  
36. *Eur J Pain* 2006 May;10(4):287-333.
- (12) Brage S, Ihlebaek C, Natvig B, Bruusgaard D. [Musculoskeletal disorders as causes of sick leave and disability benefits]  
49. *Tidsskr Nor Laegeforen* 2010 December 2;130(23):2369-2370.

- (13) Ihlebaek C, Brage S, Natvig B, Bruusgaard D. [Occurrence of musculoskeletal disorders in Norway]  
14. *Tidsskr Nor Laegeforen* 2010 December 2;130(23):2365-2368.
- (14) Rokstad K, Straand J. Drug prescribing during direct and indirect contacts with patients in general practice. A report from the More & Romsdal Prescription Study 7. *Scand J Prim Health Care* 1997 June;15(2):103-108.
- (15) Hagen KB, Dagfinrud H, Moe RH et al. Exercise therapy for bone and muscle health: an overview of systematic reviews  
10. *BMC Med* 2012;10:167.
- (16) HELFO. Statistikk over fysioterapeuters takstbruk i 2011.
- (17) Fransen M, McConnel S. Exercise for osteoarthritis og the knee (Review). *The Cochrane Library* 2009;(3).
- (18) Drummond MF, Sculper MJ, Torrance GW, O'Brian BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3 ed. New York: Oxford University Press Inc; 2005.
- (19) Kvamme MK, Lie E, Kvien TK, Kristiansen IS. Two-year direct and indirect costs for patients with inflammatory rheumatic joint diseases: data from real-life follow-up of patients in the NOR-DMARD registry  
1. *Rheumatology (Oxford)* 2012 September;51(9):1618-1627.
- (20) Nord E. [Health economics--short introduction to cost-benefit analyses]. *Tidsskr Nor Laegeforen* 2002 November 20;122(28):2719-2722.
- (21) Mukuria C, Brazier J. Valuing the EQ-5D and the SF-6D health states using subjective well-being: a secondary analysis of patient data  
2. *Soc Sci Med* 2013 January;77:97-105.
- (22) Doland, Olsen JA. Kostnads effektanalyse (CEA og CUA).
- (23) Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001 July;33(5):328-336.
- (24) Olsen JA. *Helseøkonomi effektivitet og rettferdighet*. 3 ed. Oslo: J.W.Cappelens Forlag as; 2011.
- (25) Korthals-de B, I, van TM, van DH, Bouter L. Economic evaluations and randomized trials in spinal disorders: principles and methods. *Spine (Phila Pa 1976)* 2004 February 15;29(4):442-448.
- (26) Philips C. What is a QALY. Available at: URL:  
<http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/QALY.pdf>.
- (27) Nord E, Daniels N, Kamlet M. QALYs: some challenges. *Value Health* 2009 March;12 Suppl 1:S10-S15.

- (28) Lillegraven S, Kristiansen IS, Kvien TK. Comparison of utility measures and their relationship with other health status measures in 1041 patients with rheumatoid arthritis  
10. *Ann Rheum Dis* 2010 October;69(10):1762-1767.
- (29) Brazier J, Ratcliffe J. The measurement and valuation of health for economic evaluation. *Measuring and valuing health benefits for economic evaluation*. 1 ed. Oxford, UK: Oxford University Press; 2007:252-261.
- (30) Grieve R, Grishchenko M, Cairns J. SF-6D versus EQ-5D: reasons for differences in utility scores and impact on reported cost-utility  
55. *Eur J Health Econ* 2009 February;10(1):15-23.
- (31) Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life  
58. *J Clin Epidemiol* 2003 April;56(4):317-325.
- (32) Whitehurst DG, Bryan S, Lewis M. Systematic review and empirical comparison of contemporaneous EQ-5D and SF-6D group mean scores  
7. *Med Decis Making* 2011 November;31(6):E34-E44.
- (33) Lunde L. Can EQ-5D and 15D be used interchangeably in economic evaluations? Assessing quality of life in post-stroke patients  
5. *Eur J Health Econ* 2012 June 8.
- (34) Vainiola T, Pettila V, Roine RP, Rasanen P, Rissanen AM, Sintonen H. Comparison of two utility instruments, the EQ-5D and the 15D, in the critical care setting  
2. *Intensive Care Med* 2010 December;36(12):2090-2093.
- (35) Janssen MF, Pickard AS, Golicki D et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study  
1. *Qual Life Res* 2012 November 25.
- (36) Fogarty E, Walsh C, Adams R, McGuigan C, Barry M, Tubridy N. Relating health-related Quality of Life to disability progression in multiple sclerosis, using the 5-level EQ-5D  
1. *Mult Scler* 2013 February 11.
- (37) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36  
15. *J Health Econ* 2002 March;21(2):271-292.
- (38) Rivero-Arias O, Gray A, Frost H, Lamb SE, Stewart-Brown S. Cost-utility analysis of physiotherapy treatment compared with physiotherapy advice in low back pain  
11. *Spine (Phila Pa 1976)* 2006 May 20;31(12):1381-1387.
- (39) Davis JC, Marra CA, Robertson MC et al. Economic evaluation of dose-response resistance training in older women: a cost-effectiveness and cost-utility analysis  
35. *Osteoporos Int* 2011 May;22(5):1355-1366.

- (40) Whitehurst DG, Bryan S, Hay EM, Thomas E, Young J, Foster NE. Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee  
10. *Phys Ther* 2011 May;91(5):630-641.
- (41) Korthals-de B, I, Hoving JL, van Tulder MW et al. Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial  
8. *BMJ* 2003 April 26;326(7395):911.
- (42) Johnsen LG, Hellum C, Nygaard OP et al. Comparison of the SF6D, the EQ5D, and the Oswestry disability index in patients with chronic low back pain and degenerative disc disease  
2. *BMC Musculoskelet Disord* 2013;14:148.
- (43) Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders  
22. *Qual Life Res* 2008 April;17(3):485-495.
- (44) Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups  
117. *Health Econ* 2004 September;13(9):873-884.
- (45) Van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease  
36. *Health Qual Life Outcomes* 2006;4:20.
- (46) Mokkink LB, Terwee CB, Patrick DL et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes  
12. *J Clin Epidemiol* 2010 July;63(7):737-745.
- (47) Pinto D, Robertson MC, Abbott JH, Hansen P, Campbell AJ. Manual therapy, exercise therapy, or both, in addition to usual care, for osteoarthritis of the hip or knee. 2: economic evaluation alongside a randomized controlled trial  
1. *Osteoarthritis Cartilage* 2013 June 27.
- (48) Pallant J. *SPSS survival manual*. 4th ed. Berkshire: Open University Press; 2010.
- (49) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
- (50) Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-160.
- (51) Osnes-Ringen H, Kvamme MK, Kristiansen IS et al. Cost-effectiveness analyses of elective orthopaedic surgical procedures in patients with inflammatory arthropathies  
2. *Scand J Rheumatol* 2011 March;40(2):108-115.

- (52) Hurley MV, Walsh NE, Mitchell HL et al. Economic evaluation of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain. *Arthritis Rheum* 2007 October 15;57(7):1220-1229.
- (53) Lewis M, James M, Stokes E et al. An economic evaluation of three physiotherapy treatments for non-specific neck disorders alongside a randomized trial. *Rheumatology (Oxford)* 2007 November;46(11):1701-1708.
- (54) Tan SS, van Linschoten RL, van MM, Koes BW, Bierma-Zeinstra SM, Koopmanschap MA. Cost-utility of exercise therapy in adolescents and young adults suffering from the patellofemoral pain syndrome  
26. *Scand J Med Sci Sports* 2010 August;20(4):568-579.
- (55) Sach TH, Barton GR, Jenkinson C, Doherty M, Avery AJ, Muir KR. Comparing cost-utility estimates: does the choice of EQ-5D or SF-6D matter? *Med Care* 2009 August;47(8):889-894.
- (56) Hirvonen J, Blom M, Tuominen U et al. Health-related quality of life in patients waiting for major joint replacement. A comparison between patients and population controls. *Health Qual Life Outcomes* 2006;4:3.
- (57) Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005 August;14(6):1523-1532.
- (58) Cuesta-Vargas AI, Gonzalez-Sanchez M, Casuso-Holgado MJ. Effect on health-related quality of life of a multimodal physiotherapy program in patients with chronic musculoskeletal disorders  
2. *Health Qual Life Outcomes* 2013;11:19.
- (59) Salo PK, Hakkinen AH, Kautiainen H, Ylinen JJ. Effect of neck strength training on health-related quality of life in females with chronic neck pain: a randomized controlled 1-year follow-up study  
1. *Health Qual Life Outcomes* 2010;8:48.
- (60) Kauppila AM, Kyllonen E, Ohtonen P et al. Multidisciplinary rehabilitation after primary total knee arthroplasty: a randomized controlled study of its effects on functional capacity and quality of life  
3. *Clin Rehabil* 2010 May;24(5):398-411.
- (61) Saarni SI, Harkanen T, Sintonen H et al. The impact of 29 chronic conditions on health-related quality of life: a general population survey in Finland using 15D and EQ-5D  
47. *Qual Life Res* 2006 October;15(8):1403-1414.
- (62) Herdman M, Gudex C, Lloyd A et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L)  
4. *Qual Life Res* 2011 December;20(10):1727-1736.
- (63) Moseng T, Tveter AT, Holm I, Dagfinrud H. Pasient Spesifikk Funksjons Skala: Et nyttig verktøy for fysioterapeuter i primærhelsetjenesten. *Fysioterapeuten* 2013;2.



- (64) Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009;17(3):163-170.
- (65) Goldstein MS, Elliot SD, Guccione AA. The development of an instrument to measure satisfaction with physical therapy. *Phys Ther* 2000;80(9):853-863.
- (66) Di Fabio RP, Boissonnault W. Physical therapy and health-related outcomes for patients with common orthopaedic diagnosis. *J Orthop Sports Phys Ther* 1998;27(3):219-230.
- (67) Holm I, Risberg MA, Steen H. Outpatient physical therapy influences the patients' health-related quality of life. *Advances in Physiotherapy* 2005;7:40-47.

# 8 Vedlegg

## Vedlegg 1



NOKC - Norwegian Knowledge Center for  
the Health Services  
Postboks 7004, St. Olavs plass  
N-0130 Oslo  
Norway

7<sup>th</sup> March 2013

**Subject: New European Guidelines in Cost-Effectiveness assessment**

Dear Sir/Madam

The European Consortium of Health Outcomes "ECHOUTCOME" is a project funded by the European Commission under the FP7 Program with the objective to validate the scientific robustness of healthcare outcomes. The ECHOUTCOME research project has conducted the largest survey ever undertaken to specifically validate the underlying assumptions of the Quality Adjusted Life Years (QALYs) indicator for health technology assessment. The results establish that the QALY calculation is flawed and may lead to wrong and inconsistent reimbursement decisions. Consequently, the use of QALYs should no longer be recommended to assist health care decision-making in Europe. We thus conclude that cost-effectiveness assessments in health care should use other approaches and that further research is needed to develop new tools for measuring health outcomes for supporting researchers and decision-makers to rank priorities in healthcare.

Please find enclosed a hand-out presenting the new European Guidelines in Cost-Effectiveness. We trust that these guidelines will advance health technology assessment in Europe, and establish more transparent methods of reference for assessing the overall benefits and monetary value of innovative medicines and health care interventions.

Acknowledging the importance of using robust health technology assessment methods for the optimal use of innovative medicines in Europe, we remain at your disposal would you wish to organize an information and discussion session within your organization.

With best regards,

A handwritten signature in blue ink that reads 'Ariel'.

Ariel Beresniak, MD, MPH, PhD  
CEO Data Mining International  
ECHOUTCOME Project Leader  
Route de l'Aéroport, 29-31  
CP 221, CH-1215, Geneva 15  
Switzerland  
aberesniak@datamining-international.com