

Interrater reliabilitet av et klassifikasjonssystem for pasienter med uspesifikke kroniske korsryggsmerter

En metodestudie

Heléne Engberg Skaara



Mastergradsoppgave

Seksjon for helsefag, Det medisinske fakultet

UNIVERSITETET I OSLO

Oktober 2008

Forord

Masterstudiet i helsefag ved UIO har vært et spennende og utviklende studie. Det har satt meg i kontakt med mange inspirerende forskere og master studenter fra forskjellige ståsted og helsefaglige retninger.

Jeg har hatt mye støtte og god veiledning til å gjennomføre denne masteroppgaven, og det er flere som fortjener en stor takk i samband med dette!

For det første vil jeg takke min veileder Kjersti Storheim som har guidet meg gjennom denne prosessen. Du har vært klar og tydelig og gitt meg mange konstruktive tilbakemeldninger! Jeg har satt stort pris på din positive og engasjerte måte å veilede meg på.

Takk også til Inger Holm for tilbakemeldning på oppgaven i slutfasen.

Hanne, uten din medvirkning hadde ikke denne studien blitt gjennomført. Tenk at prosjektet vi startet på for fire år siden nå endelig er ferdig.

Jeg har fått en god oppbakking fra mine kollegaer og ledere på Avdeling for fysikalsk medisin og rehabilitering på Ullevål Universitets Sykehus. Min arbeidsplass har også støttet meg med en god del permisjon som gjort det mulig å studere ved siden av jobben. Det har jeg satt stor pris på!

NFF 's Fond for kurs og videreutdanning som har bidratt med økonomisk støtte til dette studiet, og personalet på medisinsk bibliotek ved UUS som har vært meget behjelpelige fortjener også å nevnes.

En takk til familie og venner! Beate du er en engel! Din hjelp med PC spørsmål og språkvask har vært viktig! Svorskene har sitt eget språk!

Max, Felix, Axel, Håkon, Ingrid og Mari uten deres tålmodighet hadde jeg ikke klart å komme i mål med oppgaven. Nå kan vi endelig spise middag på spisestuebordet! Oliver, du er min største inspirator og har vært en uvurderlig diskusjonspartner og støttespiller gjennom hele prosessen!

Heléne

Sammendrag

Studie Design: En metodologisk reliabilitets studie.

Formål: Undersøke interrater reliabiliteten av et klassifiseringssystem for uspesifikke kroniske korsryggsmerter utviklet av Peter O'Sullivan.

Teoretisk forankring: Mange ulike faktorer er involvert i og påvirker uspesifikke kroniske korsryggsmerter og dets forløp. Problemer med å finne best mulig behandling til pasienter med uspesifikke korsryggsmerter har ført til behov for å identifisere subgrupper. Et godt klassifikasjonssystem er en forutsetning for å skape homogene subgrupper. Den kliniske anvendbarheten av et klassifikasjonssystem avhenger blant annet av dette systemets reliabilitet og generaliserbarhet.

Metode: O'Sullivans klassifikasjonssystem baserer seg på en kombinasjon av bevegelsesanalyse og smerteanamnese. I studien ble pasienter med uspesifikke kroniske korsryggsmerter videofilmet ved standardiserte bevegelser og intervjuet (smerteanamnese). To terapeuter klassifiserte pasientene uavhengig av hverandre i tråd med O'Sullivan's klassifikasjonssystem. Pasientene ble klassifisert til movement impairment (MI) eller motor control impairment (MCI). Gruppen MCI ble deretter subgruppert til 5 undergrupper etter smerteprovoserende bevegelsesmønstre. Usikre scorer ble markert på scoringsskjemaet. SPSS ble brukt for statistisk analyse. Reliabiliteten på klassifikasjonssystemet ble undersøkt ved hjelp av prosentvis enighet (%) og Kappa (K) (95 % konfidens intervall).

Resultater: 41 pasienter (gjennomsnitt alder 40 ± 8 år, 22 menn og 19 kvinner) ble inkludert. Det var 100 % enighet og $K= 1.00$ ved klassifisering av pasientene til MI ($n = 4$) og MCI ($n = 37$). Subgruppering av MCI gruppen ($n= 37$) til fem smerteprovoserende bevegelsesmønstre ga 65 % enighet og $Kappa = 0.52$ (0.32 - 0.71). Subgruppering av MCI ut fra bare anamnese ga en enighet på 81 % med $Kappa = 0.66$ (0.44 - 0.87). Man fant en signifikant sammenheng mellom usikkerhet ved scoring og uenighet.

Konklusjon: Interrater reliabiliteten av O'Sullivans klassifikasjonssystem, basert på videofilm og smerteanamnese, var svært god ved subgruppering til hovedgruppene MI og MCI, men moderat ved subgruppering til fem bevegelsesmønstre i hovedgruppen MCI. Subgruppering av MCI ut fra kun smerteanamnese viste god reliabilitet. Videre forskning er nødvendig for å identifisere årsakene til uenighetene ved subgruppering under MCI.

Abstract

Study Design: A methodological reliability study.

Objectives: To examine the interrater reliability of a classification system, developed by Peter O'Sullivan (O'S), for non-specific chronic low back pain patients.

Background: Many different factors involve and affect non-specific low back pain and its course. The effort of finding the most efficient treatment for this heterogeneous group has raised the question about identifying subgroups. A well documented classification system is a prerequisite to create homogenous subgroups. The utility of a classification system depends, among other factors, on its reliability and its ability to generalize.

Methods: The classification system of O'Sullivan is based on a combination of movement analysis and pain history. In this study non-specific chronic low back pain patients were video filmed during specific movements and a pain history was undertaken. Based on both video film and interview two therapists classified independently using the O'S classification system. The patients were classified into motor control impairment (MCI) or movement impairment (MI). The MCI group was further sub grouped into 5 categories depending on painful movement pattern. When it was difficult to decide a score this was marked on the score table. SPSS was used for statistic analyses. The reliability of the classification system was calculated by percentage of agreement (%) and Kappa (*K*) (95 % confidence interval).

Results: 41 patients (mean age 40 ± 8 years, 22 male and 19 female) were included. There was a 100 % of agreement and $K = 1.00$ classifying patients to MI ($n = 4$) and MCI ($n = 37$). Sub grouping the MCI patients into five different painful movement patterns the agreement was 65 % and Kappa = 0.52 (0.32 - 0.71). Sub grouping the MCI patients from only the pain history showed an agreement of 81% with Kappa = 0.66 (0.44 - 0.87). The relation between insecurity setting the score and disagreement was found significant.

Conclusions: The interrater reliability of the O'Sullivan classification system, based on video and interview (pain history), was excellent when classifying into the two main groups MCI and MI. When classifying MCI into five subgroups the reliability was found to be moderate. Classifying MCI into subgroups from only the interview (pain history) showed good reliability. Further research is needed to identify sources of disagreement when classifying into the MCI subgroups.

INNHOLDSFORTEGNELSE

FORORD	I
SAMMENDRAG	II
ABSTRACT	III
INNHOLDSFORTEGNELSE	IV
<i>Liste over tabeller</i>	<i>vii</i>
<i>Liste over figurer</i>	<i>vii</i>
<i>Liste over vedlegg</i>	<i>viii</i>
<i>Forkortelser</i>	<i>viii</i>
1. INNLEDNING	1
1.1 BAKGRUNN FOR OPPGAVEN	1
1.2 FORMÅL MED STUDIEN OG PROBLEMSTILLING	2
1.3 OPPBYGNING OG AVGRENSING AV OPPGAVEN	3
1.4 BEGREPSAVKLARINGER	4
2. TEORETISK BAKGRUNN	6
2.1 USPESIFIKKE KORSRYGGSMERTER	6
2.1.1 <i>Diagnostiske utfordringer</i>	6
2.1.2 <i>Behov for å subgruppere</i>	8
2.2 SUBGRUPPERING AV PASIENTER MED USPESIFIKKE KRONISKE KORSRYGGSMERTER	8
2.2.1 <i>Generelt om klassifikasjonssystemer</i>	8
2.2.2 <i>Organisatorisk stamme for klassifikasjonssystem</i>	9
2.2.3 <i>Fysioterapirelevante klassifikasjonssystemer for uspesifikke korsryggsmarter</i>	10
2.2.4 <i>Oppsummering av presenterte klassifikasjonssystemer</i>	16
2.3. METODER FOR Å VURDERE KVALITETEN PÅ ET KLASSIFIKASJONSSYSTEM	18
2.3.1 <i>Reliabilitet</i>	18
2.3.2 <i>Validitet</i>	19
2.3.3 <i>Kriterer som bidrar til reliabilitet og validitet av klassifikasjonssystemer</i>	20
2.3.4 <i>Reliabilitet og validitet av klassifikasjonssystemer for LBP</i>	22
3. MATERIALE OG METODE	29
3.1 BAKGRUNN FOR VALG AV DELTAGERE	29
3.2 UTVALGET	29
3.2.1 <i>Inklusjonskriterier</i>	30
3.2.2 <i>Eksklusjonskriterier</i>	30
3.3 TERAPEUTENE	31

3.4 STUDIEDESIGN OG PROSEDYRER	32
3.4.1 Videofilming av bevegelsesanalysen	33
3.4.2 Anamnese	34
3.5 DATAINNSAMLING	35
3.5.1 Demografiske data	35
3.5.2 Randomisert rekkefølge ved subgruppering av utvalget	36
3.5.3 Subgruppering	36
3.5.4 Subgruppering ved bare anamnese	37
3.6 DATABEHANDLING	38
3.6.1 Variabler	38
3.6.2 Statistiske metoder	38
3.7 ETIKK	40
4. RESULTATER	41
4.1 DEMOGRAFISKE DATA	41
4.1.1 Deltagere	41
4.1.2 Fordeling av filming og anamneseopptak mellom terapeutene	42
4.2 RELIABILITET	43
4.2.1 Subgruppering av control impairment og movement impairment	43
4.2.2 Subgruppering av pasienter med control impairment	43
4.2.3 Sammenslåing av mønster etter smerteretning ved subgruppering av control impairment	45
4.2.4 Subgruppering av control impairment ved kun anamnese	45
4.3 SAMMENHENG MELLOM UENIG OG USIKKER VED SCORING	46
4.4 OPPSUMMERING AV RESULTATENE	47
5. DISKUSJON	48
5.1 UTVALG	48
5.1.1 Inklusjon og eksklusjon	48
5.1.2 Deltakelse	50
5.1.3 Ekstern validitet	50
5.2 METODE	51
5.2.1 Valg av O'Sullivan's klassifikasjonssystem i metodestudien	51
5.2.2 Valg av design	53
5.2.3 Observasjon som metode ved registrering av bevegelsesmønster	54
5.2.4 Valg av video ved analyse av bevegelsesmønster	57
5.2.5 Anamnese som grunnlag for subgruppering	58
5.2.6 Opptaksprosedyrene og scoringsprosedyren	59
5.2.7 Terapeutene	60
5.2.8 Klassifikasjonssystemets utvikling under studien	61
5.3 STATISTIKK	62
5.3.1 Valg av statistiske metoder	62
5.3.2 Statistisk usikkerhet	63

5.4 RESULTATER	64
<i>5.4.1 Demografiske data</i>	64
<i>5.4.2 Subgruppering til movement impairment eller control impairment</i>	65
<i>5.4.3 Subgruppering av pasienter med control impairment til fem ulike kategorier</i>	65
<i>5.4.4 Interrater reliabilitet ved bare anamnese sammenlignet med annamnese og observasjon</i>	68
<i>5.4.5 Sammenheng mellom uenig og usikker ved scoring</i>	69
5.5 INTERRATER RELIABILITET	70
5.6 KLINISKE IMPLIKASJONER	72
6. OPPSUMMERING OG KONKLUSJON	75
7. REFERANSER	77
VEDLEGG 1	89
VEDLEGG 2	92
VEDLEGG 3	98
VEDLEGG 4	99
VEDLEGG 5	111
VEDLEGG 6	115
VEDLEGG 7	120
VEDLEGG 8	121
VEDLEGG 9	122
VEDLEGG 10	123
VEDLEGG 11	125
VEDLEGG 12	126

Liste over tabeller

Tabell 1	Hovedtrekk ved ulike klassifikasjonssystemer
Tabell 2	Reliabilitet og validitet for ulike klassifikasjonssystemer
Tabell 3	Gradering av Kappa verdier
Tabell 4	Demografiske data
Tabell 5	Kappaverdi og prosentvis enighet ved subgruppering av movement impairment og control impairment
Tabell 6	Kappaverdi og prosentvis enighet ved subgruppering av control impairment til 5 ulike mønstre
Tabell 7	Kappaverdi og prosentvis enighet ved sammenslåing av de 5 mønstergruppene til 3 større grupper.
Tabell 8	Kappaverdi og prosentvis enighet ved subgruppering av control impairment ved kun anamnese.
Tabell 9	Sammenheng enig - uenig og sikker - usikker

Liste over figurer

Figur 1	Organisasjonsstamme for klassifikasjonssystem (Buchbinder et al,1994)
Figur 2	Subgruppering av kroniske korsryggsmerter O'Sullivan 2005 (O'Sullivan, 2005)
Figur 3	Flytskjema over metodestudien
Figur 4	Subgruppering av feiltilpasset motorisk kontroll til movement impairment eller control impairment
Figur 5	Subgruppering fra controlimpairment til fem ulike bevegelsesmønstre
Figur 6	Eksempel på ulike måter å bøye ryggen på (bilde fra Sahrman, 2002)
Figur A	Klassifikasjonsprosess av O'Sullivan hentet fra Vibe Fersum's artikkel 2008 (Vibe et al, 2008)

Liste over vedlegg

- Vedlegg 1 Godkjenning fra Regional komité for medisinsk forskningsetikk
- Vedlegg 2 Godkjenning fra Datatilsynet
- Vedlegg 3 Pasientsamtykke
- Vedlegg 4 Skiveprotesestudien
- Vedlegg 5 O'Sullivan's klassifikasjonssystem
- Vedlegg 6 Hovedkriterier for subgruppering ut fra anamnese og video
- Vedlegg 7 Bevegelses- og smerteskjema for videofilming
- Vedlegg 8 Anamneseskjema
- Vedlegg 9 Skåringskjema
- Vedlegg 10: Oswestry liste for funksjonsbegrensning
- Vedlegg 11: FABQ PA+W
- Vedlegg 12: HSCL

Forkortelser

FABQ - Fear Avoidance Beliefs Questionnaire

HSCL-25 - Hopkins Symptom Check List

ICD - The International Classification of Disease

ICDIH - The International Classification of Disease, Activities and Handicap

ICF - The International Classification of Functioning, Disability and Health

LBP - Low Back Pain

MI - Movement Impairment

MCI - Motor Control Impairment

ODI - Oswestry Disability Index

QTFC - Quebec Task Force Classification

RCT - Randomised Controlled Trial

1. Innledning

1.1 Bakgrunn for oppgaven

I klinikken har vi behov for å gi hver enkelt pasient optimal behandling og ivareta pasienten på alle funksjonsnivå. Diagnosebetegnelsen ”uspesifikke korsryggsmerter” er generell og gir lite informasjon om hva som ligger bak smertene, forløpet og håndtering av lidelsen. Gruppen med uspesifikke korsryggsmerter er en heterogen gruppe som grunnet sin kompleksitet er vanskelig å diagnostisere og finne en optimal behandling for (Fritz et al., 2007; Borkan et al., 1998). Dette har bidratt til at utredning og behandling kan virke tilfeldig. Det har derfor innenfor ryggforskningstiljøet lenge vært et behov for gode klinisk anvendbare klassifikasjonssystemer for å subgruppere pasienter som faller inn under betegnelsen ”uspesifikk korsryggsmerte” (Borkan et al., 2002). Utfordringen ved utvikling av et klassifikasjonssystem er å finne godt avgrensede grupper som ivaretar alle faktorer som kan påvirke smerten.

I takt med nyere forskning rundt uspesifikke korsryggsmerter i løpet av de siste årene er det utviklet mange ulike typer av klassifikasjonssystem. Dagens forståelse av ryggproblemet er preget av en bio-psyko-sosial tilnærming til forskjell fra den tidligere mer biomedisinske eller psykologiske forståelsen. Innenfor fysioterapien har den biomedisinske delen av forskningen de senere årene hatt fokus på neuromuskulære forhold hos uspesifikke ryggpasienter. Denne forandringen gjenspeiles i utviklingen av klassifikasjonssystemene. Problemet med forskning rundt klassifikasjonssystemene har imidlertid vært dårlig metodisk kvalitet. Mange klassifikasjonssystem er dessuten komplekse og vanskelige å adaptere til klinikken. Gode klinisk anvendbare klassifikasjonssystemer må være både reliable og valide. De skal fungere i klinikken så vel som i forskningsøyemed og gi pasienten og behandlere mulighet til en felles forståelse og målsetning (Borkan et al., 1998). For arbeidsgivere og trygdevesen kan et godt dokumentert og klinisk brukbart klassifikasjonssystem bidra til tidlig identifisering og igangsetting av tiltak for pasienter med risiko for å utvikle langvarige plager og dermed ha store økonomiske konsekvenser (Childs et

al., 2004). Ved utvikling av et nytt klassifikasjonssystem er det derfor viktig at flere miljøer tester det ut. Resultater og erfaringer fra flere studier og ulike miljøer er med på å forbedre klassifikasjonssystemet (Borkan et al., 2002).

1.2 Formål med studien og problemstilling

Hensikten med denne metodestudien er å vurdere reliabilitet av et relativt nytt behandlingsrelatert biopsykososialt klassifikasjonssystem. Systemet er utviklet av manuellterapeut Peter O'Sullivan for pasienter med uspesifikke kroniske korsryggmerter. Klassifikasjonssystemet til O'Sullivan ble valgt da det med sin biopsykososiale forståelse for smerteutvikling passer for de ryggpasientene jeg ser til daglig i klinikken. I denne studien gjøres en interrater reliabilitets test mellom to fysioterapeuter som begge har gått 3 helgekurs med O'Sullivan, gjort inngående studier av litteratur og praktisert systemet på pasienter i klinikken siden 2002. Med studien ønsket vi å se om to fysioterapeuter med moderat erfaring i klassifikasjonssystemet kunne teste systemet på pasienter med uspesifikke ryggmerter og oppnå tilfredsstillende reliabilitet. Reliabiliteten er viktig å teste før et klassifikasjonssystem tas i bruk i klinikken. Særlig med tanke på hvor mye det vil kreves for å tilegne seg ferdighetene som må til for å beherske systemet og for hvilke yrkesgrupper det er tiltenkt. Reliabiliteten vil dessuten være grunnleggende for testing av klassifikasjonssystemets validitet. Klassifikasjonssystemet var i 2004, da denne studien startet, relativt nytt. Gruppen i Australia som utviklet klassifikasjonssystemet var i gang med å teste det for reliabilitet, men resultatene var ikke publisert. Selv om systemet er reliabilitetstestet av gruppen som har utviklet det er det viktig å gjenta dette i andre miljøer.

Følgende problemstilling ønskes besvart:

Er klassifikasjonssystemet til O`Sullivan reliabelt med hensyn til interrater reliabilitet?

Jeg ønsker med denne problemstillingen å finne ut om to fysioterapeuter med lang

erfaring i ryggbehandling og moderat erfaring i klassifikasjonssystemet, klarer å klassifisere likt, blindet for hverandres resultat.

1.3 Oppbygning og avgrensning av oppgaven

Oppgaven er en metodestudie der et relativt nytt klassifikasjonssystem, utviklet for pasienter med uspesifikke kroniske korsryggsmerter, testes for intertetterreliabilitet. 41 pasienter inngår i studien.

Teorikapittelet i oppgaven har sitt hovedfokus på sentrale begreper som uspesifikke korsryggplager, klassifikasjonssystemer, reliabilitet og validitet. Jeg har valgt å presentere noen klassifikasjonssystemer som er mest relevante for fysioterapi. Systemene jeg beskriver er valgt på bakgrunn av at de er godt beskrevet i litteraturen, at de er delvis testet for reliabilitet og validitet og at de er klinisk anvendbare. Studien er en metodestudie der et klassifikasjonssystem testes for intertetterreliabilitet og begreper som reliabilitet og validitet er derfor viktige å belyse teoretisk.

I min metodestudie har jeg testet reliabiliteten av den delen av klassifikasjonssystemet til O'Sullivan som omhandler uspesifikke kroniske korsryggsmerter og beskriver derfor først og fremst denne delen. Jeg utdyper ikke behandlingsdelen av klassifikasjonssystemet i denne oppgaven og går heller ikke inn på de biologiske og psykologiske forklaringsmodellene på smerte generelt.

I midtre delen av oppgaven presenteres metode og resultater. Pasientene i min metodestudie inngår i en større prospektiv multisenterstudie hvor jeg har fått tilgang til alle baselinedata. Ut fra dette har jeg trukket ut demografiske data og noen kliniske variabler relevante for min studiepopulasjon.

Metode og resultater diskuteres opp mot teori og klinisk empiri i siste delen av oppgaven og avsluttes med en oppsummering.

1.4 Begrepsavklaringer

Uspesifikke kroniske korsryggsmerter kan forstås som et symptom med varighet i over 3 måneder, som kan rapporteres subjektivt av pasienten, men som ikke har en spesifisert fysiologisk årsak og derfor ikke kan påvises av andre på en objektiv måte.

”*Flagg*” er spesielle karakteristika i pasientens sykehistorie og kliniske bilde som er viktige å fange opp. Flaggene har farge etter hvilken type karakteristika man finner.

”Røde flagg” indikerer alvorlige underliggende funn og ”Gule flagg” indikerer betydningsfulle psykososiale funn

Klassifikasjonssystem er et strukturert system der et overordnet fenomen er identifisert og brutt ned i mer homogene grupper i henhold til etablerte kriterier.

Klassifisering og subgruppering betyr i denne oppgaven å plassere pasienter med kroniske uspesifikke korsryggsmerter i riktige grupper etter gitte kriterier ut fra et klassifikasjonssystem utarbeidet for pasienter med kroniske korsryggsmerter.

Bevegelsesmønstre er den innlærte måte en kroppsdel eller hele kroppen beveger seg på i ulike situasjoner og oppstår ved at ulike muskelgrupper rekrutteres etter en gitt rekkefølge i ulike deler av bevegelsen. Forandring av anatomiske strukturer og smerte er to faktorer som kan påvirke og forandre muskelrekrutteringen og dermed også bevegelsesmønstret.

Funksjonell instabilitet forstås her som en smertebetenget ubalanse i rekrutteringen og timingen mellom mobiliserende og stabiliserende muskulatur.

Dysfunksjon er i denne oppgaven satt i direkte samband med bevegelsesmønstre og betyr en uhensiktsmessig og feiltilpasset måte å utføre en bevegelse på.

Observatør er her det samme som terapeut som via filmopptak og anamnese subgrupperer pasientene i denne reliabilitetsstudien.

Smerte provoserende bevegelsesmønster eller - retning er her ment å beskrive i hvilken retning smerten oppstår eller øker ved bevegelse. For eksempel ved å bøye seg fremover, bakover eller til en av sidene.

Bio-psyko-sosial modell er her satt i sammenheng med klassifikasjonssystem der man tar hensyn til at smerten kan oppstå, påvirkes og vedlikeholdes av både biologiske, psykologiske og sosiale faktorer.

2. Teoretisk bakgrunn

2.1 Uspesifikke korsryggsmerter

En av de vanligste årsakene til kontakt med helsevesenet er ryggproblemer. Av 2 millioner ryggkonsultasjoner i primærhelsetjenesten per år utgjør 10 % fastlege konsultasjoner, 20 - 30 % fysioterapeut konsultasjoner og 70 - 80 % kiropraktor konsultasjoner (Werner & Indahl, 2005). Årlig koster disse ryggglidelsene samfunnet rundt 15 - 17 milliarder kroner i form av medisinske utgifter, tapt arbeidsinnsats og produksjonstap (Lærum et al., 2007; NAV, 2008). I 2007 utgjorde sykefravær grunnet ryggplager totalt 700.359 dagsverk. Mindre enn 50 % som er sykemeldt i mer enn seks måneder grunnet ryggsmerter kommer tilbake til jobb. Hvert år blir rundt 5000 personer uføretrygdet på grunn av ryggglidelser (NAV, 2008).

2.1.1 Diagnostiske utfordringer

Korsryggsmerter er en av de vanligste lidelsene som før eller senere rammer opp mot 80 % av befolkningen (Long et al., 1996; Lærum et al., 2007; Walker, 2000). For de fleste går akutte ryggplager over i løpet av 6 uker, men tilbakefall er vanlig og hele 60 % får tilbakefall i løpet av et år (Hestbaek et al., 2003; Waddell, 2004). For 20 % vedvarer plagene i mer enn 6 måneder (Hestbaek et al., 2003). Ryggglidelser grupperes derfor ofte etter tidsperspektiv som akutte, subakutte eller kroniske hvor retningslinjer er satt opp med definisjon og forslag til håndtering innenfor de ulike gruppene (Abenhaim et al., 2000; Lærum et al., 2007; Waddell, 2004).

I nasjonale og internasjonale retningslinjer for behandling av korsryggsmerter anbefales en initial diagnostisk klassifikasjon. En diagnostisk triade som skiller mellom alvorlig spinal sykdom, nerverotsproblem og uspesifikke årsaker til ryggsmertene er vanlig (Koes et al., 2001). Det er også vanlig å bruke "flagg" når anamnese og kliniske undersøkelser indikerer sannsynlig underliggende årsak og forløp av korsryggsmertene. "Røde flagg" indikerer mulig alvorlig patologi som

trenger videre utredning, eksempelvis infeksjon, tumor eller fraktur. ”Gule flagg” indikerer økt risiko for at tilstanden blir langvarig grunnet faktorer som er psykososiale, eksempelvis angst, depresjon, smertemestring og jobbforhold. ”Grønne flagg” er faktorer som indikerer akutte uspesifikke korsryggsmerter med god prognose og kort tids varighet (Lærum et al., 2007).

Av alle pasienter med korsryggsmerter er 80 - 90 % uspesifikke (Waddell, 2004). Til tross for at mange av pasientene i denne gruppen har radiologiske forandringer er det ikke påvist noen sikker sammenheng mellom billeddiagnostikk og kliniske funn hos pasienter med uspesifikke korsryggsmerter (Nachemson, 1999). Det har derfor vært vanskelig å skille mellom ulike typer ryggvondt i gruppen uspesifikke korsryggsmerter ut fra radiologiske funn (Nachemson, 1999; Kent & Keating, 2004; The Quebec Task Force on Spinal Disorders, 1987; Dillingham, 1995).

Både i forskningsmiljøer og blant klinikere synes det å være enighet om at pasienter med uspesifikke korsryggsmerter ikke er en homogen gruppe. Man har derfor lenge prøvd å finne ut om ulikhetene er av betydning for hvordan disse pasientene skal diagnostiseres og behandles. Behovet har økt med det faktum at forskning på uspesifikke ryggglidelser hittil har vist at mange behandlingsintervensjoner har lik, men ikke optimal effekt (Fritz et al., 2003). Pendelen har svingt fra hard trening av rygg ekstensorer ad modum Manniche til Indahls prinsipper om å gi minst mulig behandling og ufarliggjøre ryggsmertene (Indahl et al., 1995; Manniche et al., 1989; Manniche, 1989).

I den senere tids ryggforskning blir korsryggsmerter sett på som et flerdimensjonalt problem med mange faktorer som påvirker forløpet (Sahrmann, 1988; Borkan et al., 2002; McCarthy et al., 2004). Disse dimensjonene består av patoanatomiske, neurofysiologiske, fysisk mekaniske forhold og psykososiale faktorer og man snakker ofte om en biopsykososial smertemodell (Waddell, 2004). I en review artikkel fra 2004 fremkommer det at de fleste klassifikasjonssystemer som hittil er utviklet er basert på ren biomedisinsk tilnærming, men at det har skjedd en utvikling de senere årene mot en mer helhetlig biopsykososial tilnærming (McCarthy et al., 2004).

2.1.2 Behov for å subgruppere

En stor mengde effektstudier har blitt gjort på uspesifikke korsryggsmerter med forskjellige tiltak, men ingen behandling peker seg ut som spesielt effektiv. En mulig forklaring kan være at gruppen ”uspesifikke ryggpasienter” ikke er en homogen gruppe. Forskning som kan bidra til å identifisere subgrupper av pasienter med uspesifikke korsryggsmerter, har blitt framhevet som et viktig forskningsområde (Borkan et al., 1998; McCarthy et al., 2004). Ved å dele inn uspesifikke korsryggsmerter i mer homogene subgrupper (basert på valide kriterier) kan det tenkes at utredningen blir mer rasjonell og at behandlingen blir mer effektiv (Riddle, 1998; Borkan et al., 1998; Fritz & George, 2000). I de senere årene har man gjort studier på pasienter med uspesifikke korsryggsmerter som viser at sub gruppering av både akutte og kroniske LBP pasienter kan gi økt behandlingseffekt med de rette tiltakene (Brennan et al., 2006; Browder et al., 2007; Childs et al., 2004; George et al., 2003; Fritz et al., 2003; Gudavalli et al., 2006; Vollenbroek-Hutten et al., 2004; Fritz & George, 2000).

På bakgrunn av ulike teorier og oppfatninger om årsaker til korsryggsmerter har det i løpet av de siste tiårene blitt utviklet et konglomerat av ulike klassifikasjonssystemer. Det er naturlig å se nærmere på denne utviklingen, men først vil jeg si noe generelt om klassifikasjon.

2.2 Subgruppering av pasienter med uspesifikke kroniske korsryggsmerter

2.2.1 Generelt om klassifikasjonssystemer

Subgruppering betyr at man bryter ned en overordnet større enhet, her diagnosen uspesifikke korsryggsmerter, i mer homogene subgrupper (Petersen et al., 1999). Siden midten av 1980 tallet har man utviklet klassifikasjonssystemer for pasienter med korsryggsmerter. Klassifikasjonssystemene er utarbeidet av ulike yrkesgrupper og med ulike hensikter. Petersen et al. mener det optimale klassifikasjonssystemet

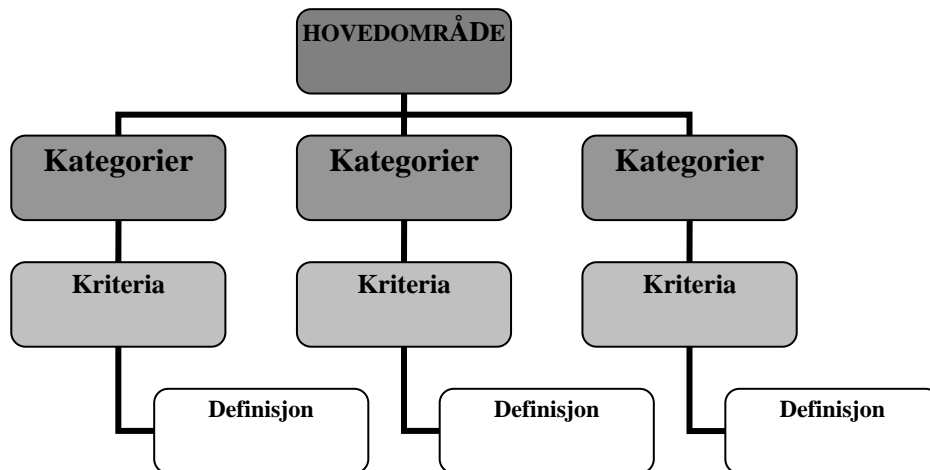
skal kunne gi en merkelapp som sier noe om årsaken til plagene, som kan predikere utviklingen av tilstanden og effekten av spesifikke terapeutiske tiltak (Petersen et al., 1999).

Det mest kjente og vanligst brukte klassifikasjonssystemet er det klassiske medisinske, The International Classification of Diseases (ICD) (World Health Organization, 2004). Den seneste versjonen ICD-10 bygger på pathoanatomiske beskrivelser, men det finnes ingen funksjonsbeskrivelser i dette systemet (ICD-10). Her er det heller ingen prosedyrer som beskriver hvilke kriterier som ligger til grunn for de ulike kategoriene. Dette systemet blir derfor lite brukbart til å fastslå patologi, sette prognoser og legge opp riktig behandling for pasienter med korsryggsmerter (Petersen et al., 1999). Reliabiliteten og validiteten har da også vist seg å være lav for dette systemet (Buchbinder et al., 1994).

I 1980 utgav WHO den første versjonen av The International Classification of Impairments, Activities and Handicaps (ICIDH) (World Health Organization, 1980). Dette er et klassifikasjonssystem hvor man i tillegg til den medisinske diagnosen beskriver pasientens funksjon både fysisk, psykisk og sosialt. Dette systemet har fått internasjonal aksept hos klinikere og forskere særlig innenfor rehabiliteringsfeltet (Soukup & Vollestad, 2001). ICIDH har blitt revidert tre ganger og i 2001 ble benevningen ICF (two level classification) (World Health Organization, 2001). Systemet inneholder nå over 9000 sykdomskategorier. ICF er fremdeles under utvikling og utprøving og man er nå i gang med å lage kjernesett innenfor ulike diagnose områder. ICF er først og fremst et deskriptivt system og ikke et måle/evalueringsredskap. Det er heller ikke relatert til behandlingstiltak.

2.2.2 Organisatorisk stamme for klassifikasjonssystem

Til grunn for de fleste nyere klassifikasjonssystem ligger en organisatorisk stamme som Buchbinder og medarbeider har beskrevet (figur 1) (Buchbinder et al., 1994).



Figur 1. Organisasjonsstamme for klassifikasjonssystem (Buchbinder et al., 1994).

Kategoriene kan være mange og skal alltid etterfølges av kriteria og definisjon. Kategoriene kan plasseres på rad og rekke som i figur 1 eller i ulike nivåer (O'Sullivan, 2005; Delitto et al., 1995). Et klassifikasjonssystem kan bygges opp på ulike måter. Det er viktig at man sikrer at pasientene grupperes etter felles egenskaper i gjensidig utelukkende kategorier. Hovedhensikten kan enten være å identifisere patologi, definere pasientens problem, predikere fremtidig status eller som guide for behandling. Noen systemer mikser disse (Feinstein, 1987; Riddle, 1998).

Det er to måter å utvikle klassifikasjonssystemer på. En statistisk tilnærming og en empirisk tilnærming. Den statistiske tilnærmingen baserer seg på forskningsbasert viten og den empiriske baserer seg på erfaringsbasert viten. Mange systemer er en blanding av disse (Feinstein, 1987; Riddle, 1998). De seneste årene har mange nye studier på korsryggsmerter blitt publisert og klassifikasjonssystemene baseres nå mer på forskningsbasert viten (O'Sullivan, 2005; Petersen et al., 2003).

2.2.3 Fysioterapirelevante klassifikasjonssystemer for uspesifikke korsryggsmerter

Det eksisterer altså en rekke ulike klassifikasjonssystem for pasienter med korsryggsmerter. Etter litteratursøking fremstår noen klassifikasjonssystem som mer hyppig og grundig beskrevne, de har blitt metodisk testet og er fysioterapirelevante. Disse vil jeg i store trekk presentere i følgende avsnitt.

Quebec Task Force Classification system (QTFC)

Quebec Task Force er et omfattende, men mye brukt klassifikasjonssystem (Billis et al., 2007). Det ble satt sammen av en tverrfaglig gruppe med eksperter innenfor ryggforskning i 1987 (The Quebec Task Force on Spinal Disorders, 1987). Initiativet ble tatt av Quebec Worker's Health and Safety Commission, da man hadde sett en stadig økende forekomst av høyt sykefravær og høye kostnader knyttet til korsryggproblematikk i Quebec. Hensiktene med klassifikasjonssystemet var mange. Man ønsket å bedre diagnostikken, lage prognoser, evaluere behandlingskvalitet og bruke systemet i forskningsøyemed (The Quebec Task Force on Spinal Disorders, 1987). Klassifikasjonssystemet kan brukes på alle pasienter med korsryggsmerter. Man kombinerer data fra symptomer og tegn, radiologiske undersøkelser, symptomenes varighet, respons av tidligere behandling og arbeidsstatus. Systemet deles i 11 kategorier der også pasienter med alvorlig patologi og nerverotaffeksjon er inkludert (Riddle, 1998). Ved kroniske plager tar systemet også hensyn til psykososiale aspekter (The Quebec Task Force on Spinal Disorders, 1987). Til forskjell fra behandlingsrelaterte klassifikasjonssystemene tar QTFC verken med ryggens bevegelsesutslag eller pasientens smerte ved ulike bevegelser. Behandlingen av de ulike gruppene er ikke definert og man har isteden via litteraturgjennomgang gitt generelle anbefalinger (Riddle, 1998).

Mc Kenzie

Mc Kenzie systemet ble utviklet i begynnelsen av 1980 tallet og er den mest brukte behandlingsmetoden i fysioterapi for pasienter med korsryggsmerter (Billis et al., 2007). Klassifikasjonssystemet er bygget opp på empirisk bakgrunn, med hensikt å guide fysioterapeuter i valg av behandling. Det er basert på en biomedisinsk forståelse (Billis et al., 2007). Via klinisk erfaring har man laget pasientkategorier ut fra respons på en spesifikk behandling. Pasientkategoriene har så blitt nøye beskrevet og satt i system. Klassifikasjonssystemet er primært utviklet for pasienter med akutte og subakutte korsryggsmerter. Klassifikasjonen baserer seg på informasjon fra

anamnese om smertenes lokalisasjon, intensitet og smerte respons ved provokasjon og avlastning (Petersen et al., 1999). Man undersøker holdning og forskjellige belastninger ved trunkus bevegelser. I tillegg undersøkes hofte og IS ledd og man utfører en neurologisk undersøkelse. Man klassifiserer pasientens problemer til 1 av 13 kategorier. Kategori 1: Posturalt syndrom som kommer av ugunstig holdning. Kategori 2 - 6: Dysfunksjonelt syndrom kommer ofte sekundært til posturalt syndrom, men betyr mer vedvarende endringer. Kategori 7 - 13: Forstyrrelse i et bevegelses segment (Riddle, 1998). McKenzie laget systemet slik at hver kategori krevde ulik behandling.

Delitto et al.

Delitto og medarbeider lanserte i 1995 et klassifikasjonssystem for pasienter med akutte og subakutte korsryggsmerter hovedsakelig med uspesifikk diagnose. Klassifikasjonssystemet har som hensikt å gi retningslinjer om behandlingsstrategier. (Delitto et al., 1995). Systemet er primært bygget opp etter biomedisinsk forståelse. Systemet deles inn i 3 nivåer. Først skilles pasienter som skal til spesialist fra de som kan klare seg med fysikalsk behandling. Deretter inndeles pasientene som skal til konservativ fysikalsk behandling etter varighetsgrad: akutt, subakutt, og pasienter som har aktivitets- og arbeidsintoleranse grunnet høye fysiske krav. Ut fra disse tre gruppene deles pasientene ytterligere inn i behandlingsgrupper på bakgrunn av symptomer, funn og funksjonsstatus. Målet er å redusere symptomer og komme rask tilbake i normal funksjon. Pasientene evalueres underveis og kan bytte behandlingsstrategi (Delitto et al., 1995). Systemet har vært under utvikling de senere årene med forbedring av undersøkelseskriterier for subgruppering og oppdaterte intervensjoner for de ulike gruppene (Fritz et al., 2007). Man har også tatt hensyn til psykososiale faktorer i det modifiserte klassifikasjonssystemet (George et al., 2003).

Sahrmann, Maluf og Van Dillen

Sahrmann og medarbeider lanserte i 2000 et klassifikasjonssystem for uspesifikke kroniske korsryggsmerter (Maluf et al., 2000). Systemet har en biomedisinsk tilnærming og baserer seg på anamnese og klinisk undersøkelse. Hensikten er å gi retningslinjer for behandling til fysioterapeuter. Systemet deler inn pasientene i 5 kategorier med bakgrunn i smerteprovoserende bevegelsesretning: 1. Fleksjon, 2. Ekstensjon, 3. Rotasjon, 4. Rotasjon med fleksjon og 5. Rotasjon med ekstensjon. Man legger stor vekt på funksjon i dagliglivet og knytter dette opp mot undersøkelsesfunnene. Den kliniske undersøkelsen ser etter muskulær stabilitet, asymmetri, holdning og bevegelse av lumbalcolumna, bekken og hofter (Maluf et al., 2000). Systemet er laget for å avdekke uhensiktsmessige bevegelsesmønstre som man mener er med på å vedlikeholde smerten og behandlingen tar sikte på å normalisere dette (Van Dillen et al., 2003).

Petersen et al.

Petersen og medarbeidere har utviklet et diagnostisk klassifikasjonssystem hvor hensikten er å identifisere klinisk homogene subgrupper med uspesifikke korsryggsmerter i forhold til antatte symptomatiske strukturer (Petersen et al., 2003). Systemet er basert på en biomedisinsk forklaringsmodell. Han bygger sine kategorier ut fra prevalens studier der man har lagt anestetiske blokkader på ulike symptomatiske strukturer. Systemet er primært utviklet for å brukes i behandlingsstudier i primærhelsetjenesten (Petersen et al., 2003).

Klassifikasjonssystemet til Petersen et al. er bygget opp av 13 hovedkategorier. Kategori 1-11 identifiserer ulike smerte-produserende strukturer i korsryggen. Kategori 12 identifiserer ryggsmerter av ikke organisk type der pasienten viser tegn til somatisering. Kategori 13 er pasienter som ikke kan plasseres i noen av de 12 tidligere gruppene (Petersen et al., 2003). Etter hvert som man ut fra kliniske effektstudier finner hvilken behandling som er mest effektiv for hver av de ulike

diagnostiske kategoriene, er hensikten med systemet å hjelpe terapeuter til å finne riktig behandling.

O'Sullivan

O'Sullivan publiserte i 2005 et nytt klassifikasjonssystem for pasienter med uspesifikke kroniske rygg smerter. Klassifikasjonssystemet har sin base i QTFC. Hensikten med klassifikasjonssystemet er å gi retningslinjer om behandlingsstrategier. Klassifiseringen bygger på anamnese, klinisk presentasjon og fysisk undersøkelse. De siste to årene har systemet blitt videre utviklet og i dag inngår også spørreskjema for angst, depresjon og fear-avoidance-beliefes (Vibe et al., 2008). Systemet til O'Sullivan har en biopsykososial tilnærming og er det eneste av de behandlingsrelaterte systemene som tar hensyn til at psykologiske faktorer kan bidra til utvikling og opprettholdelse av kroniske smerter (O'Sullivan, 2005; Billis et al., 2007; Dankaerts et al., 2006; Vibe et al., 2008). Ut fra forskjellige underliggende mekanismer som gir endret motorisk respons, deles pasienten med kroniske korsrygg smerter i grove trekk inn i tre hovedgrupper (figur 2):

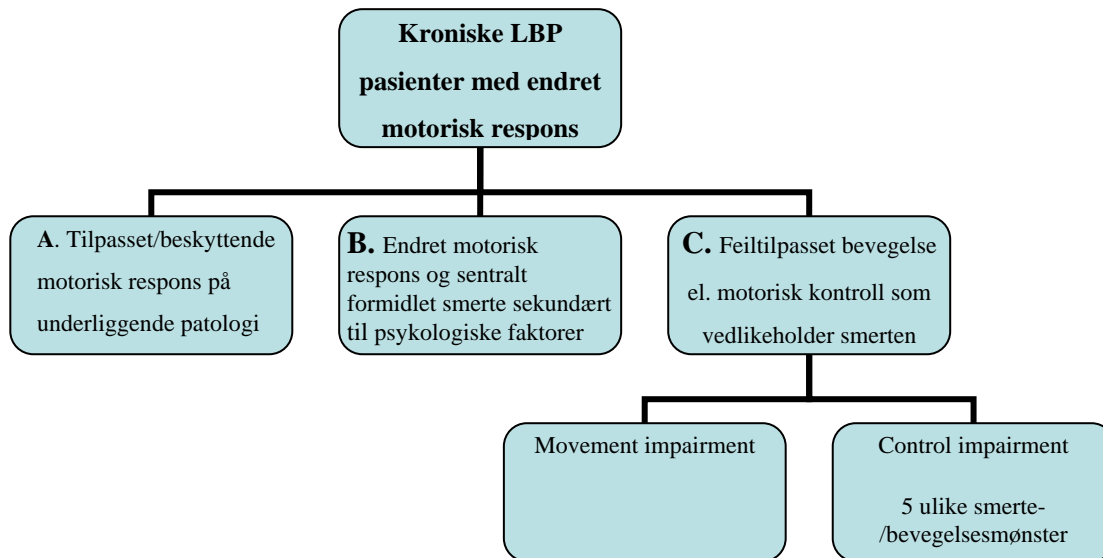
A. Forandring av motorisk respons grunnet underliggende patologi, såkalte adaptive. Herunder ligger "Røde flagg" og spesifikke korsryggdiagnoser.

B. Atferdbasert forandring av motorisk respons.

Herunder ligger uspesifikke korsrygg smerter med sentralt mediert rygg smerte så kallet "Gule flagg".

C. Forandring av motorisk respons grunnet feiltilpasset bevegelsesmønster, så kalte maladaptive.

Herunder ligger uspesifikke korsrygg smerter med perifert mediert rygg smerte.



Figur 2. Subgruppering av kroniske korsryggsmerter O'Sullivan 2005 (O'Sullivan, 2005)

Den siste gruppen C deles videre opp i to kategorier: 1) nedsatt bevegelighet i korsryggen, *movement impairment (MI)* og 2) nedsatt motorisk kontroll av korsryggen, *motor control impairment (MCI)*. For begge gruppene tar man hensyn til dysfunksjon og smerteprovoserende bevegelsesretning. Gruppen med control impairment (MCI) er delt i 5 undergrupper etter smerteprovoserende bevegelsesmønstre: *fleksjonsmønstre*, *fleksjonsmønstre med shift*, *aktivt ekstensjonsmønstre*, *passivt ekstensjonsmønstre* og *multidireksjonalt mønstre*.

Det er laget kriterier for hva som kjennetegner de ulike mønstrene (vedlegg 6) (Dankaerts et al., 2006; O'Sullivan, 2005). O'Sullivan har laget sitt system ut fra at hver kategori og undergruppe krever sin spesifikke behandling. Behandlingene i de ulike undergruppene går først og fremst ut på motorisk læring, den er behovs tilpasset og krever pasientens forståelse, medvirkning og oppfølging. Psykososiale faktorer som i variert grad påvirker pasientens smertemønstre blir lagt inn i behandlingsstrategien. Klassifikasjonssystemet har de seneste årene vært under utvikling og har etter at denne studien startet blitt utviklet og nye kategorier har kommet til. Blant annet har man delt gruppen med uspesifikke perifert medierte ryggsmerter i to grupper; en for korsryggsmerter og en for bekkensmerter. Klassifikasjonssystemets teoretiske grunnlag og behandlingsprinsipper blir

grundigere presentert sammen med en oppdatert oversikt over klassifikasjonssystemets kategorier og subgrupperinger fra den nylig publiserte artikkelen til Vibe Fersum i (vedlegg 5) (Vibe et al., 2008).

2.2.4 Oppsummering av presenterte klassifikasjonssystemer

En oppsummering av hovedtrekkene ved de presenterte klassifikasjonssystemene gis i tabell 1.

Alle systemene unntatt QTFC og Petersen er behandlingsrelaterte. Petersen og QTFC er mer diagnostisk og prognostisk rettet.

McKenzy og Delitto' klassifikasjonssystemer er hovedsakelig rettet mot akutte og subakutte korsryggsmerter, O'Sullivan og Sharmann fremst mot uspesifikke kroniske korsryggsmerter og QTFC og Pedersen omfatter både akutte og kroniske korsryggsmerter.

QTFC var tidlig ute med å inkludere psyko-sosiale faktorer, men først etter at plagene hadde vart en tid. Delitto tar hensyn til psyko-sosiale faktorer i sitt modifiserte klassifikasjonssystem, mens O'Sullivan er det eneste systemet som i Billis artikkel blir presentert som et biopsykososialt klassifikasjonssystem (Billis et al., 2007; George et al., 2003)

Tabell 1. Hovedtrekk ved ulike klassifikasjonssystemer (Billis et al., 2007; Delitto et al., 1995; Maluf et al., 2000; O'Sullivan, 2005; Riddle, 1998; The Quebec Task Force on Spinal Disorders, 1987)

Klassifikasjons System	Klassifikasjons Domene	Hensikt	Subgrupper NSLBP	Klassifikasjons Kriterier	Klassifikasjons Paradigme
QTF – Quebec Task Force	LBP akutt og kronisk	Diagnose gruppering, prognostisk kvalitetskontr. og forskning	4 mulige grupper baserte på sympt respons på tidl beh., radiolog us, varighet, arbeidsstatus	Anamnese, klinisk presentasjon og medisinske undersøkelse	Biomedisinsk
Mc Kenzie	LBP akutte og subakutte	Kliniske retningslinjer for valg av riktig behandling	Postural, Dysfunksjon Derangement	Anamnese, klinisk presentasjon og undersøkelse	Biomedisinsk
Delitto et al.	LBP akutte og subakutte	Kliniske retningslinjer for valg av riktig behandling	Manipulasjon Spesifikke øvelser Stabilisering Traksjon	Anamnese, klinisk presentasjon og fysiske tester	Biomedisinsk
Sahrmann et al.	NS CLBP	Kliniske retningslinjer for valg av riktig behandling	Fem smerteretninger	Klinisk presentasjon (symptomer) og undersøkelse	Biomedisinsk
Petersen et al.	NS LBP akutt og kronisk	Diagnose, stille prognose og forskning	Smerteprod. strukturer, 11 undergrupper, ikke organisk smertegruppe 12	Anamnese og klinisk undersøkelse	Biomedisinsk
O'Sullivan	NS CLBP	Kliniske retningslinjer for valg av riktig behandling	Control impairment og movement impairment. Basert på bakenforliggende faktor til smerten. Fem smerteretninger/ mønster	Anamnese, klinisk presentasjon og undersøkelse, spørreskjemaer	Biopsyko-sosial

LBP – Low back pain, CLBP – Chronic low back pain, NSLBP – Nonspecific low back pain

2.3. Metoder for å vurdere kvaliteten på et klassifikasjonssystem

Når en ny metode utvikles trenger den å testes for både validitet og reliabilitet. Dette kan ses på som en kontinuerlig flertrinnsprosess. Man begynner alltid med å teste ut reliabiliteten, en metode som ikke er reliabel kan ikke heller være valid (McDowell & Newell, 1996).

2.3.1 Reliabilitet

Reliabilitet kan også beskrives som repeterbarhet, reproducerbarhet eller test-retest. Reliabilitet sier noe om en tests pålitelighet og hvor reproducerbar den er (Peat et al., 2002). Dersom man gjentar en test på samme subjekt ønsker man å få samme svarmønster. Jo mindre variasjon man får frem ved repetererte målinger desto høyere er reliabiliteten (Polit & Beck, 2004).

Man kan undersøke flere komponenter ved reliabilitet. Instrumentetsreliabilitet (instrument reliabilitet), reliabilitet mellom to testere (interrater reliabilitet), reliabilitet ved flere målinger av samme testere (intrarater reliabilitet) eller reliabiliteten ved flere målinger av samme subjekt (intrasubjekt reliabilitet) (Domholdt, 2000). Instrumentets reliabilitet innbefatter både tekniske måleinstrument, spørreskjemaer og observasjons målinger. Ved den sistnevnte er det testeren som er selve instrumentet. Ved både interrater og intrarater reliabilitet vil opphavet til variasjon kunne ligge hos både testere og objekt (Domholdt, 2000). Ved bruk av hjelpemidler som video kan objektets variasjoner fjernes og man ser da utelukkende på testernes variasjoner. Noen målinger innenfor fysioterapi kan være lite reliable grunnet fenomenets stadige variasjoner, som for eksempel kronisk smerte. De fleste test- retest målinger er utsatt for en kombinasjon av instrumentfeil, testers feil og subjektets variasjon (Domholdt, 2000).

Mange faktorer kan spille inn og påvirke reliabiliteten og det er ved planlegging av en reliabilitetsstudie viktig å ta høyde for disse faktorer. Variasjonen kan minskes ved

å standardisere testene og trene terapeuter til å utføres tester likt (Strenger et al., 1997; Domholdt, 2000).

Ved reliabilitet er tanken at en måling er reliabel når den kan reproduseres så nært opp til første måling som mulig. Dette måles med en eller annen form for korrelasjons koeffisient. Ved kontinuerlige data brukes Pearson product moment korrelasjon eller Intraclass korrelasjon. Ved rangerte kontinuerlige data brukes Spearman rank order correlation eller Kendall's tau. Ved kategoriske data brukes Cohen's kappa (Domholdt, 2000). Korrelasjons koeffisient på 1.0 indikerer en perfekt samhörighet mellom to målinger. Man kan diskutere hvor mye mindre enn 1.0 korrelasjonen kan være for å kalles reliabel. Man har prøvd å sette tallgrenser og beskrive reliabiliteten med ord, men ulike typer korrelasjons koeffisienter og matematiske formler gjør det vanskelig å lage den type grenseverdier (Strenger et al., 1997).

For å styrke reliabilitetsinformasjonen er det viktig også å beskrive omfanget av variasjon på en repetert måling og dette måles med standard error of measurement (SEM). Det er ikke alltid at en høy korrelasjonskoeffisient er assosiert med liten SEM eller lav korrelasjonskoeffisient med stor SEM. Det er derfor viktig å få med informasjon om begge disse i en presentasjon av resultater (Altman, 2000; Domholdt, 2000).

2.3.2 Validitet

Høy reliabilitet er nødvendig, men ikke tilstrekkelig for høy validitet. En reliabel måling blir altså valid dels ved å være reliabel og dels ved å frembringe meningsfull informasjon. Det er viktig å måle det man har til hensikt å måle (Domholdt, 2000).

Det finnes to distinkte former av validitet, intern validitet og ekstern validitet. En studie har intern validitet om målingene og metodene er korrekte og reliable, om målingene måler det de skal måle og om feilen mellom observatørene og innenfor

subjektene er små.

De mest vanlige former for intern validitet er logisk validitet (face validity), konstruksjonsvaliditet (construct validity), innholdsvaliditet (content validity), og kriterievaliditet (criterion validity).

Logisk validitet viser i hvilken utstrekning en metode måler det den har til hensikt å måle. Denne type validitet vurderer om et instrument/spørreskjema oppfyller sin hensikt med å samle inn eksakt og korrekt informasjon om karakteristika, sykdom eller eksponering av et subjekt (Peat et al., 2002). *Konstruksjonsvaliditeten* viser til den teoretiske basisen som ligger til grunn for en test eller en måling. Det er viktig å være klar over hvilket fenomen man ønsker å måle og definere begrepene nøyaktig og gjøre dem målbare (Domholdt, 2000). Konstruksjonsvaliditeten viser i hvilken grad en test samstemmer med en annen test på forventet måte. Den benyttes ofte i situasjoner når en gullstandard ikke er tilgjengelig (Peat et al., 2002).

Innholdsvaliditet viser i hvilken grad delene i en test eller en måling er et representativt utvalg for det som skal måles. (Domholdt, 2000; Peat et al., 2002).

Kriterievaliditet viser i hvilken grad en måling er systematisk relatert til andre valide målinger og resultat. Kriterievaliditet brukes når man har en gullstandard å sammenligne med som måler det samme (Domholdt, 2000; Peat et al., 2002).

Korrelasjons koeffisient brukes ofte ved undersøkelse av kriterievaliditet.

Ekstern validitet sier noe om resultatene fra en klinisk kontrollert studie kan appliseres på en større populasjon, såkalt generaliserbarhet. I kliniske forsøk skal den eksterne validiteten defineres ved at man har klare inklusjons- og eksklusjonskriterier (Peat et al., 2002). God ekstern validitet avhenger også av at studien er randomisert. Generaliserbarheten bedømmes subjektivt (Peat et al., 2002).

2.3.3 Kriterer som bidrar til reliabilitet og validitet av klassifikasjonssystemer

Buchbinder et al. har utarbeidet et metodisk grunnleggende verktøy som inneholder sju seksjoner for å systematisere og standardisere vurderingen av

klassifikasjonssystemer (Buchbinder et al., 1996) Følgende seksjoner er satt opp og har underrubrikker med spørsmål som klassifikasjonssystemet bør kunne svare tilfredsstillende på:

1. *Hensikt*: Er hensikten, populasjonen og setting klart spesifisert?
2. *Innhold validitet (content validity)*: Er domenet og alle spesifikke eksklusjoner fra dette domenet klart spesifisert? Er alle relevante kategorier inkludert? Er oppdelingen av kategorier passende i forhold til hensikten? Er kategoriene innbyrdes (gjensidig) utelukkende? Var utviklingsmetoden passende? Om den er multiaksial, er da kriteriene for innholdsvaliditet tilfredsstillende for hver enkelt akse?
3. *Logisk validitet (face validity)*: Er merkingsnomenklaturen på kategoriene tilfredsstillende? Er kriteriene for å fastslå inklusjon i hver kategori klart spesifisert? Hvis ja, virker disse kriteriene relevante? Har kriteriene vist seg å være valide? Har kriteriene vist seg å være reliable? Er definisjon på kriteriene klart spesifisert? Om systemet er multiaksialt, er kriteriene for kriterievaliditet tilfredsstillende for hver tilleggsakse?
4. *Klinisk brukbarhet (feasibility)*: Er klassifikasjonen enkel å forstå? Er klassifikasjonen enkel å utføre? Består den av bare klinisk undersøkelse? Kreves det spesielle egenskaper, kunnskaper, verktøy eller utdanning? Hvor lang tid tar det å utføre?
5. *Konstruksjons validitet (construct validity)*: Skiller systemet mellom grupper (entities) som er ment å være ulike, på en hensiktsmessig måte? Fungerer systemet tilfredsstillende sammenlignet med andre system som klassifiserer samme diagnose?
6. *Reliabilitet*: Gir klassifikasjonssystemet konsekvente resultater når man klassifiserer under de samme betingelsene (test – retest)? Er intra- og interaterer reliabiliteten tilfredsstillende?
7. *Generaliserbarhet*: Har det blitt brukt i andre studier og/eller andre sammenhenger?

(Buchbinder et al., 1996)

I to review artikler, en av Riddel 1998 og en av Petersen 1999, ble de vanligst brukte og mest dokumenterte klassifikasjonssystemer for uspesifikke korsryggsmerter analysert med hensyn til Buchbinders sju kriterier (Petersen et al., 1999; Riddle, 1998). Blant disse finner man blant annet Mc Kenzie, Quebec Task Force og Delitto som er beskrevet her i tidligere avsnitt. Ingen av klassifikasjonssystemene oppfylte alle kriteriene. Spesielt dårlig rapportert var reliabiliteten, konstruksjonsvaliditeten og generaliserbarheten (Riddle, 1998).

2.3.4 Reliabilitet og validitet av klassifikasjonssystemer for LBP

En gullstandard for klassifikasjonssystem for korsryggsmerter savnes fortsatt, de fleste mest brukte klassifikasjonssystemene har svak dokumentasjon for reliabilitet og validitet (Billis et al., 2007; McCarthy et al., 2004; Petersen et al., 1999; Riddle, 1998). I 2007 gjorde Billis og medarbeidere en analyse av klassifikasjonssystemer som hittil er publisert for å blant annet å se på kulturelle forskjeller ved systemene. Blant disse finnes også klassifikasjonssystemene som er beskrevet i denne oppgaven. Hun fant at det fortsatt er for svak reliabilitet, innholds (content) validitet og generaliserbarhet blant systemene. Det er avgjørende at det er stor enighet blant klinikerne ved bruk av et klassifiseringssystem og at det trengs adekvate sterke reliabilitetsstudier (Billis et al., 2007). For å tilfredsstille innholds (content) validitet er det foreslått at en biopsykososial tilnærming skal inngå ved klassifikasjon av LBP pasienter. Billis understreker i sin oversikt at man også må styrke klassifikasjonssystemene ved å tilføre sosialkulturelle faktorer til sin klassifisering (Billis et al., 2007).

Jeg gir under en oversikt over de tidligere presenterte klassifikasjonssystemenes reliabilitet og validitet. I tabell 2 er klassifikasjonssystemenes reliabilitet og validitet oppsummert.

Quebec Task Force Classification system

Det er ikke gjort noen reliabilitetsstudier på hele QTFC systemet.

Klassifikasjonssystemet er meget omfattende og er derfor brukt og vurdert i forenklete versjoner. Studier på systemets prediktive og diskriminerende evne har blitt gjort, men ikke for alle kategorier i systemet (Atlas et al., 1996; BenDebba et al., 2000; Loisel et al., 2002). I artikkelen til Billis og medarbeider fra 2007 har QTFC som vist i tabell 6 en partiell reliabilitet, men forfatteren skriver samtidig i artikkelen at det ikke er gjort noen reliabilitetsstudier på systemet. Kritisk gjennomgang av QTFC systemet viser at det møter krav om målsetning, logisk validitet og generaliserbarhet og møter delvis krav om innholdsvaliditet, klinisk anvendbarhet og begrepsvaliditet (Billis et al., 2007).

Mc Kenzie

I 2005 publiserte Clare og medarbeidere en reliabilitetsstudie på Mc Kenzie klassifikasjon på både nakke- og rygg smerter (Clare et al., 2005). I studien refererer de til fire tidligere studier som har testet systemets reliabilitet og understreker at den svake reliabiliteten skyldes flere faktorer. I to av studiene var terapeutene dårlig trent i Mc Kenzie systemet, og i tre av studiene var det få terapeuter, hvilket minsker generaliserbarheten (Kilpikoski et al., 2002; Riddle & Rothstein, 1993; Kilby et al., 1990). I to av studiene brukte man ikke sjanskorrigerende statistikk og i en studie brukte man ikke systemet på en korrekt måte. Clare og medarbeidere ønsket å lage en studie som unngikk disse svakhetene (Clare et al., 2003). De ble inkluderte 25 LBP og 25 nakkesmerte pasienter. To terapeuter (totalt 7 par), som var godt trent i Mc Kenzie metoden, undersøkte sammen pasienten og scoret deretter blindet fra hverandre. Studien viste høy reliabilitet både ved syndrom klassifisering (3 syndrom) og subsyndrom klassifisering (20 forskjellige) med en kappa på 1.0 og 0.89 for LBP, men med lavere kappa for nakkeklassifisering, 0.63 respektive 0.84 (Clare et al., 2003). Billis og medarbeidere mener i sin artikkel at Mc Kenzie systemet er partielt reliabelt. Kritisk gjennomgang av Mc Kenzie systemet viser at krav blir møtt når

det gjelder anvendbarhet, begrepsvaliditet og generaliserbarhet. Krav blir delvis møtt om innholdsvaliditet og logisk validitet (Billis et al., 2007).

Delitto et al.

Klassifikasjonssystemet til Delitto et al. har blitt reliabilitetstestet i 2004 av Heiss og medarbeidere (Heiss et al., 2004). Systemet har tidligere vært testet i to studier, en i 1992 og en i 2000 (Delitto et al., 1992; Fritz & George, 2000). Hensikten med studien til Heiss var å finne nøkkelpunkter ved uenighet for å kunne utvikle systemet og øke dets reliabilitet. Terapeutene som skulle teste systemet leste en artikkel og fikk en dags kurs i bruk av systemet. Førstifem pasienter med akutt LBP ble klassifisert. Terapeutene som var randomiserte i par undersøkte pasienten (Heiss et al., 2004). Kappa verdien ble 0.14 ved syndromklassifisering respektive 0.15 ved klassifisering til fire behandlingsgrupper. Den endelige kappa verdien ble regnet ut for tre av terapeutene da den fjerde hadde gjort feil i en av kategoriene. De konkluderte med en interrater kappa på 0.45. De to tidligere studiene viste kappa verdier på -0.1 og 0.53 respektive 0.49 og 0.56 (Heiss 2004). I artikkelen til Billis har systemet til Delitto ikke møtt krav til reliabilitet. Kritisk gjennomgang viser at systemet møter krav om målsetning og anvendbarhet og delvis møter krav om innhold og logisk validitet samt generaliserbarhet (Billis et al., 2007).

Sahrman, Maluf og Van Dillan

Systemet til Sahrman ble i 1998 reliabilitets testet av Van Dillen et al. (Van Dillen et al., 1998). Man undersøkte nittifem pasienter med LBP og 43 personer uten LBP. Man valgte å ta med matchede kontroller uten LBP for å kunne teste spesifisitet og sensitivitet. Fem terapeuter undersøkte pasientene i randomiserte par. I hvert par undersøkte den ene terapeuten mens den andre observerte. Terapeutene var alle godt trent i systemet og hadde alle i variert grad vært med på å utvikle systemet. Scoringene ble gjort blindet fra hverandre. Ved data analysen brukte man vektet enighetsprosent og vektet kappa. Man kom frem til en kappa som varierte fra 0.87 til 1.00 for symptom klassifisering og 0.00 til 0.78 for holdnings- og

bevegelsesklassifisering (Van Dillen et al., 1998). Systemet ble også reliabilitetstestet i 2002 av White og Thomas som fikk en lavere kappa 0.02 til 0.62 (White & Thomas, 2002). I artikkelen til Billis beskrives reliabiliteten på Sahrman's system som partiell. Systemet møter krav om målsetning og klinisk anvendbarhet, den møter delvis krav om innhold, logisk og begrepsvaliditeten (Billis et al., 2007).

Petersen et al.

I 2004 publiserte Petersen og medarbeidere en interrater reliabilitetsstudie på sitt nyutviklede klassifikasjonssystem (Petersen et al., 2004). Nitti pasienter inngikk i studien og hver pasient ble undersøkt av to fysioterapeuter. Til sammen 4 fysioterapeuter deltok i studien. Terapeutene klassifiserte blindet fra hverandre. Resultatene viser at reliabiliteten for hele systemet er lav med 39 % enighet, men at det er moderat til god reliabilitet i de enkelte kategoriene med kappa mellom 0.44 - 1.00. To av kategoriene har kappa lavere enn 0.40 (Petersen et al., 2004). Klassifikasjonssystemet er i Billis artikkel vurdert som partielt reliabel. Kritisk gjennomgang av systemets kvalitet viser at krav om målsetning blir møtt, men krav om anvendbarhet, innhold og logisk validitet og om generaliserbarhet blir bare delvis møtt. Systemet har ikke møtt krav om konstruksjonsvaliditet (Billis et al., 2007).

O'Sullivan

En interrater reliabilitetsstudie har blitt publisert i 2006 av Dankaerts et al. (Dankaerts et al., 2006). Den bestod av to delstudier. Den første var en interraterreliabilitets test mellom to meget erfarne terapeuter hvor av den ene var systemets utvikler. Pasienter med kroniske uspesifikke korsryggmerter ble rekruttert fra en privatklinikk og inkludert etter gitte kriterier der man blant annet bare inkluderte pasienter med motorcontrol impairment (MCI). 35 pasienter ble inkludert og klassifisert av de to terapeutene blindet for hverandres scoringer. Resultatene viste nesten perfekt enighet mellom de to med kappa på 0.96 og en enighets prosent på 97% (Dankaerts et al.,

2006). Scoringene ble brukt som gullstandard i den andre delen av studien. 25 pasienter ble randomisert fra den første delen av studien for å delta i den andre delen av studien (Dankaerts et al., 2006). Pasientene ble filmet og smerteanamnese tatt opp. 13 terapeuter deltok i denne studien. Gruppen med terapeuter kom fra Norge og Australia og ble delt i to grupper etter erfaring med systemet. Alle terapeutene fikk tilsendt video og anamnese og klassifiserte uavhengig av hverandre. Resultatet av denne studien viste at alle terapeutene sammen scoret en kappa på 0.61 (0.47 - 0.80) med 70 % (60 - 84%) enighet. De med moderat erfaring med systemet scoret en kappa på 0.55 (0.47- 0.64) med 65% (60 - 72%) enighet og de med lang erfaring i systemet scoret en kappa på 0.71 (0.58 - 0.80) med 78% (68 - 84%) enighet. I alle gruppene var enigheten lavere når man scoret ut fra bare anamnese. Man fant i den blandede gruppen en kappa på 0.32 (0.13 - 0.54) med 48% enighet, i den moderate gruppen en kappa på 0.28 (0.13 - 0.37) med 44% enighet og i den erfarne gruppen en kappa på 0.40 (0.39 - 0.54) med 54% enighet (Dankaerts et al., 2006). Da ikke hele klassifikasjonssystemet ble undersøkt og testet for reliabilitet har Billis i sin oversiktsartikkel satt systemet som partielt reliabelt. Kritisk gjennomgang av systemets kvalitet viser at krav om målsetning og anvendbarhet blir møtt. Krav om innhold og logisk validitet blir delvis møtt. Systemet har ikke møtt krav om begrepsvaliditet og generaliserbarhet (Billis et al., 2007).

Doktorgradstipendiat Kjartan Vibe Fersum ved Universitetet i Bergen har gjort en interrater reliabilitets test av hele O'Sullivans klassifikasjonssystem og en RCT studie med klassifisering og behandling ad modum O'Sullivan. Reliabilitetsstudien som nylig er blitt utgitt (article in press) har vist liknende resultat som Dankaerts studie, men Vibe Fersum har testet en større del av systemet på pasienter med uspesifikke kroniske korsryggmerter og inkluderer pasienter med MI og bekkenrelaterte ryggmerter (Vibe et al., 2008). Vibe Fersum inkluderte sine pasienter fra både primærhelsetjenesten og Rygg-Nakke poliklinikken ved Haukeland Universitets Sykehus. Han inkluderte 26 pasienter etter gitte inklusjons- og eksklusjonskriterier (Vibe et al., 2008). Fire fysioterapeuter med mange års erfaring i undersøkelse og behandling av pasienter med korsryggmerter var eksaminatorer. 3 av terapeutene var manuellterapeuter og den fjerde var Peter O'Sullivan som også utviklet systemet.

Terapeutene ble godt trent i systemet av Peter O'Sullivan før studien. Det ble brukt en test - retest design. O'Sullivan utarbeidet en klassifiseringsmanual. Pasientene gjennomgikk en full anamnese og full fysikalsk undersøkelse av alle fire eksaminatorene. I tillegg brukte man andre medisinske tester og radiologisk bildediagnostikk. Undersøkelsen tok en time og etter undersøkelsen fylte terapeutene ut et klassifikasjonsskjema. Resultatene er bare presentert med prosentvis enighet i de første tre nivåene av systemet (vedlegg 5). Man hadde en prosentvis enighet på 75 % respektive 99 % ved subgruppering til MI og MCI. Av 25 pasienter med uspesifikke korsryggsmerter fant man 24 med MCI og 1 med MI. Ved klassifisering av pasientene til ulike smerteprovoserende retninger innen gruppen MCI fikk man en kappa på 0.82 (range 0.66 - 0.90). I denne studien vurderte og scoret man også grad av psykososiale faktorer som influerte på smertene. Her fant man en kappa på 0.65 (range 0.57 - 0.74) (Vibe et al., 2008).

Tabell 2. Reliabilitet og validitet for ulike klassifikasjonssystemer (Billis et al., 2007; May et al., 2006)

Klassifik. system	Reliabilitet	Validitet					
		Hensikt	Innhold	Logisk	Klinisk brukbar	Konstruksjon	Generaliserbar
QTF Sptizer 1987 Wernecke 2004	Delvis Delvis	Ja Ja	Delvis Delvis	Ja Ja	Delvis Delvis	Delvis Nei	Ja Delvis
Mc Kenzie McKenzy 1981 Razmjou 2000 Kilpikoski 2002 Clare 2005	Delvis κ 0.45 - 0.96 κ 0.88 - 1.00 Delvis κ 0.6/0.7 κ 0.35 - 1.0 κ 0.66 - 1.0	Ja	Delvis	Delvis	Ja	Ja	Ja
Delitto Delitto 1995 George 2005 Fritz 2003 Fritz 2000	Nei Nei Nei κ 0.49/0.56	Ja Nei Ja	Delvis Delvis Delvis	Delvis Ja Ja	Ja Delvis Nei	Nei Delvis Nei	Delvis Delvis Nei
Sahrmann Van Dillan 1998, 2003	Delvis (deler av systemet god reliabilitet) κ 0.0 - 0.78 κ 0.87 - 1.0	Ja	Delvis	Delvis	Ja	Delvis	Delvis
Petersen Petersen 2003, 2004	Delvis (deler av systemet moderat reliabelt) κ 0.26 - 1.0	Ja	Delvis	Delvis	Delvis	Nei	Delvis
O'Sullivan O'Sullivan 2005, Dankaert 2006	Delvis (deler av systemet moderat reliabelt) Kappa 0.47 to 0.80	Ja	Delvis	Delvis	Ja	Nei	Nei

3. Materiale og metode

3.1 Bakgrunn for valg av deltagere

På Ullevål Universitets Sykehus (UUS) startet i 2004 en stor norsk multisenterstudie, videre kalt Skiveprotesestudien, der man ønsket å sammenligne effekten av to forskjellige behandlingsmetoder; skiveprotesekirurgi vs. trening og kognitiv intervensjon på pasienter med uspesifikke kroniske korsryggsmerter.

Da studien ble planlagt var jeg i gang med å ta i bruk O'Sullivan's klassifikasjonssystem og behandlingsopplegg for denne pasientgruppen. Dette systemet var reliabilitetstestet i Australia av miljøet som utviklet det (presentert på kurs og konferanser, men ikke publisert) og jeg ønsket å gjøre det samme i klinikken på UUS. Utvalget som skulle inngå i skiveprotesestudien er en selektert gruppe innen uspesifikke kroniske ryggsmerter, hvor pasientene ble inkludert ut fra grad av degenerative forandringer etter spesielle radiologiske kriterier.

Ved å klassifisere pasientene før intervensjonen gjorde vi det mulig å senere undersøke subgruppens respons på de to intervensjonene og undersøke prediktiv verdi av systemet. Først var det viktig å reliabilitetsteste klassifikasjonssystemet på det selekterte utvalget.

3.2 Utvalget

Pasientene som inngår i min metodestudie ble rekruttert via Skiveprotesestudien (vedlegg 4). Pasientene fikk ved inklusjon til Skiveprotesestudien samtidig informasjon om deltagelse i metodestudien. Dette var frivillig og pasienten godkjente deltakelse ved eget skriv som tilhører metodestudien (vedlegg 3).

Inklusjon til hovedstudien ble utført i fellesskap av leger på Ortopedisk senter (ryggkirurg) og Ryggpoliklinikken (fysikalsk medisinsk lege). Det var en fysikalsk-medisinsk lege og to ortopedkirurger som sto for inklusjonen. For at en pasient skulle

bli inkludert krevde man at både kirurg og fysikalsk-medisiner var enige om inklusjonen. Dette for å sikre en mest mulig homogen populasjon i studien.

3.2.1 Inklusjonskriterier

- Alder 25 – 55 år
- Korsryggsmarter som hovedsymptom med varighet i mer enn 1 år.
- Gjennomgått strukturert fysioterapi eller kiropraktorbehandling uten tilfredsstillende effekt.
- Oswestry score ≥ 30 % poeng
- Degenerative forandringer i mellomvirvelskiven i en eller begge nederste lumbale nivåer. Degenerasjon bedømmes på MR etter følgende kriterier:
 - Høydereduksjon av mellomvirvelskiven med minst 40 %.
 - Modic forandringer type I og/eller II.
 - High Intensity Zone i skiven.
 - Morfologiske forandringer vurdert som endret signaltetthet i skiven

For at man skal kunne klassifisere skiven som degenerativ må enten det første kriteriet alene eller minst to av de andre kriterier være til stede. MR bildene vurderes av to uavhengige observatører.

3.2.2 Eksklusjonskriterier

- Kronisk generalisert smertesyndrom
- Degenerasjon påvist ved bildediagnostikk i mer enn to nivåer. For at et skivenivå skal klassifiseres som normalt må ikke skivehøyden være redusert med mer enn 40 % og ikke ha mer enn et av de øvrige kriterier for degenerativ skivepatologi. Bedømmingen av en normal skive skjer ved to uavhengige observatører. Ved uenighet bedømmes bildene av en tredje observatør og utfallet avgjøres ved simpelt flertall.
- Symptomgivende spinal stenose
- Skiveprolaps eller rescesstenose med nerverotsaffeksjon
- Spondylolyse

- Spondylolisthese, isthmisk
- Artritt
- Tidligere fraktur L1 – S1
- Psykisk eller somatisk sykdom som gjør at pasienten ikke egner seg for et av behandlingsalternativene
- Forstår ikke norsk skriftlig og muntlig
- Medikamentmisbruker
- Osteoporose
- Medfødt eller ervervet deformitet
- Tidligere ryggoperert (med unntak av inntil to prolapsoperasjoner med residiv i aktuelle skivenivå(er))

Ved inklusjon til Skiveprotesestudien ble pasienter med neurologiske utfall, annen spesifikk patologi og ”Røde flagg” ekskludert (Gruppe A, figur 2). Det samme gjelder pasienter med generaliserte smerter og smerter primært relatert til psykososiale faktorer, såkalte ”Gule flagg” (Gruppe B, figur 2). De inkluderte pasientene blir derfor en begrenset populasjon innenfor gruppen med kroniske uspesifikke korsryggsmerter, inkludert fremst ut fra radiologiske kriterier og funksjonsnivå motvarende ≥ 30 % på Oswestry 2.0 (Gruppe C, figur 2).

3.3 Terapeutene

To fysioterapeuter var hovedansvarlige for metodestudien. Fysioterapeutene hadde 20 respektive 15 års erfaring med kroniske muskelskjelettlidelser med hovedvekt på pasienter med kroniske korsryggsmerter. Terapeutene har arbeidet de siste 10 årene på samme fysikalsk medisinske avdeling med de samme pasientgruppene. De har deltatt på 3 av kursene til Peter O’Sullivan som har blitt holdt i Norge. Artikler og beskrivelse av klassifikasjonssystemet er blitt nøye gjennomgått. Terapeutene har hatt

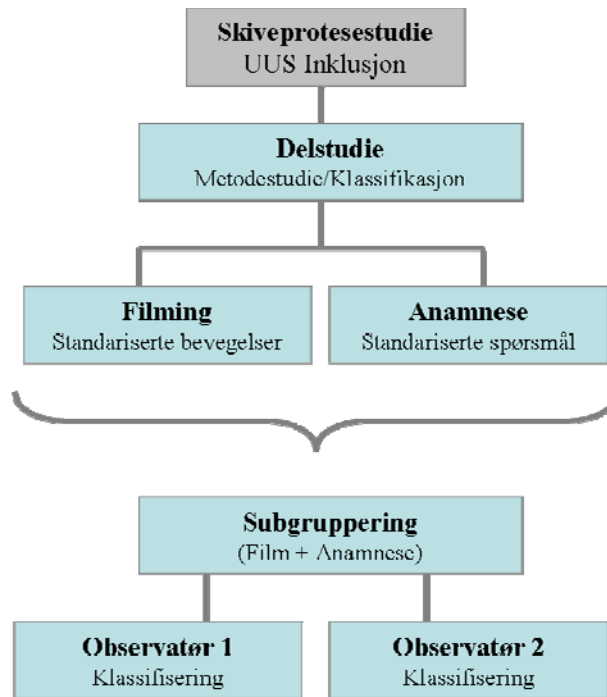
fire møter med to andre erfarne terapeuter som også har gått de samme kursene og sammen vurdert og diskutert pasienter ut fra O'Sullivan's klassifikasjonssystem.

3.4 Studiedesign og prosedyrer

Denne studien er en metodestudie der et klassifikasjonssystem, basert på bevegelsesanalyse og smerteanamnese, har blitt reliabilitetstestet på en selektert gruppe pasienter med uspesifikke kroniske korsryggsmerter etter gitte inklusjons- og eksklusjonskriterier.

Pasienter inkludert i Skiveprotesestudien som samtykket til å delta i denne metodestudien ble, før behandlingsstart i Skiveprotesestudien, innkalt til video- og smerteanamnese opptak (figur 3). Bevegelsesundersøkelsen baseres på O'Sullivan's klassifiseringssystem, der man vurderer hvilke muskulære strategier pasienten bruker i lumbal-/bekken regionen under bevegelse og i hvilke bevegelsesretninger de får smerte (vedlegg 5 og 6). Anamnesen fokuserer fremst på bevegelser, bevegelsesretninger og stillinger som gir pasienten smerter/smertelettelse (vedlegg 8). Fysioterapeutene vekslet på å foreta anamnese- og videoopptak av pasientene.

I forkant av metodestudien ble 10 pasienter brukt som pilotgruppe for å trene på filmprosedyren og anamneseopptak, og for å diskutere og løse eventuelle problemstillinger rundt dette. Pilotgruppen ble også brukt for å trene på samkjøring i subgrupperingene. Pilotpasientene er ikke medregnet i resultatene. Basert på videoopptak og smerteanamnese ble det gjort en interraterreliabilitets test mellom 2 terapeuter. Terapeutene klassifiserte pasientene uavhengig av hverandre ut fra videoopptak og smerteanamnese og var blindet for hverandres klassifisering.



Figur 3. Flytskjema

3.4.1 Videofilming av bevegelsesanalysen

Klassifiseringen baseres delvis på videoopptak av pasientenes bevegelser. Dette er de samme bevegelsene som O'Sullivan har brukt i sine undervisningsfilmer og senere Dankerts har brukt ved sin reliabilitetsstudie (Dankaerts et al., 2006). Før filmingen gikk vi igjennom hvilke bevegelser pasienten skulle gjøre og hver bevegelse ble så filmet på avstand og i nærbilde. Dette for å se helhetsbevegelsen første gang med kompensatoriske bevegelser og deretter de lumbale segmentene i bevegelse. Det er laget egne skjema for registrering av de ulike bevegelsene og eventuell smerte under utførelse av bevegelsene ved filmingen (vedlegg 7).

I tråd med vanlig klinisk praksis var pasienten kun iført undertøy under utføring av bevegelsene som ble filmet. Pasientene ble filmet fra siden og bakfra. De er derfor normalt sett vanskelig å identifisere. Det var imidlertid et tilbud om anonymisering ved at man fikk utlevert briller og hette ved filmingen.

Filmingen tok ca. 20 minutter og anamneseopptak ca. 20 minutter. Terapeutene fordelte oppgavene etter hvordan det passet tidsmessig. Valg av videofilm hadde også tidsmessige aspekter for terapeutene som skulle utførte testen i sin kliniske hverdag. Av praktiske grunner filmet terapeutene og tok opp anamnese etter hvem som hadde mulighet i forhold til at dette ble gjort på dagtid når man normalt behandlet pasienter i klinikken. I praksis ble det derfor slik at noen ganger tok samme terapeut opp både anamnese og videofilm, andre ganger filmet den ene og den andre tok opp anamnese. Det ble ikke gjort forsøk på å subgruppere under filming og anamneseopptak. Terapeutene hadde ikke lov å diskutere pasientene seg imellom.

Følgende bevegelser filmes:

<i>Fra siden</i>
Stående holdning
Bekkentilt
Fremoverbøyning
Oppreising fra fremoverbøy
Bakoverbøy
Knebøy
Sittende holdning
Sittende krum-svai
Sittende til stående

<i>Bakfra</i>
Stående holdning
Bekkentilt
Fremoverbøyning
Oppreising fra fremoverbøy
Bakoverbøy
Sidebøy
Etbenstående(begge ben)
Knebøy
Sittende holdning
Sittende krum-svai
Sittende til stående

3.4.2 Anamnese

O’Sullivan legger stor vekt på anamnesen ved subgruppering. Han er opptatt av å finne bakenforliggende årsaker til at smerten vedlikeholdes (O’Sullivan, 2005). Anamnesen er en måte å få frem pasientens mer helhetlige smertebilde, funksjonsproblemer og ressurser. Ved anamneseopptak brukes den samme malen som har blitt presentert i artikler og kurs til O’Sullivan (vedlegg 8). Terapeuten skrev

ned pasientens svar skriftlig. For å standardisere brukte vi hovedrubrikkene i skjemaet som spørsmål. Dette gikk vi nøye igjennom før vi startet studien.

3.5 Datainnsamling

3.5.1 Demografiske data

Demografiske data og et bredt utvalg av effektvariabler ble innsamlet gjennom Skiveprotesestudien, kodet og lagt inn i SPSS sentralt ved St.Olav Hospital i Trondheim. Etter at alle pasienter var inkludert fikk jeg tilgang til disse dataene for pasientene inkludert i min metodestudie. Jeg valgte ut relevante data for å gi en deskriptiv beskrivelse av pasientpopulasjonen i min metodestudie og lagret dem i egen SPSS fil.

Funksjonsstatus blir vurdert ut fra Oswestry 2.0 (Baker et al., 1990). Oswestry 2.0 er et spørreskjema som er vanlig brukt ved ryggstudier og som skal måle hvordan ryggplagene påvirker pasientens evne til å klare seg i dagliglivet. Skjemaet er oversatt til norsk og validert av Grotle et al. i 2003 (Grotle et al., 2003). Skjemaet består av 10 spørsmål som hver har 6 svars alternativ. Pasientens scoring regnes om til en skala fra 0-100 % hvor 0 % betyr ingen funksjonsbegrensning og 100 % betyr maksimal funksjonsbegrensning (vedlegg 10).

Smerte måles med VAS; visuel analog scale (Karoly & Jensen, 1987). VAS er delt i to målinger, grad av ryggsmerte og grad av bensmerte. Pasienten ble bedt om krysse av på en 10 cm lang linje mellom to ytterpunkter: ingen smerte --- verst tenkelige smerte.

Utrygghet og redsel forbundet med ryggplagene måles med FABQ; fear avoidance beliefs questionnaire (Waddell et al., 1993). FABQ er et spørreskjema som ble utviklet av for å kvantifisere i hvor stor grad pasientens tro på at fysisk aktivitet og arbeide kan gi økt smerte og risiko for tilbakefall. Spørreskjemaet er delt i to hovedgrupper: FABQ-W som relaterer spørsmål til arbeidet og FABQ-PA som relaterer spørsmål til fysisk aktivitet. Skalan for FABQ-W går fra 0-42, og skalan for

FABQ-PA fra 0-24. For begge skalaene indikerer 0 at pasienten har liten tro på at arbeide eller fysisk aktivitet gir økt smerte. Den norske versjonen av FABQ har blitt utarbeidet og validert av Grotle et al. i 2006 (Grotle et al., 2006) (vedlegg 11).

Angst, depresjon og somatisering måles med HSCL-25 (Derogatis et al., 1974). HSCL-25 er et spørreskjema som ble utviklet for å kvantifisere grad av angst, depresjon og somatisering. HSCL-25 består av 25 spørsmål som kartlegger angst, depresjon og somatisering. Skjemaet indikerer uspesifikke psykosomatiske symptomer og plager og ikke nødvendigvis en psykiatrisk diagnose. Pasientens svar scores på en skala fra 1- 4 hvor 1=ingen symptomer og 4=svært mye symptomer (vedlegg 12).

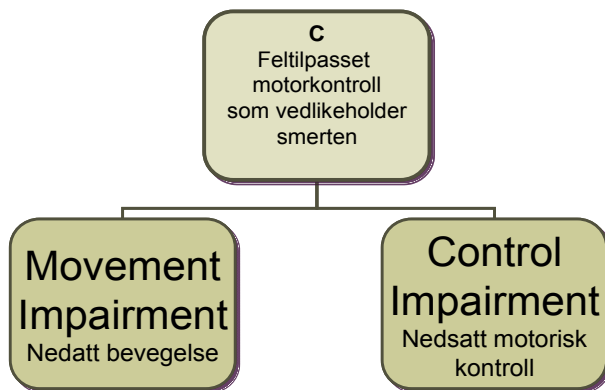
3.5.2 Randomisert rekkefølge ved subgruppering av utvalget

Før selve klassifiseringen begynte ble rekkefølgen på pasientene randomisert ved loddtrekking gjort av uavhengig person. Det er satt opp en rekke for terapeut 1 og en annen rekke for terapeut 2. Dette ble gjort for å unngå å huske rekkefølger av pasienter ved eventuell intrarater reliabilitets test senere. Man ville også redusere eventuell læringseffekt dvs. at vurderingen av filmene ble endret underveis etter hvert som erfaringen til terapeutene økte.

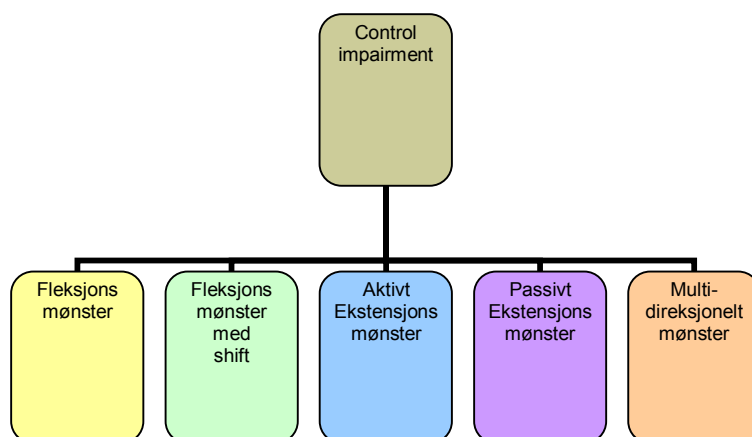
3.5.3 Subgruppering

Filmen og anamnesen danner sammen grunnlaget for å kunne velge subgruppe. Pasientene i metodestudien subgrupperes først til to hovedgrupper; den ene med nedsatt bevegelse, *movement impairment* og den andre med nedsatt motorisk kontroll, *control impairment* (Figur 4). Deretter subgrupperes de videre i gruppen under nedsatt motorisk kontroll (Figur 5) til 5 ulike smerteprovoserende bevegelsesstrategier: *fleksjonsmønster*, *fleksjonsmønster med shift*, *aktivt ekstensjonsmønster*, *passivt ekstensjonsmønster* og *multidireksjonalt mønster* (vedlegg 6). Ved subgrupperingen bruker vi samme skjema som O'Sullivan, men har lagt til en kolonne for å krysse for om vi er sikre eller usikre på vår klassifisering

(vedlegg 9). Resultatene ble lagt i lukkede konvolutter og en uavhengig medarbeider har deretter plottet inn klassifiseringsresultatene i SPSS.



Figur 4. Subgruppering fra feiltilpasset motorkontroll til movement impairment eller control impairment



Figur 5. Subgruppering fra control impairment til fem ulike smerteprovoserende bevegelsesmønstre

3.5.4 Subgruppering ved bare anamnese

Etter at alle pasientene var klassifisert ut fra anamnese og film ble det gjort ytterligere en klassifisering ut fra bare anamnese. Begge terapeutene scoret, blindet fra hverandre, mønstre ut fra bare anamnesen. Tidsmessig hadde vi ikke anledning til å også gå igjennom alle filmene ytterligere en gang. Grunnen til at jeg valgte å se på anamnesen separat var at dette også var gjort i Danaerts studie og det ville da være mulig å jevnføre dette i studiene (Dankaerts et al., 2006).

3.6 Databehandling

Alle data ble lagt inn på SPSS (Statistical Package for Social Science versjon 14) og koblet til registreringsnummer. Statistiske tabeller er laget ut fra resultatene av utregningene i SPSS.

3.6.1 Variabler

Populasjonens alder, kjønn, smertestatus, funksjonsstatus, sykemeldingsstatus, bevegelsesredsel, angst og depresjon ved inklusjon blir beskrevet som deskriptiv statistikk. Variablene som fremkommer ved klassifiseringen er kategoriske nominale.

3.6.2 Statistiske metoder

1) Reliabilitet

Kappa statistikk: For å kartlegge enighet mellom 2 testere brukes Cohen's Kappa. Ved interraterreliabilitet blir grad av enighet mellom testere vurdert ved hjelp av prosentvis enighet og kappakoeffisient (k) med 95 % konfidens intervall (1.96 fra mean). Konfidens intervallet er viktig for å vise spredningen. Kappa kontrollerer for sannsynligheten å få enighet ved sjanse og indikerer om enigheten er bedre enn ved sjanse. Kappa koeffisient brukes for å teste enighetsgrad ved kategoriske nominale variabler. Kappakoeffisienten k har et maksimum på 1.00 og et minimum på 0.00. En kappa på 1.00 indikerer en helt perfekt enighet og en kappa på 0.00 indikerer at resultatet ikke er bedre enn ved sjanse. Ved en negativ kappa er enigheten dårligere enn ved sjanse. Kappaverdiene mellom 1.00 og 0.00 kan tolkes etter følgende tabell.

k verdi	enighetsgrad
< 0.20	dårlig
0.21-0.40	svak
0.41-0.60	moderat
0.61-0.80	god
0.81-100	svært god

Tabell 3. Gradering av Kappa verdier (Altman, 2000; Landis & Koch, 1977)

Dette er retningslinjer og gir bare en pekepinn. Ved utregning av kappa må den oppnådde verdien vurderes opp mot på forhånd satte krav. Hvilken grad av enighet man lavest kan akseptere må stå i forhold til hvilket fenomen man måler.

2) Sammenheng – forskjeller i proporsjoner

For å undersøke om *uenighet* og *usikkerhet ved scoring* har noe sammenheng brukes Kji- kvadrat test for uavhengighet. Variablene som skal undersøkes for sammenheng er kategoriske: usikker/sikker og enig/uenig.

Kji-kvadrat (chi-square) test for uavhengighet (independence) brukes ved mulighet for 2 x 2 tabell. Det er en ikke- parametrisk hypotese test, en eksakt likverdighetstest for å sammenligne to forhold (mål). Testen ser på proporsjonen mellom de ulike alternativene som faller i ulike kategorier når en frekvens tabell har blitt laget. Ved mindre antall N tenderer x^2 verdiene å bli for store og dette korrigeres ved bruk av sammenhengs korreksjon (continuity correction så kallet Yates correction). Yates correction tar ikke bort kravet til størrelsen på forventete frekvenser. Fisher`s Exakt test skal derfor brukes ved få observasjoner (totalt $N = 30$ eller færre) eller hvis en eller flere av cellene i en 2 x 2 tabell har forventet verdi på < 5 .

Observed frequencies er de antall som faller i de ulike cellene etter klassifikasjonen. *Expected frequencies* er de antall som skulle vært i cellen om det ikke var noen som helst relasjon mellom kategoriene (i min studie uenighet og usikkerhet). Ved statistisk beregning er et signifikansnivå på 5 % valgt som grense for statistisk signifikans (Altman, 2000; Polit & Beck, 2004).

3) Styrkeberegning

Antall pasienter i studien ble først beregnet ut fra litteratur gjennomgang (Essendrop et al., 2002). En styrkeberegning ble senere gjort når studien til Dankaert ble publisert. Ved å bruke 40 pasienter og beregne en sannsynlig kappa på 0.58 som er valgt fra en tidligere interrater reliabilitets studie på samme klassifiserings system (Dankaerts et al., 2006) får vi ved et 95 % konfidensintervall en range på 0.45 - 0.70. Dette regnes ut ved å omregne prosenttallene av korrekt og ikke korrekt klassifisering

av de fem mønstrene i Dankærts artikkel og sette dem pluss antall ønskede observasjoner inn i formelen for å beregne $se(k) = 0.065$. Deretter beregnes 95% konfidensintervall med Kappa 0.58 med følgende formel: fra $k - 1.96 se(k) = 0.45$ til $k + 1.96 se(k) = 0.70$ (Altman, 2000).

3.7 Etikk

Metodestudien ble godkjent som en tilleggsundersøkelse til "Skiveprotesestudien" av Regional Komité for Medisinsk Forskningsetikk Øst-Norge (REK 1) den 10.05.04 og godkjent av Datatilsynet den 24.09.04 (vedlegg 1 og 2).

Selv om situasjonen som skal videofilmes representerer en vanlig klinisk situasjon er man imidlertid klar over at videoopptak kan være forbundet med ekstra belastning for pasienten. Pasientene skulle bli filmet avkledd i bare undertøy. Pasientene ble informert om dette før de sa seg villige til å delta i metodestudien. Mulighet til anonymiseringen med hodehette og briller ble gitt alle pasienter. Videofilmkassetene ble overført til CD og arkivert i et låsbart skap etter forskrifter fra Datatilsynet (vedlegg 2). Filmene kan bare vises i samband med presentasjon av studien på UUS. All annen bruk må hver enkel pasient gi skriftlig tillatelse til. Filmene skal slettes når studien er ferdig.

Ved intervensjonsstudier er det viktig som forsker å forholde seg nøytral. Ved filmingen visste vi hvilken intervensjon pasienten var randomisert til og diskuterte det derfor minst mulig med pasienten. Som fysioterapeut og forsker var det viktig å ikke være behandler i intervensjonsgruppen der de filmene pasientene fikk tverrfaglig ryggrehabilitering. Selve intervensjonen ville ikke ha noen betydning for subgrupperingen, men det ville innebære at vi kunne ha sett pasientene over tid i forskjellige aktiviteter og under samtaler, hvilket kunne ha påvirket våre resultater i metodestudien.

4. Resultater

4.1 Demografiske data

Av over 3000 pasienter som ble vurdert på ryggseksjonen (Ortopedisk senter) på UUS i løpet av inklusjonsperioden 2004-2007, ble 276 pasienter vurdert å være mulige kandidater til å inngå i protesestudien. Etter nærmere seleksjon ble 222 av disse ekskludert og 54 inkludert. Av de 54 inkluderte i protesestudien var 51 villige til også å delta i min metodestudie. Det ble ikke oppgitt noen begrunnelse for de som ikke ønsket å bli inkludert i min studie.

4.1.1 Deltagere

Fra høsten 2004 til sommeren 2007 ble 41 pasienter (22 menn og 19 kvinner) i alderen 25-55 år rekruttert fra protesestudien til å delta i denne metodestudien. Pasientene scoret i gjennomsnitt 43 % på Oswestry hvor den laveste på 28 % er en pasient som ble inkludert ved en feilregning av Oswestry ved inklusjon. Pasientene hadde gjennomsnittlig rygg/hofte smerter på 72 mm på VAS skala og gjennomsnittlig lår/ legg smerte på 47 mm på VAS skala (tabell 4). Pasientene hadde ved inklusjon i gjennomsnitt hatt smerter kontinuerlig i 333 ± 114 uker (range 219-447). Det laveste var kontinuerlige smerter de siste 8 uker og det høyeste de siste 7 årene. 9 pasienter var tidligere prolapsoperert (mikrokirurgi).

Pasientene i metodestudien scoret i gjennomsnitt 12 ± 1 (range 0 - 22) på FABQ A og gjennomsnitt 25 ± 2 (range 2 - 42) på FABQ W.

Gjennomsnittscore på HSCL- 25 var 1.86 ± 0.5 (range 1 - 3.0).

Tabell 4. Demografiske data

N = 41 (22 menn, 19 kvinner)			
	Mean	SD	Range
Alder	37 – 57	± 8	25 – 55
ODI	43 %	± 9	28 – 62
VAS rygg/hofte	72 mm	± 14	18 – 89
VAS lår/legg	47 mm	± 10	37 - 57
FABQ A – Fys akt.	12	± 1	0 – 22
FABQ W – Arbeid	25	± 2	2 – 42
HSCL 25	1.86	± 0.5	1 – 3.08

Av de 41 pasientene i metodestudien var 16 pasienter sykemeldt ved inklusjon, 8 var i arbeid, 2 var deltids sykemeldte, 1 var aktivt sykemeldt, 9 var på attføring/rehabiliterings penger, 5 hadde ikke jobb og av disse var 2 studenter, 1 var arbeidsledig, 1 var hjemmeværende og 1 var uføretrygdet. Alle de inkluderte pasientene hadde en eller annen gang vært sykemeldt grunnet ryggsmarter.

4.1.2 Fordeling av filming og anamneseopptak mellom terapeutene

Fordelingen av hvem som filmet og hvem som tok opp anamnese ble som tidligere nevnt vilkårlig gjort i forhold til hvem som hadde kapasitet. Hos 10 av pasientene filmet terapeut 1 og terapeut 2 tok opp anamnese. Hos 19 pasienter filmet terapeut 2 og terapeut 1 tok opp anamnese. Som tidligere presisert hadde terapeutene ikke lov å diskutere pasienten etter film og anamnese opptakene. Hos 11 av pasientene var det en av terapeutene som både filmet og tok opp anamnese. Her fordelte det seg jevnt mellom terapeutene med 5 respektive 6. Dette skjedde spesielt i ferier eller ved sykdom. Det var ikke mulig å randomisere hvem som skulle filme og intervju pasientene da dette ble gjort i arbeidstiden og vi la det inn etter kapasitet og håpet det skulle fordele seg noenlunde jevnt over de tre årene.

4.2 Reliabilitet

4.2.1 Subgruppering av control impairment og movement impairment

Tabell 5 viser Kappa koeffisient og prosentvis enighet ved scoring av movement impairment og control impairment mellom to terapeuter. Prosentvis enighet ligger på 100 % og kappa er på 1.00 som etter Altmans skala klassifiseres som perfekt enighet. Majoriteten av pasientene, 37 av 41, ble subgruppert som control impairment.

Tabell 5. Kappaverdi og prosentvis enighet ved subgruppering av movement impairment og control impairment. N 41

		Terapeut 2		
		Movement imp	Control imp.	Total
Terapeut 1	Movement imp.	4 100 % enighet	0	4
	Control imp.	0	37 100 % enighet	37
	Total	4	37	41

4.2.2 Subgruppering av pasienter med control impairment

Tabell 6 viser kappa og prosentvis enighet mellom to terapeuter, ved subgruppering av 37 pasienter med control impairment. Prosentvis enighet mellom de to terapeutene ligger på 65 % og kappa verdien på 0.52 (0.32-0.71). Etter Altmans skala regnes en kappa på 0.52 som moderat enighet. Den laveste verdien i konfidens intervallet er på 0.32 og regnes som svak enighet og den høyeste er på 0.71 og regnes som god enighet. Største uenigheten ses i øverste raden på tabellen der terapeut 1 har satt fleksjonsmønster og terapeut 2 har satt aktivt ekstensjonsmønster på 6 av de samme pasientene. En konsekvent feil i en og samme celle drar ned kappa verdien. Totalt 13 av 37 pasienter var terapeutene uenige om.

Tabell 6. Kappaverdi og prosentvis enighet ved subgruppering av control impairment til 5 ulike mønster. N 37

N = 37 K= 0. 52(0.32-0.71) 65 % enighet (24/37)		Terapeut 2					
		Fleksj.	Fleksj. m shift	Aktiv Ekstensj.	Passiv Ekstensj.	Multidir.	Total
Terapeut 1	Fleksj.	5	0	6	1	1	13
	Fleksj. m shift	0	7	1	0	0	8
	Aktiv Ekstensj.	0	0	11	1	0	12
	Passiv Ekstensj.	0	1	0	0	0	1
	Multidir.	0	1	1	0	1	3
	Total	5	9	19	2	2	37

Manuell Kappa utregning: Målt eksakt enighet $24/37 = 0.65$. Forventet enighet i cellene i diagonalen: multipliserer antall (eks.fleksj. 5×13) og dividerer med total (37). Summere alle celler i diagonalen og del på total ($10.06/37$). Jeg får da tallet 0.27. Tallet settes inn i formel $0.65 - 0.27 / 1 - 0.27 = 0.52 k$. Spredning: $Se(k) = 0,098$. Ved 0.95% konfidens intervall (1.96 fra mean) $0.098 \times 1.96 = 0.192$ $K = 0.52 \pm 0.192(0.32 - 0.71)$.

Fordeling av de forskjellige mønstrene blant de pasienter vi var enige om (N 24): Fleksjon ($5/24$) 21 %, fleksjon med shift ($7/24$) 29 %, aktiv ekstensjon ($11/24$) 45 %, passiv ekstensjon ($0/37$) 0 %, multidireksjonalt ($1/24$) 4 %.

Prevalens av de ulike mønstrene i den totale gruppen: det blir vanskelig å få frem riktig prevalens da vi ikke har en terapeut som anses å være en slags gullstandard (Dankerts et al. 2003).

4.2.3 Sammenslåing av mønster etter smerteretning ved subgruppering av control impairment

Tabell 7 viser enigheten ved sammenslåing av mønster i 3 større grupper (fleksjon, ekstensjon og multidireksjonal). Disse tre hovedgruppene representerer smerteretningen, som er den viktigste faktoren ved klassifisering. Den prosentvise enigheten øker fra 65 % til 68 % med 25 enige score av totalt 37. Kappa derimot sank til 0.44. Nå er det 8 pasienter som terapeut 1 har scoret fleksjonsmønster der terapeut 2 scoret ekstensjonsmønster. Til tross for større prosentvis enighet er også uenigheten i en og samme celle større hvilket gir minsket kappa. Kappa på 0.44 regnes etter Altmans skala som moderat enighet. Spredning $k = 0.44$ (0.20 - 0.68) med $sd(k) = 0.121$

Tabell 7. Kappaverdi og prosentvis enighet ved sammenslåing av de 5 mønstergruppene til 3 større grupper. N=37

N = 37 K = 0.44 (0.20-0.68) 68 % enighet (25/37)		Terapeut 2			
		Fleksj.	Ekstensj.	Multidir.	Total
Terapeut 1	Fleksj.	12	8	1	21
	Ekstensj.	1	12	0	13
	Multidir	1	1	1	3
	Total	14	21	2	37

4.2.4 Subgruppering av control impairment ved kun anamnese

Tabell 8 viser kappaverdi og prosentvis enighet mellom 2 terapeuter ved subgruppering av pasienter med control impairment ut fra kun anamnese. Subgruppering også nå til 3 hovedgrupper (fleksjon, ekstensjon, multidireksjonell), da man ut fra anamnesen kun kan analysere retningen pasienten opplever som mest smertefull. Prosentvis enighet ligger nå på 81 % og kappaverdien på 0.66 (0.44-0.87). Kappa på 0.66 regnes som god enighet. Den laveste verdien på 0.44 regnes som

moderat enighet og det høyeste på 0.87 regnes som svært god enighet. Fortsatt er det 6 pasienter som blir scoret som fleksjonsmønster av terapeut 1 og ekstensjonsmønster av terapeut 2. Totalt 7 av 37 pasienter var vi uenige om. Alle disse 7 finner man blant de 13 uenige i tabell 3.

Tabell 8. Kappaverdi og prosentvis enighet ved subgruppering av control impairment ved kun anamnese. N=37

N = 37 K= 0. 66 (0.44-0.87) 81 % enighet (30/37)		Terapeut 2 kun anamnese			
		Fleksj.	Ekstensj.	Multidir.	Total
Terapeut 1 kun anamnes	Fleksj.	13	6	1	20
	Ekstensj.	0	16	0	16
	Multidir	0	0	1	1
	Total	13	22	2	37

4.3 Sammenheng mellom uenig og usikker ved scoring

Tabell 9 viser en 2 x 2 tabell med fordelingen mellom usikker og uenig ved scoring. Ingen celler (0 %) har forventet verdi mindre enn 5. Minste forventet verdi er 6,32 (13x18/37). I tabell 5 er alle usikre score tatt med. Hos flere pasienter er det bare en av terapeutene som er usikker. Ved gjennomgang av scoreskjemaene var det bare ved 8 score at begge terapeutene var usikre på samme pasient.

Tabell 9. Sammenheng enig - uenig og sikker - usikker N=37

		Sikker – usikker på mønster		
		sikker	usikker	Total
N= 37				
P= 0.052				
Mønster	Enig- % av total	15 40.5%	9 24.4%	24 64.9%
	Uenig % av total	3 8.1%	10 27.0%	13 35.1%
	Total % av total	18 48.6%	19 51.4%	37 100%

Chi-square test for independence viser en p-verdi på 0.052. Ingen forventet verdi er mindre enn 5 men da det er få observasjoner ble også en Fisher's eksakt test prøvet. Fisher's eksakt test viser en signifikant sammenheng med p verdi lik 0.038.

Sammenhengen er signifikant ved $p \leq 0.05$

4.4 Oppsummering av resultatene

- Svært god enighet ved subgruppering av control imp. og movement imp.
 $k = 1.0$
- Moderat enighet for subgruppering av control impairment til 5 mønster.
 $k = 0.52 (0.32 - 0.71)$
- Lavere kappa ved sammenslåing av mønster i 3 grupper (fleksjon, ekstensjon, og multidireksjonal) $k=0.44 (0.20 - 0.68)$
- God enighet ved subgruppering av bare anamnesen i 3 grupper (fleksjon, ekstensjon og multidireksjonal) $k=0.66 (0.44 - 0.87)$
- Sammenheng mellom uenighet i mønstersetting og usikkerhet ved mønstersetting. $p=0.052$ (Continuity Correction), $p=0.038$ (Fisher`s Exact Test)

5. Diskusjon

I den første delen av dette kapittelet diskuteres utvalg, metode og statistikk. Det vil bli lagt vekt på å diskutere valg av design og de problemstillinger som kom opp under selve gjennomføringen av metodestudien. Ved reliabilitetstesting kan resultatene bli påvirket av mange ulike faktorer. Hovedfaktorene som kan påvirke resultatene i denne studien er terapeutene som tester, selve metoden for subgrupperingen og de inkluderte ryggpasientene (utvalget). Jeg ønsker her å belyse disse faktorene.

5.1 Utvalg

5.1.1 Inklusjon og eksklusjon

Pasientene i Skiveprotesestudien er en selektert gruppe hvor utgangspunktet for inklusjon er at pasienten oppfyller et sett med definerte radiologisk degenerative forandringer i et eller to segment. Bakgrunnen for at man valgte disse kriteriene er tidligere kirurgisk erfaring hvor man har observert at det er pasienter med degenerasjon utelukkende i mellomvirvelskiven som antas å ha størst effekt av en skiveprotese (Siepe et al., 2006; Tropiano et al., 2005). Modic forandringer langs endeplatene og HIZ i annulus fibrosus er begge assosiert med høy forekomst av rygg smerter (Albert et al., 2008; Jensen et al., 2007; Kjaer et al., 2006a; Kjaer et al., 2006b; Schellhas et al., 1996; Peng et al., 2006). Teorien er da at man ved en operasjon fjerner det smertefulle stimulus. Synspunktet, om at det er visse degenerative forandringer som skulle være årsaken til rygg smerter, har først og fremst vokst frem gjennom de observasjoner man systematisk har gjort på MR (Albert et al., 2008; Kjaer et al., 2006a; Kjaer et al., 2006b). Mange er derfor nå av den oppfatning at dette er en gruppe spesifikke ryggpasienter, altså der hvor det er et spesifikt patoanatomisk substrat som er årsak til plagene (Albert et al., 2008; Kjaer et al., 2006a; Kjaer et al., 2006b). Dette synet har imidlertid blitt kontradiktert av andre. Mange billedfunn som anses å kunne forårsake smerteplager er også sett hos friske smertefrie personer (Jarvik & Deyo, 2002; Carragee et al., 2005). Ut fra dagens

kunnskap må man derfor kunne konkludere med at det ikke er noen sikker sammenheng mellom kliniske funn og radiologiske forandringer og at pasienter inkludert i denne studien er å anse som pasienter med uspesifikke korsryggsmerter.

For å bli inkludert i Skiveprotesestudien krevdes en Oswestry funksjons score på \geq 30 % (poeng). Grensen på 30 % ble valgt da dette er vanlig brukt ved andre studier hvor operasjon inngår som en av intervensjonene (Brox et al., 2006; Fritzell et al., 2001; Schellhas et al., 1996). Man ønsket derfor en såpass redusert funksjon at det var forsvarlig å vurdere operasjon. At man i denne studien har satt et krav om grad av funksjons nedsettelse gjør at utvalget også på dette punktet er en selektert gruppe i forhold til tidligere publiserte studier hvor dette systemet har vært brukt (Dankaerts et al., 2006; Vibe et al., 2008). I studien til Vibe Fersum var den gjennomsnittlige Oswestry skåren på 21.2 % mens i min studie var den på 41 % (Vibe et al., 2008). Pasientene i min studie hadde også en mye lenger symptomvarighet og kravet til inklusjon var ryggplager i minimum et år. I Vibe Fersums studie ble pasienter med sykemeldning grunnet ryggplager mer enn 4 måneder ekskludert. Oswestry inngår ikke som en del av O'Sullivans klassifikasjonssystem. Man vurderer funksjonen men subgrupperer ikke etter grad av generell funksjonsnedsettelse. I min metodestudie brukes de inkluderte pasientenes Oswestry scor for å beskrive utvalget.

Dankaerts ekskluderer pasienter som er tidligere ryggopererte i sin interrater reliabilitets studie (Dankaerts et al., 2006). I denne studien har vi tatt med pasienter som er tidligere ryggoperert forutsatt at de har gjort et mikrokirurgisk inngrep. Mikrokirurgiske inngrep affiserer bevegeligheten lite og pasientene har etter operasjonen ingen restriksjoner og kan komme direkte tilbake i vanlig aktivitet. Vi har i klinikken subgruppert og behandlet mikroopererte pasienter etter O'Sullivans prinsipper og hatt god respons ved motorisk endring og læring som intervensjon. Fersum ekskluderte bare ryggoperasjoner som påvirket lumbalcolumna eksempelvis fusjonsoperasjoner (Vibe et al., 2008). Disse var også ekskluderte i vår studie.

Ved inklusjon i Skiveprotesestudien var pasientene allerede screenet for "røde flagg" og "gule flagg" før de ble inkludert i studien da dette var en del av eksklusjonskriteriene. Vi antok derfor at de pasientene som ble inkludert i

metodestudien hadde psykososiale faktorene av underordnet betydning for smerten. Den del av klassifikasjonssystemet som blir testet i denne metodestudien er altså etter O'Sullivan's klassifikasjonssystem for pasienter som antas å ha uspesifikke kroniske korsryggmerter som primært vedlikeholdes av nedsatt bevegelse eller feiltilpasset motorisk kontroll.

5.1.2 Deltakelse

Intervensjonene; skiveproteseoperasjon eller tverrfaglig rehabilitering kan også tenkes å ha betydning for hvem som ble inkludert til deltakelse i studien. Deltakelse i studien innebærer at pasientene må være villige til å motta en av de to behandlingene avhengig av loddtrekningen. Hos noen av de pasienter som takket nei var begrunnelsen at de hadde trent mye og ønsket bare operasjon, mens andre ikke ønsket operasjon.

5.1.3 Ekstern validitet

Ved hjelp av inklusjons- og eksklusjonskriterier har man selektert pasienter med uspesifikke korsryggmerter som har felles degenerative karakteristika ut fra MR - funn. Det er viktig ved besvarelse av problemstillingen i denne oppgaven å presisere at klassifikasjonssystemet blir testet på en selektert gruppe av pasienter med uspesifikke korsryggmerter med avgrenset radiologisk degenerasjon.

Fordelen med å delta i en større studie er den grundige screeningen som pasientene gjennomgår. Det gir også muligheter til senere å bruke resultatene fra klassifiseringen til for eksempel å undersøke dets prediktive verdi. Dersom klassifikasjonen viste seg å ha en god reliabilitet på dette selekterte pasientmaterialet kunne man se om det er enkelte undergrupper som responderer bedre på den operative eller den konservative behandlingen.

I min jobb som fysioterapeut i tredjelinjetjenesten vil man alltid møte et selektert utvalg av ryggpasienter. Jeg mener at det er viktig også for denne spesialiserte delen av fysioterapifaget å teste ut nye klassifikasjonssystem og behandlingsparadigmer.

Da klassifikasjonssystemet har blitt forandret og utvidet de siste årene er det også viktig å presisere hvilke deler av systemet som er testet ut i denne undersøkelsen (fargelagt del i figur A i vedlegg 5).

5.2 Metode

5.2.1 Valg av O'Sullivan's klassifikasjonssystem i metodestudien

Da jeg kom i kontakt med O'Sullivan's klassifikasjonssystem var det først og fremst behandlingsdelen jeg fattet interesse for. For å kunne ta i bruk O'Sullivan's behandlingsprinsipper må man først klare å klassifisere riktig. Herunder diskuterer jeg noen av faktorene som bidrar til at jeg har valgt O'Sullivan's klassifikasjonssystem i metodestudien.

I den senere tids ryggforskning blir korsryggsmerter sett på som et flerdimensjonalt problem med mange faktorer som påvirker forløpet (Borkan et al., 2002; McCarthy et al., 2004). Som tidligere nevnt består disse dimensjonene av patoanatomiske, neurofysiologiske, fysisk mekaniske forhold og psykososiale faktorer og man snakker ofte om en biopsykososial smertemodell (Waddell, 2004). Gordon Waddell hevder at økningen av kroniske ryggglidelser virker å bero mer på samfunnet og medisinen's forståelse enn endring i den biologiske sykdommen (Waddell, 1993). Man har i review artikler som omhandler klassifikasjonssystemer for LBP diskutert om de bør være biopsykososialt rettet for å være valide (Billis et al., 2007; Dankaerts et al., 2006; McCarthy et al., 2004).

Mc Kenzie, Delitto og Sahrman har en biomedisinsk forståelse som er basert på at mekanisk vevsskade gir pasienten økte smerter. Mc Kenzie og Delitto retter hovedsakelig sitt fokus mot akutte korsryggsmerter. Sahrman's klassifikasjonssystem er begrenset for bruk av kronisk uspesifikke korsryggsmerter og man legger ingen vekt på den psykososiale siden ved ryggsmertene. QTFC ligger til grunn for O'Sullivan's klassifikasjonssystem, men QTFC subgrupperer ikke de uspesifikke kroniske korsryggsmertene spesifikt etter underliggende mekanismer for smerten og

inkluderer bare de psykososiale faktorer dersom man ikke klarte å finne en biomedisinsk forklaring på smertene. Petersen argumenterer for sitt biomedisinske klassifikasjonssystem med at LBP pasienter ofte ønsker å finne årsaken til sine smerter og at redsel kommer av uvissheten rundt smertene (Petersen et al., 2003). Wadell kritiserer på den andre siden strukturorienterte diagnostiske system for å skape bevegelsesangst hos pasienter og lage ineffektive behandlinger (Waddell, 2004). Mange strukturer i og rundt ryggsøylen kan gi smerter, men patoanatomiske funn korrelerer dårlig med smerte og grad av nedsatt funksjon (Nachemson et al. 1999). O'Sullivan mener at det er umulig å klassifisere ut fra disse og finne riktige behandlingstiltak (O'Sullivan, 2005). O'Sullivan har derfor laget et biopsykososialt klassifikasjonssystem som klassifiserer ut fra bakenforliggende mekanismer for smerten.

O'Sullivan har ved utviklingen av sitt klassifikasjonssystemet og sine behandlingsteorier utgått fra senere tids forskning omkring dyp stabiliserende muskulatur og disses innflytelse på korsryggen (Bergmark, 1989; Panjabi, 1992b; Panjabi, 1992a). Teoriene har blitt videreutviklet og inkludert i klinikken. (Maluf et al., 2000; O'Sullivan, 2000; O'Sullivan et al., 1997; Van Dillen et al., 2001; O'Sullivan et al., 2002; Hides et al., 2001; Hodges & Richardson, 1999). Studier har vist at motorkontroll og bevegelsesmønster endres sekundært til smerten (Moseley & Hodges, 2005; Hodges & Richardson, 1996; Hodges & Moseley, 2003; Hodges et al., 2003). Flere publiserte studier viser at spesifikke stabiliserende øvelser og trening av motorkontroll har effekt på korsrygg smerter (Hides et al., 2001; Moseley et al., 2002; O'Sullivan et al., 1997; Van Dillen et al., 2003).

O'Sullivan har også tatt hensyn til at psykososiale faktorer kan påvirke smerte og inkluderer dette i sitt klassifikasjonssystem (figur A i vedlegg 5). Psykososiale faktorer som primær årsak til utvikling og opprettholdelse av kronisk smerter såkalte "gule flagg", er en av hovedkategoriene i O'Sullivans klassifiseringssystem for kroniske korsrygg smerter (vedlegg 5). O'Sullivan er i dag et av få behandlingsrelaterte klassifikasjonssystemet som tar med de psykososiale faktorene i subgrupperingen av LBP pasienter (Billis et al., 2007).

I behandlingsopplegget for denne pasientgruppen har O'Sullivan hovedvekten på motorisk læring, men dette skjer innenfor en kognitiv ramme der man er bevisst på at psykososiale faktorer kan influere på og vedlikeholde smerte (figur A i vedlegg 5). I studien til Fersum subgrupperte man også MI og MCI og bekkensmerte pasientene etter grad av psykososiale faktorer som påvirket pasientens smerteadferd (Fersum 2008). I min metodestudie og i Dankaerts studie er ikke disse faktorer blitt vurdert og klassifisert (Dankaerts et al., 2006). Denne delen er ny i klassifikasjonssystemet og krever en annen form for anamnese og klinisk undersøkelse.

Klinikken jeg har jobbet ved i mange år er spesialisert på kroniske muskelskjelettlidelser og har lang tradisjon for en biopsykososial sykdomsforståelse. En del av denne behandlingen består i bevisstgjøring, normalisering av bevegelsesmønster og trening av motorkontroll med pasientens funksjonsbehov i fokus. Med dette som bakgrunn var valget av O'Sullivans klassifikasjonssystem og behandlingsprinsipper naturlig. Klassifikasjonssystemet til O'Sullivan er også enkelt å anvende i klinikken og ikke altfor tidkrevende.

5.2.2 Valg av design

Som tidligere nevnt må et system være reliabelt for å kunne være valid, men en god reliabilitet medfører ikke nødvendigvis en god validitet (Peat et al., 2002). Det er derfor viktig at man først reliabilitetstester klassifikasjonssystemet. I vår daglige praksis var vi først og fremst opptatt av å samkjøre oss fysioterapeuter imellom og valgte derfor å teste interraterreliabiliteten. Det hadde vært naturlig også å teste intrareterreliabiliteten, men det la seg ikke gjøre tidmessig da vi trengte å vente noen måneder før vi så på filmene igjen for ikke å huske scoringene fra første gang. Studier har vist at intrareterreliabiliteten gir generelt litt høyere Kappa verdi enn interraterreliabiliteten (Vroomen et al., 1999).

Et problem var at vi ikke hadde en gullstandard for klassifikasjonssystemet. Når vi klassifiserte så vi bare på om vi subgrupperte likt eller ulikt, men vi kan ikke si noe om vi begge hadde riktig eller feil scor eller hvem av oss som hadde riktig eller feil. Vi risikerer å være enige om en feilaktig subgruppering. Videofilmene og

anamnesene gir oss imidlertid en mulighet til å få andre erfarne terapeuter til å klassifisere pasientene. Vi kan dermed kontrollere våre resultater. I Dankaerts studie brukte man to meget erfarne klinikere (hvorav den ene var opphavsmannen) som gullstandard (Dankaerts et al., 2006). Man fant i denne studien høy interraterreliabilitet mellom de erfarne terapeutene ved klassifisering, men vet imidlertid fortsatt ikke om deres vurdering er valid. Validitet kan etableres gjennom klinisk randomiserte kontrollerte studier. Dette kan eksempelvis gjøres ved å først subgruppere et gitt pasientutvalg og deretter randomisere disse til enten ordinær behandling eller den for subgruppene beste behandling. Deretter sammenligner man pasientene som fikk behandling som matchet sin subgruppe mot de som fikk en ikke matchet behandling (Brennan et al., 2006).

5.2.3 Observasjon som metode ved registrering av bevegelsesmønster

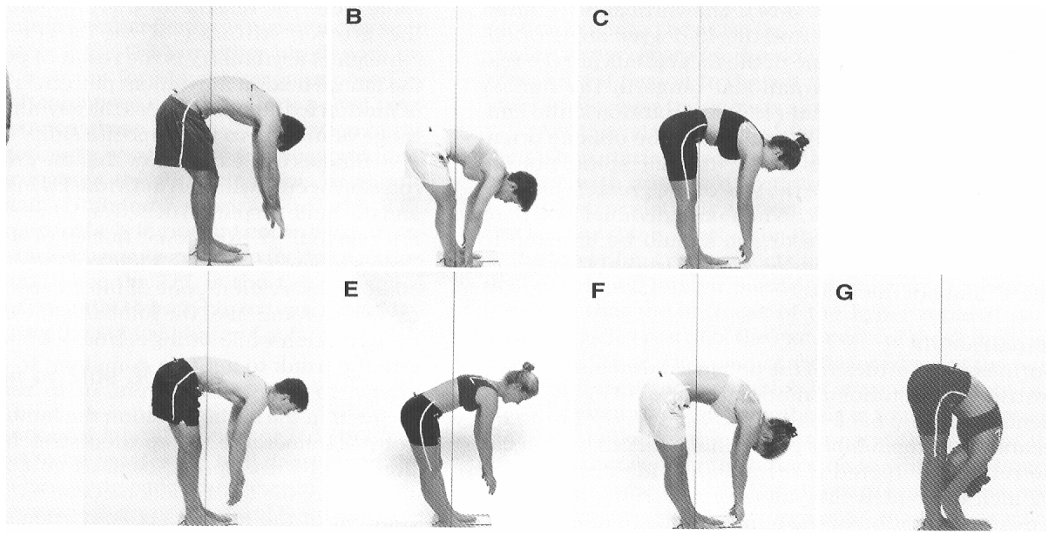
A conclusion based upon a badly conceived experiment is usually further from the truth than one based on clinical observation. Sir Robert Platt 1963

Observasjon og bevegelsesanalyse er en viktig del av den fysioterapeutiske funksjons- undersøkelse, behandling og evaluering. Vi ser i klinikken og i litteraturen at måten pasienten beveger seg på kan forandres i forbindelse med smerter i og rundt ledd (Hides et al., 1996; Sahrman, 2002; Sterling et al., 2001), men vi har ingen standardiserte måter å formidle det på. I min metodestudie bruker jeg observasjon som en viktig informasjonskilde ved subgruppering. Jeg vil her diskutere ulike faktorer som kan påvirke reliabiliteten ved bruk av observasjon.

All menneskelig sanseoppfattelse er selektiv. Ulike mennesker legger merke til ulike forhold ved en observasjon. Hva som observeres blir avhengig av hvem som observerer (Carlsson, 1990). Hver observatør har sin egen referanseramme som bestemmer hva han eller hun observerer i en viss situasjon. For å komme til rette med dette problemet ved bruk av observasjon i vitenskaplige sammenheng bestemmer man på forhånd hva som skal observeres. På denne måten blir det også mulig å trene flere observatører til å observere på en lik måte (Carlsson, 1990).

Tolkning av et statisk fenomen kan være vanskelig, man har for eksempel funnet variasjon ved beskrivelse av lumbale skiveprolaps ved MR undersøkelse (Brant-Zawadzki et al., 1995; Milette et al., 1999). Enda vanskeligere blir det når man skal tolke dynamiske bevegelser enten over et ledd for eksempel kneledd eller en rekke sammensatte bevegelser/funksjoner over mange ledd og i flere plan, for eksempel ganganalyse, humeroscapulære rytme og fleksjon av ryggen. Det er da viktig at man har en relativ lik forståelse for hva som skal undersøkes og hvordan man vurderer det man ser. Dette for å kunne formidle funn mellom behandlere og for å kunne diagnostisere likt (Kibler et al., 2002). Man må være enig om de samme anatomiske referansepunkter og fysiologiske prinsipper og ha det samme teorigrunlaget for analysen.

Bevegelser består ofte av mange sammensatte komponenter og det er derfor vanskelig å beskrive dem på en enkel og presis måte (Hickey et al., 2007). Det er store variasjoner i menneskers anatomi, holdning, bygning, bevegelsesmåter, rytme, timing, alder, kjønn, med mer. Dette innebærer også at jo flere komponenter som er involvert desto vanskeligere blir det å lage standardiserte kategorier (Hickey et al., 2007). Det er gjort forsøk med å utforme kategorier med beskrivelser på humeroscapulær rytme, som er en stor utfordring da man har flere plan, flere ledd og mange muskelgrupper involverte. (Hickey et al., 2007; Kibler et al., 2002; Sahrman, 2002). Det samme er gjort for kroniske ryggpasienter (O'Sullivan, 2000; Sahrman, 2002) Figur 7 viser eksempel på ulike måter å bøye ryggen.



Figur 7. Eksempel på ulike måter å bøye ryggen på, bilde fra Sahrman's bok *Diagnosis and treatment of movement impairment syndromes* (Sahrmann, 2002)

I min metodestudie var observasjon og analyse av bevegelse av stor betydning ved subgrupperingen. Vi hadde enkelte ganger store problem å tolke pasientens bevegelsesmønstre. Pasientenes variasjon i anatomi og størrelse i min studie gjorde det av og til vanskelig å se strukturene nøye nok. Vi prøvde å bruke teip på pilotpasientene for å markere SIAS (spina iliaca anterior superior) og SIPS (spina iliaca posterior inferior) men hos de adipøse pasientene ble dette mer mot sin hensikt da teipen beveger seg med huden og ikke med beinstrukturene.

De fleste studier viser generelt moderat interrater reliabilitet ved observasjon av bevegelsesmønstre (Dankaerts et al., 2006; Kibler et al., 2002). Det har blitt diskutert om den moderate interrater reliabiliteten beror på at observatøren har vanskelig for å overføre det man ser og tolker i skriftlig form eller som i min studie over i en definert subgruppe (Kibler et al., 2002; Hickey et al., 2007).

Nivå på overensstemmelse er avhengig av hvilken variabel som skal observeres. Diskrete eller diskontinuerlige variabler er lettere (hø eller ve hånd). Kontinuerlige variabler og beskrivelse av sammensatte bevegelser som bevegelsesmønstre er vanskeligere (humeroscapulær rytme) (Kibler et al., 2002). Å observere bevegelse i lumbalcolumna er vanskelig da det består av flere ledd som beveges samtidig og henger sammen med bevegelse i hofteleddene og thoracalcolumna.

5.2.4 Valg av video ved analyse av bevegelsesmønstre

Video opptak er vanlig å bruke i klinikken for å analysere bevegelseskvalitet (Delitto et al., 1992; Eastlack et al., 1991; Fritz et al., 2000; Kibler et al., 2002; Krebs et al., 1985). Videoopptak har blitt anbefalt i tidligere studier som et alternativ ved en test - retest design for å utforske interrater reliabilitet ved klassifisering av pasienter med LBP (Dankaerts et al., 2006; Delitto et al., 1992). Videofilming ble valgt i min studie for at det er minst belastende for pasienten, tar kortest tid og terapeutene vurderer eksakt samme pasientsituasjon. Det ville sikre at pasienten utfører de samme bevegelsene for begge observatørene (Delitto et al., 1992; Domholdt, 2000; Fritz et al., 2000; Van Dillen et al., 1998).

Filmen kan ses i slow motion og kan ses om igjen flere ganger. Vi har også i etterkant av studien mulighet for å se på video opptakene sammen og diskutere valg av subgruppe. I tillegg gir dette oss mulighet til å kunne trekke inn flere fysioterapeuter for klassifisering av pasientenes bevegelsesmønstre og dermed være i stand til å gjøre metodeundersøkelsen best mulig.

På den andre siden er videofilm av standardiserte bevegelser ulik den vanlige kliniske undersøkelsessituasjonen. Det å ikke kunne være "hands on" på pasienten, belaste, avlaste og stille supplerende spørsmål underveis når man undersøker er begrensende i forhold til vanlig praksis. Til tross for at kvaliteten på filmingen var god i min studie var det imidlertid flere pasienter hvor man fikk ulikt inntrykk av pasienten ut fra anamnesen i forhold til filmen. Det hadde da vært ønskelig å kunne teste pasienten ytterligere eller se på andre typer bevegelser. Det kunne også vært aktuelt med oppfølgningsspørsmål for å klargjøre klassifiseringen.

Et av argumentene for å bruke video filming ved reliabilitetstest av bevegelsesmønstre baseres på tidligere reliabilitetsstudier der man har sett at reliabilitet blir dårligere ved økning av symptom i forbindelse med flere eksaminasjoner etter hverandre eller eksaminasjoner på ulike dager (Delitto et al., 1992; Fritz et al., 2000; Van Dillen et al., 1998). Jeg mener at dette ikke bør ha noen større betydning ved bruk av klassifikasjonssystemet til O'Sullivan da pasienten ikke klassifiseres etter grad av smerter eller utfall, men etter smerteprovokasjon og

pasientens bevegelsesstrategier. I den kliniske hverdagen vil pasientens symptomer stadig variere og klassifikasjonssystemet må da tåle disse variasjoner. Det er i midletid viktig at pasienten klarer å utføre eksaminasjonen relativt likt for observatørene. I studien til Vibe Fersum kontrollerte man for dette ved å sette en grense hvor stor forandring på VAS man kunne tillate mellom eksaminasjonene (Vibe et al., 2008).

Nye teknisk avanserte målemetoder er under utvikling for å prøve og gi et mer nøyaktig svar på hva som virkelig skjer ved bevegelse under hud og ytre muskulatur. Man prøver å matche den kliniske observasjon mot det biomekaniske systemet ved å ta i bruk instrumenter som kan måle bevegelige komponenter som røntgen gjennomlysning (Deutsch et al., 1996), EMG (Hodges & Richardson, 1996; Kelly et al., 2005), ultralyd (Critchley & Coutts, 2002), 3 dimensjonal videoanalyse med markører (Szeto et al., 2005; Taylor et al., 2001) og kombinasjon 3D og EMG (Bernhardt et al., 1998; Lin et al., 2005; Dankaerts et al., 2007). Ulempen med røntgen er at man bare ser to plan av 3 plans bevegelse. EMG er følsom for støy og kan bli upresis og ultralyd må appliseres likt for ikke å gi feilaktige vinkler og er derfor følsom for bevegelse. Avanserte tekniske målmetoden som blant annet 3D video med markører og kombinasjon 3D video og EMG er under stadig utvikling og noen av dem er i dag ikke brukbare til daglig i klinikken, men er av betydning i forskningssammenheng. De er ofte teknisk avanserte og krever mye tid. Det er imidlertid viktig å forske videre på dette, da det i fremtiden forhåpentligvis kan brukes som gullstandard.

5.2.5 Anamnese som grunnlag for subgruppering

Anamnesen har en viktig rolle i O'Sullivan's klassifikasjonssystem. Jeg valgte i min metodestudie et strukturert standardisert intervju. Å standardisere metoden er anbefalt ved reliabilitets testing for å skille metode feil fra subjektets variasjon (Domholdt, 2000). Et standardisert intervju kan lede til ulikheter i hvor utdypende og nøyaktig pasientene svarer. I metodestudien kunne vi derfor be pasienten om å utdype sitt svar, men det var viktig å ikke lede pasienten mot noen spesielle retninger (Svensson, 1984). Ved et standardisert intervju blir det opp til pasientens evne å uttrykke seg og

beskrive sine smertesymptomer i motsetning til et åpent intervju hvor man kan stille oppfølgende spørsmål for å få en klarere smerte og funksjonsbilde. Ved åpent intervju er risikoen derimot at man underveis ubevisst sikter seg inn mot en retning og dermed stiller utdypende spørsmål som underbygger det man misstenker. Dette er ikke mulig i like stor grad ved å ha standardiserte spørsmål og man unngår derved denne feilen. Betydningen av anamnese i kombinasjon med undersøkelse blir diskutert under diskusjon av resultatene.

5.2.6 Opptaksprosedyrene og scoringsprosedyren

Videofilming krever riktig utstyr, lokaliteter og lysforhold. Vi fikk hjelp av sykehusets fototekniske avdeling for å få tatt de beste opptakene der man kunne se flest mulig detaljer.

Når prosjektkoordinatoren hadde en ny pasient til filming ble pasienten satt opp på den av terapeutene som hadde mulighet. Ved film- og anamneseopptak av pilotpasientene og i begynnelsen av metodestudien var begge terapeutene til stede under filming og anamneseopptak, en av terapeutene filmet og en tok opp anamnesen. Prosedyren tok cirka 40 minutter når utstyret var satt opp. Ofte hadde vi 2-3 pasienter på samme dag. Vi diskuterte om terapeutene som var i samme rom ved filming og anamneseopptak kunne påvirke resultatene, men konkluderte med at selve klassifikasjonen kom til å finne sted når alle pasienter var inkluderte (inkludert beregnet til 3 år). Under filmingen var vi også nøye med aldri å diskutere den inkluderte pasienten og unnvike ledende spørsmål underveis som røpet hvilken subgruppe vi vurderte. I noen interrater-test studier undersøker terapeuter pasientene sammen, men scorer individuelt rett etter undersøkelsen (Heiss et al., 2004; Clare et al., 2005). Dette er en mer følsom situasjon da man ved klinisk undersøkelse ofte følger opp funn etter hvert som man undersøker og man kan ved å se på en kollega undersøke, forstå hvilken retning han/hun har tenkt. Dette unngikk vi med bruk av standardiserte bevegelser og anamnesespørsmål. Det var også liten mulighet for oss å huske alle pasientene etter tre år med filming og dessuten behandling av andre LBP pasienter i klinikken samtidig. Siste året var vi alene om filming og anamnese for å unngå påvirkning fra hverandre.

Etter tre år når alle pasienter var inkludert ble alle filmer og anamneser vurdert og subgruppert i tilfeldig rekkefølge. Terapeutene scoret blindet for hverandre. Ved subgruppering av pilotpasientene oppdaget vi at noen pasienters filmer og anamneser ga sprikende svar. Vi valgte derfor å føye til en rubrikk i skjemaet som ga oss mulighet til å si om vi var sikre eller usikre på vår subgruppering. Dette for å se om det var sammenheng mellom usikkerhet og ulik scoring.

5.2.7 Terapeutene

I løpet av de 3 årene som vi inkluderte og filmet pasientene har vi stadig sett på pasienter i avdelingen sammen og klassifisert sammen. Det å jobbe sammen, undersøke pasienter sammen og kontinuerlig diskutere funn og definisjoner anses som viktig for å få en høy interrater reliabilitet (Strender et al., 1997).

Studier hvor man ser på reliabilitet ved observasjon av bevegelsesmønstre, har vist at reliabilitet blir høyere jo mer trente observatørene er på oppgaven (Dankaerts et al., 2006; Hickey et al., 2007; Kibler et al., 2002). Ved bruk av O'Sullivan klassifikasjonssystem ser det ut at det kreves mye trening i systemet for å øke reliabiliteten (Dankaerts et al., 2006; Vibe et al., 2008). I min metodestudie ligger de to terapeutene på samme nivå med andre terapeuter som er moderat trente i klassifikasjonssystemet (Dankaerts et al., 2006). Derimot finner man ingen sammenheng mellom erfaring i faget og god reliabilitet, men derimot erfaring i systemet som skal testes (Eastlack et al., 1991). Det er også diskutert en artikkel til Heiss et al fra 2004 om eldre, erfarne terapeuter kan ha vanskeligere for eller er uvillige til å lære nye ting, mens de nyutdannede terapeutene er mer åpne for å lære noe nytt å følge beskrivelsen til punkt og prikke. De nyutdannede har da større mulighet å skåre mer likt enn de som gjør ting på sin egen måte (Heiss et al., 2004).

Jeg må derfor ta høyde for at de ovenfor diskuterte faktorene har spilt en rolle i min studie. Begge observatørene var meget erfarne i faget, men bare moderat trente i klassifikasjonssystemet. Jeg tror derimot ikke reliabiliteten i vår metodestudie hadde noe å gjøre med uvillighet til å lære nytt. Vi var derimot meget åpne for å lære klassifikasjonssystemet og følge beskrivelsen nøyaktig.

Jeg mener at det er viktig at et klassifikasjonssystem er laget på en såpass solid måte at det er nok å være moderat erfaren i systemet. Klassifikasjonssystemer styrkes ved å ha klare kriterier for kategorisering som er lette å finne ved anamnese og kliniske tester (Heiss et al., 2004; Petersen et al., 1999). Her kan man tenke seg at klassifikasjonssystemet til O'Sullivan kan videreutvikles og kategoriernes kriterier og definisjoner gjøres ennå mer presise. Dette for å gjøre klassifikasjonssystemet tilgjengelig for flest mulig. Ellers blir det bare brukbart for en liten gruppe og blir mindre alment gjeldende og generaliserbart.

Jeg har en dobbel rolle i denne studien både som forsker og kliniker og har sammen med min kollega både vært med og planlagt, gjennomført og klassifisert pasienter. Jeg har deretter vært den som har jobbet med den statistiske analysen og fått frem resultatene og senere tolket og diskutert resultatene. Dette er kanskje ikke det mest optimale med tanke på verdien av å være mest mulig nøytral gjennom hele forskningsprosessen. Da dette er en masterstudie og det er lite midler til bruk av annen personell må man derfor gjøre kompromisser. Jeg valgte først og fremst å bruke en uavhengig person til å punche alle dataene fra de to lukkede konvoluttene med scoringsresultatene. Jeg har gjennom hele prosessen med planlegging, gjennomføring og tolking av resultater hatt som mål å være mest mulig objektiv.

5.2.8 Klassifikasjonssystemets utvikling under studien

I en publisert artikkel fra 2007 beskriver Fritz og medarbeider utviklingen av klassifikasjonssystemet til Delitto. Hun poengterer blant annet at klassifikasjonssystemet er bygget opp etter ekspertisens empiri og den forskning som er tilgjengelig på det tidspunktet systemet presenteres (Fritz et al., 2007). Et klassifikasjonssystem skal også evalueres i randomiserte kontrollerte intervensjonsstudier (McDowell & Newell, 1996).

Etter hvert som forskningen vinner nye kunnskaper er det viktig at et klassifikasjonssystem kan modifiseres og utvikles både hva gjelder utvikling av kliniske prediksjonsregler, nye undersøkingskriterier for klassifikasjon og optimal behandling for hver subgruppe. Nye forsknings resultat skal regelmessig inkorporeres

i det eksisterende systemet (Fritz et al., 2007) Det er derfor en lang prosess å få et klassifikasjonssystem både reliabelt og valid.

Klassifikasjonssystemet til O'Sullivan er relativt nytt og har vært under stadig utvikling. Et resultat av dette er at har man endt opp med nye kategorier. Systemet er i dag presentert som et mer helhetlig system som er beregnet for alle kroniske korsryggmerter. Pasientene subgrupperes ned gjennom ulike nivåer (figur A i vedlegg 5) (Vibe et al., 2008). Innenfor gruppen med perifert medierte korsryggmerter skiller man på om pasienten har motor kontroll problem (MCI) eller bevegelses problem (MI) som årsak til smerten. For å subgruppere til disse på en optimal måte kreves en del klinisk tester i tillegg til bevegelsesanalyse og anamnese. Vi ble oppmerksomme på dette når Dankaerts studie ble publisert og vi allerede var i gang med studien (Dankaerts et al., 2006). I metodestudien inngikk ikke kliniske tester av pasientene og for å kunne fullføre subgrupperingen satte vi opp klare kriterier for MI og MCI som passet ved bruk av bare film og anamnese (vedlegg 6).

Utviklingen av systemet har vært en av de store utfordringene i denne studien da inklusjonen tok 3 år og systemet i mellomtiden hadde utviklet seg så pass mye at man egentlig burde lagt opp metoden for subgrupperingen noe annerledes. Jeg hadde valgt en klinisk undersøkelse av pasientene ved subgrupperingen istedenfor film og anamnese. Film og anamnese egner seg best ved subgruppering av de 5 smerte provoserende bevegelsesmønstrene, men ikke ved subgruppering til de ulike kategoriene tidligere i klassifikasjonssystemet. Skal man i fremtiden også subgruppere etter grad av psykososiale faktorer som påvirker pasientens smerteadferd må man nok teste systemet med mer dyptgående anamnese og klinisk undersøkelse.

5.3 Statistikk

5.3.1 Valg av statistiske metoder

For å måle grad av enighet ble kappa analyse valgt. Dette er den eneste metoden å bruke ved kategoriske nominale data (Altman, 2000). Tidligere ble prosentvis enighet

brukt alene, men den kontrollerer ikke for enighet ved sjanse. Da de fleste interrater reliabilitets studier på klassifikasjonssystem har valgt å oppgi enighet med både kappa og prosentvis enighet blir sammenligning lettere. (Clare et al., 2005; Dankaerts et al., 2006; Fritz & George, 2000; Heiss et al., 2004; Van Dillen et al., 1998; White & Thomas, 2002). Ved reliabilitetstester der man har kategoriske ordinale data, det vil si mange graderinger som kriterier for subgruppering, kan en vektet kappa brukes. Vektet kappa betyr at man vurderer noen grupper mer beslektet og nærmere enighet enn andre grupper. Ved kontinuerlige data brukes en intra-class korrelasjons koeffisient eller Bland Altman test. Hvilke statistiske metoder man skal bruke ved reliabilitetstudier har vært mye debattert i løpet av de siste årene (Streiner & Norman, 2003).

Innenfor O'Sullivan's klassifisering kan det la seg gjøre å bruke vektet kappa ved at man for eksempel vektet fleksjonsmønster som mer nære et multidireksjonalt mønster med hovedretning i fleksjon enn et ekstensjonsmønster. Da vi valgte å bare sette multidireksjonalt mønster og ikke velge retning når vi scoret har jeg i denne oppgaven ikke brukt vektet kappa. Jeg valgte å heller slå sammen grupper etter smerteretning for å se om det forandret min kappa verdi.

5.3.2 Statistisk usikkerhet

Ved test av sammenheng mellom usikker og uenig valgte jeg en Chi-square test for independence (2 x 2 tabell). For å si noe om sannsynligheten for sammenheng brukes p-verdi. En p-verdi på 0.05 vil si at sannsynligheten for å komme frem til samme resultat eller et bedre resultat ved tilfeldighet er lik 5 prosent. Det er viktig å forstå at p-verdi ikke står for den absolutte sannheten (Altman, 2000).

Det samme forhold gjelder ved bruk av Kappa. Kappa verdiene må ses i den sammenheng den ble testet og tabellene må analyseres. Konfidensintervallet er av stor betydning når man vurderer grad av enighet da det gir et bilde av variasjonen (Altman, 2000). Statistisk usikkerhet i denne metodestudien diskuteres under resultatene.

5.4 Resultater

Hensikten med denne studien var å undersøke interrater reliabiliteten på et nyutviklet klassifikasjonssystem på en selektert gruppe pasienter med uspesifikke kroniske korsryggsmerter. Resultatene vil i dette kapittel bli diskutert opp mot teori og empiri.

5.4.1 Demografiske data

Jeg har valgt å ta med FABQ og HSCL-25 i min presentasjon av demografiske data for å se om gruppen som er inkludert i denne studien skiller seg fra øvrige pasienter med uspesifikke kroniske korsryggsmerter (tabell 4).

Pasientgruppen som inngår i min studie likner pasientene i andre store ryggstudier der pasienter med uspesifikke korsryggsmerter er rekrutterte fra en sykehuspoliklinikk (Fritzell 2001, Brox 2003, Fairbank 2005). Vi fant i vår studie at pasientene scorer høyt på Oswestry, HCSL og VAS ved inklusjon. Dette har til dels sammenheng med de strengere inklusjonskriteriene som brukes der operasjon er en av intervensjonene.

Pasientene i min studie ligger på en gjennomsnittlig HSCL på 1.86. HSCL indikerer uspesifikke psykosomatiske symptomer og plager og ikke nødvendigvis en psykiatrisk diagnose. Skår på ≥ 1.7 ved HSCL-25 indikerer en pasient med stort forbruk av helsetjenester (Sandanger et al., 1998).

I andre ryggstudier der kroniske ryggpasienter har blitt rekruttert fra allmennpraktikere, ses en noe lavere score på HSCL-25 og VAS men lik score på FABQ (Storheim et al., 2003; Vibe et al., 2008). I min studie ligger gjennomsnittet på VAS (ryggsmerter) på 72 mm og HSCL på 1.86 mens i Storheims og Vibe Fersums studier ligger gjennomsnittet på VAS på 55 mm respektive 60 mm og gjennomsnittscore på HSCL på 1.5 respektive 1.53 ved baseline (Storheim et al., 2003; Vibe et al., 2008).

Vibe Fersum har ekskludert pasienter som har vært sykemeldte mer en 4 måneder (Vibe et al., 2008). Her skiller våre utvalg seg betydelig da flere av pasientene i

metodestudien hadde vært sykemeldte lenger enn 4 måneder, 9 pasienter var på rehabiliteringspenger og en var uføretrygdet. Pasientene i Vibe Fersums studie hadde en Oswestry på gjennomsnitt 21 % mens i min studie lå gjennomsnittet på 43 %.

5.4.2 Subgruppering til movement impairment eller control impairment

Interrater reliabiliteten er meget god ved klassifisering til control impairment eller movement impairment (tabell 5). Terapeutene var enige i klassifikasjon av alle pasientene. Dette kan bero på at kriteriene for de to kategoriene er godt definert og at vi ut fra filmen satte klare kriterier. Normalt skal dette klassifiseres ved hjelp av kliniske segmentale leddtester, men vi hadde valgt film og lagde da kriterier ut fra dette. Når pasienten ikke klarte å ta ut bevegelse i smertefull retning og det ble et "stunt stopp" ble den definert som movement impairment. Ved control impairment derimot klarte pasienten å ta ut hele bevegelsen, men den kunne være hakkete, skjev og/eller langsom med stopp midtveis. For å skille disse to gruppene brukes hovedsaklig filmen ved klassifiseringen og man har kun to valgmuligheter og gruppene er klart avgrenset fra hverandre. Få alternativer og klare kriterier bidrar til at man får høy reliabilitet. Diskrete eller diskontinuerlige variabler er lettere å klassifisere enn kontinuerlige (Kibler et al., 2002). Det ble en hovedvekt av pasienter med control impairment (37 av 41). Vibe Fersum bruker samme klassifikasjonssystem på uselekterte ryggpasienter og han har fått relativt lik fordeling i sin studie. Han fant 1 pasient med movement impairment av 26 pasienter (Vibe et al., 2008). I Dankaerts studie var movement impairment et eksklusjonskriterium, da man primært ønsket å se på klassifisering av control impairment til de 5 undergruppene (Dankaerts et al., 2006). Det er derfor vanskelig å si noe om hvordan dette fordeler seg i større grupper med pasienter med uspesifikke kroniske korsryggsmerte.

5.4.3 Subgruppering av pasienter med control impairment til fem ulike kategorier

Interrater reliabiliteten er moderat med en kappa på 0.52 (0.32 - 0.71) ved subgruppering av control impairment til fem ulike kategorier (tabell 6). Dette

stemmer med mange av de andre klassifikasjonssystemene som viser moderat interarterreliabilitet ved subgruppering (Billis et al., 2007; McCarthy et al., 2004). Dankaerts hadde i sin studie en Kappa verdi på 0.55 (0.47 - 0.64) ved interraterreliabilitet blant terapeutene med moderat erfaring med klassifikasjonssystemet til O'Sullivan. Dette ligger på samme nivå som våre resultat, men vi hadde en større spredning 0.32 - 0.71. Dankaerts hadde i sin studie 13 terapeuter som klassifiserte 24 pasienter. Dette gir et større antall pasient tester som vil bidra til en mindre spredning enn i min studie. Den laveste verdien på 0.32 anses som svak reliabilitet og mitt ønske ved start var å få den laveste verdien på moderat nivå dvs. mellom 0.41 - 0.60 og den høyeste på god nivå dvs. mellom 0.61 - 0.80. Et større antall pasienter og/eller flere terapeuter kunne ha strammet inn spredningen i min studie. Til tross for at noen forfattere mener en kappa mellom 0.40 - 0.60, er akseptabel reliabilitet, er det flere forskere som mener at et system eller instrument bør ha en kappa på over 0.75 (god reliabilitet) for å anses som anvendbart (Streiner & Norman, 2003). Andre setter grensen til over 0.85 (svært god reliabilitet) for å kunne vurdere et system tilfredsstillende (McDowell & Newell, 1996). Fersum hadde i sin studie 4 terapeuter som klassifiserte 24 pasienter. Her var alle meget godt trent i klassifikasjonssystemet, en av de fire var dessuten opphavsmannen og man fikk en Kappa på 0.82 (0.66 - 0.90) (Vibe et al., 2008).

Vi var uenige om 13 av 37 pasienter. For 6 av de 13 pasientene hadde vi samme uenighet det vil si at den ene klassifiserte til aktivt ekstensjonsmønster og den andre til fleksjonsmønster. Disse to er motpoler med smerter i helt forskjellige bevegelsesretninger. I forhold til behandling er dette å anse som en uakseptabel uenighet. Dersom man har en høy prosent av samme type uenighet dras kappa verdien ned (tabell 6). En nærmere undersøkelse viste at for de 6 pasientene hvor vi begge uenige var vi usikre på klassifiseringen. Disse pasientene gir også ulike tolknings- muligheter ut fra anamnesen i forhold til på filmen. Det kan se ut til at den ene terapeuten valgte å stole på anamnesen, mens den andre valgte å legge hovedvekt på filmen når man var usikker. Her kommer kanskje en svakhet med selve studiedesignen (filmingen og anamnesen) frem. Ved usikkerhet har man ikke mulighet til å komplettere undersøkelsen med provokasjon eller avlastning for å finne

bedre ut av pasientens smerteretning. Vi har heller ikke mulighet til å stille mer klargjørende spørsmål. Dette viser at til tross for å ha standardisert bevegelsesanalysen og anamnesen kan våre tolkninger av samme observasjon og tekst gir to forskjellige svar.

Uenigheten kan også bero på at våre kriterier for klassifisering til de ulike gruppene ikke er helt gjensidig utelukkende og at en pasient derfor kan passe i flere grupper (Buchbinder et al., 1996). Man kunne ha bestemt i forveien at ved ulikheter i anamnese og film velges for eksempel anamnesen. En annen forklaring kan være at dette er pasienter med multidireksjonale smerter eller at det er pasienter som ikke har et motorcontrol problem som den primære grunnen for deres smerte, som vi ikke klarte å fange opp via film og anamnese.

Ved å slå sammen de 5 subgruppene under MCI til 3 større grupper i analysen, gjennom endrede kriterier for gruppene til fleksjon, ekstensjon og multidireksjonal, ville jeg se om det kunne forbedre kappa verdien (tabell 7). Dette viste seg ikke å være tilfelle, Kappa verdien sank, mens prosentvis enighet økte. Til tross for færre subgrupper og flere pasienter som vi er enige om, har antall uenige i samme celle økt. Det er vanlig at K ikke alltid følger den prosentvise enigheten. Dette skjer ofte ved skjev distribusjon når prevalensen er lav "the low base-rate problem" hvor bare 1 eller 2 uenigheter mellom terapeutene kan redusere K verdien merkbart (Clare et al., 2005).

Da mange ulike fordelinger i en tabell kan gi samme kappa verdi er det derfor viktig å lese tabellene (Altman, 2000). Man bør alltid presentere både prosentvis enighet og kappa verdien som tar høyde for enighet ved sjanse sammen med spredningen som viser hvor stor variasjon det er.

Prevalensen av de ulike mønstrene er vanskelig å si noe om. Min studie inneholder et relativt lite antall pasienter for å si noe om prevalens og gruppen hvor vi scoret likt er for lite til å dra noen konklusjoner. Av de 24 vi scoret likt på var 11 aktivt ekstensjonsmønster, 7 fleksjon med shift, 5 fleksjonsmønster, 1 multidireksjonalt og ingen passiv ekstensjon. I studien til Dankaert var det flest fleksjonsmønster (8), deretter aktivt ekstensjonsmønster (6) og multidireksjonalt (5) og til sist fleksjon med

shift (3) og passiv ekstensjon (3). Også hans gruppe var for liten til å kunne dra noen konklusjoner om prevalens (N 25) (Dankaerts et al., 2006).

Dankaerts fant at fleksjon med shift var lettest å identifisere og gav størst enighet, mens aktiv ekstensjon var vanskeligst. Vi hadde i vår studie størst enighet hos pasienter med fleksjon med shift hvor vi var enige i 70 % av scorene (7 av 10). For gruppen med aktiv ekstensjon var vi enige i 55 % av scorene (11 av 20). Det var størst uenighet om passiv ekstensjon, men her er antallet pasienter lavt. Vibe Fersum fant i sin studie det lettest å subgruppere passivt ekstensjonsmønster og deretter fleksjonsmønster. Han fant aktivt ekstensjonsmønster som vanskeligst å score likt på (Vibe et al., 2008).

Pasientene som inngår i denne metodestudien er et avgrenset utvalg, basert på grad av degenerative funn sentrert til nedre delen av korsryggen, som anses som en homogen gruppe for kirurgi. Ved bruk av O'Sullivan's klassifikasjonssystemet viser det seg å andre siden at dette er en heterogen gruppe med ulike grad av smerte, ulike type smertestrategier og motorisk respons.

5.4.4 Interrater reliabilitet ved bare anamnese sammenlignet med anamnese og observasjon

Interrater reliabiliteten økte dersom man klassifiserte ut fra bare anamnese (tabell 8). Pasientene ble gruppert i 3 hovedgrupper og vi fant da en lik klassifikasjon på alle uten 7 pasienter. Av de 7 vi var uenige om, er 6 i ruten for uenighet om fleksjon eller ekstensjon. Disse er de samme 6 pasientene som vi tidligere var uenige om. Dette kan tolkes til at det ikke bare er divergensen mellom anamnese og film som skaper uenighet men at også anamnesen alene er vanskelig å score likt.

Det er gjort få studier som evaluerer anamneseopptak hos pasienter med ryggsmarter. Van Dillen fant i sin studie fra 1998 at undersøkelse bare basert på observasjon ga dårligere kappa enn ved anamnese pluss observasjon (Van Dillen et al., 1998). Man har i en studie på nerverotsyndrom funnet at anamnesen alene ved diagnosesetting hadde en kappa på 0.40 men økte etter klinisk undersøkelse til 0.66 (Vroomen et al., 2000). Det samme finner Dankaerts i sin reliabilitetsstudie der reliabiliteten sank

betydelig med bare anamnese sammenlignet med anamnese pluss observasjon (Dankaerts et al., 2006). I vår studie derimot fikk vi høyere interrater reliabilitet ved bare scoring av anamnese enn ved både observasjon og anamnese. Divergens mellom mine og Dankaerts resultater kan være forårsaket av rekkefølgen man gjorde subgrupperingen på. Dankaerts subgrupperte først ut fra anamnesen, mens vi startet med subgruppering ut fra både anamnese og videofilm (Dankaerts et al., 2006). Det kan tenkes at man i min studie var mer trent til å subgruppere ut fra anamnesen i andre omgangen og derfor scoret med større enighet. Man har i tidligere studier sett en lavere reliabilitet ved bruk av visuelle informasjonskilder (Strender et al., 1997; Van Dillen et al., 1998). Dette kan stemme med vår studie der vi får lavere kappa når vi kombinerer anamnese og videofilm.

5.4.5 Sammenheng mellom uenig og usikker ved scoring

Som tidligere nevnt var noen pasienter vanskelige å klassifisere. Vi valgte derfor å føye på en rubrikk i scoringsskjemaet hvor terapeuten kunne angi om det var usikkerhet ved scoringen. Ved 8 av 19 usikre scoringer var begge terapeutene usikre på samme pasient (tabell 9). Totalt var vi derfor usikre på 27 av 82 scoringer (33 %). Det ser ut å være en sammenheng mellom usikkerhet og uenig i klassifisering av mønster med en $p = 0.052$ ved chi-square test for independence og en $p = 0.038$ ved Fisher's eksakt test. En usikkerhet på 33 % er relativt høy og innebærer at man må se over årsakene til usikkerheten. En kan tenke seg at usikkerhet hører sammen med terapeutenes erfaring i klassifikasjonssystemet. Det at vi var moderat erfarne i systemet som vi testet kan være bidragene til den moderate reliabiliteten. De fleste studier der man tester interrater reliabiliteten på klassifikasjonssystemer konkluderer med at jo mer trente terapeutene er i systemet de tester desto høyere reliabilitet får man (Clare et al., 2005; Dankaerts et al., 2006; Vibe et al., 2008). En mulig forklaring kan også være at systemet var nytt når vi startet og at det ennå ikke var nok gjennomarbeidet. Systemet har også blitt revidert etter at vi startet studien. Sammenhengen mellom usikker og uenig kan kanskje forklare den moderate reliabiliteten. Det viser seg å være de samme pasientene som vi er uenige om som vi også er usikre på ved scoring. Som jeg tidligere har diskutert kan det tenkes at disse

ikke er verken movement impairment eller control impairment det vil si ikke har et problem med motorisk kontroll som den primære grunnen for deres smerte. Dette kan man ikke kontrollere for i ettertid. Å redusere usikkerheten ville kanskje innebære en redusert uenigheten som i sin tur ville kunne øke kappa verdien.

5.5 Interrater reliabilitet

I 2006 ble en review artikkel publisert som systematisk undersøker reliabiliteten på ulike fysioterapeutiske undersøkelsesmetoder som er vanlig brukte ved uspesifikke korsryggsmerter (May et al., 2006). Man undersøkte palpasjon, observasjon, symptom respons og klassifikasjons system. De fleste studiene brukte kappa eller ICC. Generelt var reliabiliteten på moderat som det beste. Den metodiske kvaliteten var også bare moderat. Mange vanlig brukte undersøkelsesmetoder hadde meget lav reliabilitet eller hadde sprikende bevis på sin reliabilitet. Dette sier noe om hvor vanskelig det er å reliabilitetsteste fysioterapeutiske undersøkelsesmetoder.

I artikkelen ble følgende retningslinjer satt opp for bedret reliabilitet (May et al., 2006):

1. Deltakerne bør være pasienter med den studerte tilstand, representative for klinisk praksis og rekruttert tilfeldig eller fortløpende
2. Prosedyrer bør ha en klar anvendbar definisjon med entydig og reproduserbar beskrivelse
3. Flere målinger bør gjøres for å redusere bias, undersøkerne må være blindet i forhold til hverandre; resultatene forsegle; og at det er en uavhengig bedømmer
4. Stort antall pasienter, minst 50 og mange par med testere
5. En standardisert måling av testresultatet
6. Frekvens av resultat og enighet må rapporteres så at prevalens og bias kan vurderes
7. Den korrekte statistikken bør rapporteres med spredningsmål
8. Studien bør rapporteres etter STARDs konvensjoner (Bossuyt et al. 2003)

Forslagene til May gjør det mulig å bedre reliabiliteten ved å metodisk være korrekt og fjerne mest mulig med støy. Hvis jeg ser på min studie kunne den styrkes ved at jeg hadde brukt flere terapeuter med multiple tester mellom terapeutene og hatt større antall pasienter. Det å bare øke antall pasienter er ofte ikke nok for å styrke interrater reliabiliteten (May et al 2006).

Interrater reliabiliteten i min studie ligger på Kappa lik 1.0 som anses å være perfekt enighet på å skille mellom control impairmet og movement impairment . Ved subgruppering til de 5 bevegelsesmønstrene ligger Kappa på 0.52 som kan anses som en moderat enighet. Svaret på min problemstilling om systemet til O'Sullivan er reliabelt må derfor bli delt. Jeg har ikke har testet hele O'Sullivans klassifikasjonssystemet og har en selektert gruppe av uspesifikke LBP. Ut fra disse rammene ser det ut som om det var lettere å skille movement impairment fra control impairment, men at subgrupperingen til de fem smerteprovoserende bevegelsesmønstrene ikke var like lett.

Man kunne tenke seg at pasientene i min studie var vanskeligere å klassifisere siden de har lenger symptom varighet og et betydelig lavere funksjonsnivå enn pasientene i Vibe Fersum's studie (Vibe et al., 2008). En langvarig kronifisering kan medføre at pasientene får sekundære plager som gir et mer uklart klinisk bilde, altså at de blir vanskeligere å klassifisere. Mine resultat viser at dette fenomenet i noen grad kan ha inntruffet.

Reliabiliteten i min metodestudie er på nivå med en tidligere publisert studie som også inkluderte uspesifikke kroniske ryggpasienter med lavt funksjonsnivå og mangeårige ryggplager (Dankaerts et al., 2006). Begge disse viser moderat reliabilitet når moderat erfarne terapeuter klassifiserer. I Dankaerts studie har imidlertid terapeutene med lang erfaring like god reliabilitet som de erfarne terapeutene i studien til Vibe Fersum. Dette indikerer at grad av erfaring har mer å si enn grad av funksjonsnedsettelse og at det teoretiske grunnlaget som O'Sullivan bygger på viser seg også hos pasienter med mangeårige ryggmerter.

Den moderate reliabiliteten kan som tidligere nevnt bety at terapeutene ikke er erfarne nok i klassifiseringssystemet eller at klassifikasjonssystemet ikke er optimalt

og at noen pasienter faller utenfor klassifikasjonssystemet. Utvalget og metoden har også blitt diskutert som mulig årsak til den moderate reliabiliteten. Det er vanskelig å plukke ut hvilken eller hvilke av disse faktorene som påvirket resultatene i min metodestudie mest.

Det viktigste ved min metode studie er at klassifikasjonssystemet har blitt testet av andre enn opphavsmiljøet og at jo flere reliabilitetstester som utføres desto større er muligheten for å avdekke svakheter ved systemet som kan forbedres.

5.6 Kliniske implikasjoner

Det finnes mange ulike systemer for subgruppering av pasienter med uspesifikk kroniske korsryggmerter, men ingen av dem oppfyller alle krav til anvendbarhet og metodisk kvalitet. De fleste klassifikasjonssystemene blir fort for kompliserte å bruke i den daglige klinikken (Machado et al., 2006; Petersen et al., 2003). Spesielt systemer hvor det er undergrupper som krever mange spesifikke undersøkelser. Strukturorienterte diagnostiske klassifikasjonssystemer som det til Petersen er mest i bruk for klinisk forskning da det er meget tidskrevende og lite klinisk anvendbart. Dersom pasienten går til behandling kan de testes underveis, men for en konsultasjon med anamnese i tillegg blir de for kompliserte og tidkrevende (Werneke & Hart, 2001). Noen av de behandlingsrelaterte klassifikasjonssystemene er allment kjent blant fysioterapeuter og i bruk i flere land, som Mc Kenzie klassifikasjon. Dette systemet krever imidlertid en spesialisert videreutdanning i Mc Kenzie metoden (Machado et al., 2006). Generaliserbarheten er usikker da det ser ut som klassifikasjonssystemene krever såpass mye veiledning og utdanning innenfor de spesielle behandlingsmetodene at det blir de som utvikler dem og deres ”nærmeste krets” som bruker dem. Opplæring skal ikke kreve mer tid enn at det lar seg kombinere i dagens hektiske kliniske hverdag (Heiss et al., 2004).

Det er mange ulike forskningsmiljøer som over lang tid har innarbeidet tradisjoner og en kultur som gjør det vanskelig å samordne et felles klassifikasjonssystem (Billis et al., 2007). Grunnen til at klassifisering av LBP fortsatt ikke er allment kjent kan være

at det ennå ikke er noen system som er allment akseptert. Det bør også nevnes at de nasjonale kliniske retningslinjer for korsryggsmerter ikke legger noen nevneverdig vekt på klassifikasjons metoder (Lærum et al., 2007).

Det er først de siste årene at man har laget effektstudier i relasjon til subgruppering og fått positive resultat (Brennan et al., 2006; Browder et al., 2007; Gudavalli et al., 2006; Vollenbroek-Hutten et al., 2004). Jeg mener det viktigste i dag er å spre kunnskap om behovet for kartlegging av subgrupper og oppmuntre til videre forskning på dette området. Her er det mange små brikker som skal settes sammen til en helhet.

Jeg tror det vil være vanskelig å finne et klassifiseringssystem som kan tas i bruk i sin helhet av forskjellige yrkesgrupper. Derimot ser jeg at et systemet, som eksempelvis det til O'Sullivan, kan brukes i sin helhet av non medisinske yrkesgrupper (manuellterapeuter med rett til primærkontakt) og delvis av andre medisinske yrkesgrupper (lege, fysioterapeut, kiropraktor m.fl.). Den første delen av subgrupperingen kan brukes av primærkontakten. Her bruker man stort sett den diagnostiske triaden, røde flagg og gule flagg. Dersom pasienten presenterer en tilstand med akutte uspesifikke korsryggsmerter kan primærkontakten følge retningslinjene med å avdramatisere, gi generelle råd om aktivitet og arbeide. Ved tilbakefall eller ved risiko for langvarige plager, henvises pasientene til behandlere (eventuelt primærkontakten selv) som ytterligere klassifiserer og gir riktig intervensjon ut fra pasientens individuelle behov. Før O'Sullivans klassifikasjonssystem kan kalles reliabelt og valid må flere store RCT studier gjøres og det må testes i flere fagmiljøer og i flere kulturer (Billis 2007).

Vi vet at faktorer som kjønn, personlighet, motivasjon, familiære forhold, kultur og tidligere episoder av korsryggsmerter kan tenkes å innvirke på smertetilstanden og medføre at pasienter responderer ulikt på samme behandling (Wadell 2004, Billis 2006). Det er derfor viktig å videreutvikle klassifikasjonssystemet slik at det ivaretar flest mulige faktorer som påvirker ryggsmertene for ikke å miste helhetssynet. Det er også viktig å kunne vite noe om forventet behandling og behandlingstid for de ulike faktorene som innvirker på smerten. En systematisering av ryggbehandlingen kunne

bidra til at pasientene kommer fortere frem i systemet og at de slipper unødvendige undersøkniger og behandlinger. Et velfungerende klassifikasjonssystem kan være til gode for pasienten som dermed blir mer ivaretatt ut fra sine individuelle behov. Det har også stor betydning for behandlingsapparatet som kan gi pasienten mer målrettet tilbakemeldning, behandling og oppfølging. Ikke minst gir et velutviklet klassifikasjonssystem mulighet for arbeidsgivere og trygdevesen til tidlig å identifisere og sette i gang tiltak for pasienter med risiko for å utvikle langvarige plager (Childs et al., 2004). Hvis vi får en mer enhetlig forklaringsmodell og kommunikasjon mellom yrkesgruppene kan disse formidle større troverdighet og trygghet for pasientene. Behandlernes trygghet har stor betydning for hvordan pasienten takler sin lidelse (Coudeyre et al., 2006; Poiraudau et al., 2006; Houben et al., 2005)

6. Oppsummering og konklusjon

Hensikten med denne studien var å undersøke interrater reliabiliteten av et behandlingsrelatert klassifikasjonssystem ad modum O'Sullivan for pasienter med uspesifikke kroniske korsryggsmerter.

Studien viser at erfarne klinikere fra samme klinikk, men med moderat erfaring i klassifiseringssystemet, klassifiserer med moderat kappa ut fra anamnese og videofilm. Nedre del av vårt konfidensintervall gir Kappa verdier som er under det man vanligvis anser som akseptabelt mens øvre del av konfidensintervallet er regnet som god Kappa (tabell 2 og tabell 6). Andre studier som bruker video som metode ved klassifikasjon viser liknende resultat (Dankaerts et al., 2006; Delitto et al., 1995; Fritz et al., 2000; Van Dillen et al., 1998).

Klassifikasjonssystemet til O'Sullivan er i to studier testet for interrater reliabilitet. I den ene studien er deler av klassifikasjonssystemet testet basert på video (Dankaerts et al., 2006). Den videobaserte delen stemmer med våre resultat, men har et mindre konfidens intervall. Dankaert et al. har i sin studie brukt flere terapeuter og får da flere test tilfelle som kan gi mindre konfidens intervall (Dankaerts et al., 2006; May et al., 2006). Den delen av studien som bruker klinisk undersøkelse istedenfor video viser høyere interrater reliabilitet i Dankarts studie (Dankaerts et al., 2006). De samme resultatene viser en nylig presentert studie som også har brukt anamnese og klinisk undersøkelse istedenfor video (Vibe et al., 2008). I begge disse studiene får man høyere reliabilitet når terapeutene er meget erfarne og godt trente i klassifikasjonssystemet. Grad av funksjonsnedsettelse ser ikke ut å ha betydning for reliabiliteten. Felles for disse to studiene er imidlertid at den som har vært ansvarlig for utviklingen av systemet også har deltatt aktivt i studiene. Vår studie er derfor den første uavhengige studien hvor O'Sullivans system ble reliabilitetstestet.

Utvalget i min metodestudie er en del av en større nasjonal ryggstudie der kirurgi er en av intervensjonene. Det blir dermed en avgrenset gruppe blant uspesifikke kroniske ryggpasienter som er inkludert basert på grad av degenerative forandringer. Systemet er derfor ikke testet på et tilfeldig utvalg av pasienter med uspesifikke

korsryggsmerter. Klassifikasjonssystemet har de siste 3 årene blitt utviklet betydelig og min studie omfatter bare en del av systemet slik det fremstår i dag.

I alt viste denne studien svært god enighet ved subgruppering av control impairment og movement impairment $k = 1.0$ og moderat enighet for subgruppering av control impairment til 5 bevegelsesmønstre $k = 0.52$ (0.32 - 0.71). Det ble lavere kappa ved sammenslåing av bevegelsesmønstre i 3 grupper (fleksjon, ekstensjon og multidir.) $k = 0.44$ (0.20-0.68) men en høyere enighet ved klassifisering av bare anamnesen i 3 grupper (fleksjon, ekstensjon og multidir.) $k = 0.66$ (0.44 - 0.87). Dette kan tyde på at videofilmen gir større sjanse for uenighet. Andre studier har imidlertid vist at testing med bare anamnese kommer dårligere ut enn anamnese pluss klinisk undersøkelse (Dankaerts et al., 2006; Vroomen et al., 2000). Det er derfor mulig at metoden videofilm ikke er den beste ved denne type klassifisering. Klinisk undersøkelse ved klassifisering er det vi normalt gjør i klinikken og er det vi er mest trent på. Mange av kriteriene til de ulike gruppene ved klassifiseringen er laget med tanke på klinisk undersøkelse.

Studien viste også at det var en sammenheng mellom uenighet og usikkerhet ved klassifisering til mønstre. Vi var sammenlagt usikre på 33 % av klassifiseringene. Det er derfor av stor betydning at subgruppene er klart definerte og gjensidig utelukkende for å minske mest mulig usikkerhet. Det er de samme pasientene (N=6) som ble klassifiserte til aktiv ekstensjon av den ene terapeuten og til fleksjon av den andre. Disse er to helt forskjellige smerteprovoserende bevegelsesmønstre. Da vi ikke har en gull standard og ingen andre har undersøkt pasientene blir det vanskelig å si noe mer presist om hvor feilen ligger. I studien til Dankerts og Vibe Fersum ser man at aktivt ekstensjonsmønstre er vanskeligst å score likt og at aktivt ekstensjonsmønstre, også i deres studie ved ulik score, ble scoret som fleksjonsmønstre (Dankaerts et al., 2006; Vibe et al., 2008).

Klassifikasjonssystemet til O'Sullivan trenger flere studier for å utvikles til et anvendbart og generaliserbart redskap. Det er også viktig å få flere uavhengige fagmiljøer til å ta i bruk og teste klassifikasjonssystemet.

7. Referanser

- Abenhaim, L., Rossignol, M., Valat, J. P., Nordin, M., Avouac, B., Blotman, F. et al. (2000). *The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine, 25, 1S-33S.*
- Albert, H. B., Kjaer, P., Jensen, T. S., Sorensen, J. S., Bendix, T., & Manniche, C. (2008). *Modic changes, possible causes and relation to low back pain. Medical Hypotheses, 70, 361-368.*
- Altman, D. G. (2000). *Practical statistics for medical research. (2nd ed ed.) London ; New York: Chapman and Hall.*
- Atlas, S. J., Deyo, R. A., Patrick, D. L., Convery, K., Keller, R. B., & Singer, D. E. (1996). *The Quebec Task Force classification for Spinal Disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. Spine, 21, 2885-2892.*
- Baker, D., Pynsent, P., & Fairbank, J. (1990). *The Oswestry Disability Index revisited. In M.Roland & J. Jenner (Eds.), Back Pain: new approaches to rehabilitation and education (pp. 175-181). Manchester: Manchester University Press.*
- BenDebba, M., Torgerson, W. S., & Long, D. M. (2000). *A validated, practical classification procedure for many persistent low back pain patients. Pain, 87, 89-97.*
- Bergmark, A. (1989). *Stability of the lumbar spine. A study in mechanical engineering. Acta Orthop.Scand.Suppl, 230, 1-54.*
- Bernhardt, J., Bate, P. J., & Matyas, T. A. (1998). *Accuracy of observational kinematic assessment of upper-limb movements. Physical Therapy, 78, 259-270.*
- Billis, E. V., McCarthy, C. J., & Oldham, J. A. (2007). *Subclassification of low back pain: a cross-country comparison. European Spine Journal, 16, 865-879.*
- Borkan, J., Van, T. M., Reis, S., Schoene, M. L., Croft, P., & Hermoni, D. (2002). *Advances in the field of low back pain in primary care: a report from the fourth international forum. Spine, 27, E128-E132.*
- Borkan, J. M., Koes, B., Reis, S., & Cherkin, D. C. (1998). *A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. Spine, 23, 1992-1996.*

- Brant-Zawadzki, M. N., Jensen, M. C., Obuchowski, N., Ross, J. S., & Modic, M. T. (1995). *Interobserver and intraobserver variability in interpretation of lumbar disc abnormalities. A comparison of two nomenclatures. Spine, 20, 1257-1263.*
- Brennan, G. P., Fritz, J. M., Hunter, S. J., Thackeray, A., Delitto, A., & Erhard, R. E. (2006). *Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. Spine, 31, 623-631.*
- Browder, D. A., Childs, J. D., Cleland, J. A., & Fritz, J. M. (2007). *Effectiveness of an extension-oriented treatment approach in a subgroup of subjects with low back pain: a randomized clinical trial. Physical Therapy, 87, 1608-1618.*
- Brox, J. I., Reikeras, O., Nygaard, O., Sorensen, R., Indahl, A., Holm, I. et al. (2006). *Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. Pain, 122, 145-155.*
- Buchbinder, R., Goel, V., Bombardier, C., & Hogg-Johnson, S. (1996). *Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? Journal of Clinical Epidemiology, 49, 141-149.*
- Buchbinder, R., Goel, V., & Bombardier, C. (1994). *A methodological framework for the critical appraisal of classification systems. Toronto: Institute for Work and Health = Institut de recherche sur le travail et la santé.*
- Carlsson, B. (1990). *Grundläggande forskningsmetodik för medicin och beteendevetenskap. (andra upplagan ed.) Stockholm: Liber.*
- Carragee, E. J., Alamin, T. F., Miller, J. L., & Carragee, J. M. (2005). *Discographic, MRI and psychosocial determinants of low back pain disability and remission: a prospective study in subjects with benign persistent back pain. Spine J., 5, 24-35.*
- Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. et al. (2004). *A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. Annals of Internal Medicine, 141, 920-928.*
- Clare, H. A., Adams, R., & Maher, C. G. (2003). *Reliability of detection of lumbar lateral shift. J.Manipulative Physiol Ther., 26, 476-480.*
- Clare, H. A., Adams, R., & Maher, C. G. (2005). *Reliability of McKenzie classification of patients with cervical or lumbar pain. J.Manipulative Physiol Ther., 28, 122-127.*
- Coudeyre, E., Rannou, F., Tubach, F., Baron, G., Coriat, F., Brin, S. et al. (2006). *General practitioners' fear-avoidance beliefs influence their management of patients with low back pain.[see comment]. Pain, 124, 330-337.*

- Critchley, D. J. & Coutts, F. J. (2002). *Abdominal muscle function in chronic low back pain patients: measurement with real-time ultrasound scanning. Physiotherapy, 88, 322-332.*
- Dankaerts, W., O'Sullivan, P. B., Burnett, A. F., & Straker, L. M. (2007). *The use of a mechanism-based classification system to evaluate and direct management of a patient with non-specific chronic low back pain and motor control impairment--a case report. Man.Ther., 12, 181-191.*
- Dankaerts, W., O'Sullivan, P. B., Straker, L. M., Burnett, A. F., & Skouen, J. S. (2006). *The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. Man.Ther., 11, 28-39.*
- Delitto, A., Erhard, R. E., & Bowling, R. W. (1995). *A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Physical Therapy, 75, 470-485.*
- Delitto, A., Shulman, A. D., Rose, S. J., Strube, M. J., Erhard, R. E., Bowling, R. W. et al. (1992). *Reliability of a clinical examination to classify patients with low back syndrome. Phys Ther Pract, 1, 1-9.*
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). *The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behavioral Science, 19, 1-15.*
- Deutsch, A., Altchek, D. W., Schwartz, E., Otis, J. C., & Warren, R. F. (1996). *Radiologic measurement of superior displacement of the humeral head in the impingement syndrome. Journal of Shoulder & Elbow Surgery, 5, 186-193.*
- Dillingham, T. (1995). *Evaluation and management of low back pain: an overview. State of the Art Reviews, 9, 559-574.*
- Domholdt, E. (2000). *Physical therapy research principles and applications. (2nd ed ed.) Philadelphia: Saunders.*
- Eastlack, M. E., Arvidson, J., Snyder-Mackler, L., Danoff, J. V., & McGarvey, C. L. (1991). *Interrater reliability of videotaped observational gait-analysis assessments. Physical Therapy, 71, 465-472.*
- Essendrop, M., Maul, I., Laubli, T., Riihimaki, H., & Schibye, B. (2002). *Measures of low back function: a review of reproducibility studies. Clin.Biomech.(Bristol., Avon.), 17, 235-249.*
- Feinstein, A. R. (1987). *Clinimetrics. New Haven: Yale University Press.*
- Fritz, J. M., Cleland, J. A., & Childs, J. D. (2007). *Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. Journal of Orthopaedic and Sports Physical Therapy, 37, 290-302.*

- Fritz, J. M., Delitto, A., & Erhard, R. E. (2003). *Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. Spine, 28, 1363-1371.*
- Fritz, J. M., Delitto, A., Vignovic, M., & Busse, R. G. (2000). *Interrater reliability of judgments of the centralization phenomenon and status change during movement testing in patients with low back pain. Archives of Physical Medicine and Rehabilitation, 81, 57-61.*
- Fritz, J. M. & George, S. (2000). *The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. Spine, 25, 106-114.*
- Fritzell, P., Hagg, O., Wessberg, P., & Nordwall, A. (2001). *2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. Spine, 26, 2521-2532.*
- George, S. Z., Fritz, J. M., Bialosky, J. E., & Donald, D. A. (2003). *The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. Spine, 28, 2551-2560.*
- Grotle, M., Brox, J. I., & Vollestad, N. K. (2003). *Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. J.Rehabil.Med., 35, 241-247.*
- Grotle, M., Brox, J. I., & Vollestad, N. K. (2006). *Reliability, validity and responsiveness of the fear-avoidance beliefs questionnaire: methodological aspects of the Norwegian version. J.Rehabil.Med., 38, 346-353.*
- Gudavalli, M. R., Cambron, J. A., McGregor, M., Jedlicka, J., Keenum, M., Ghanayem, A. J. et al. (2006). *A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for chronic low back pain. European Spine Journal, 15, 1070-1082.*
- Heiss, D. G., Fitch, D. S., Fritz, J. M., Sanchez, W. J., Roberts, K. E., & Buford, J. A. (2004). *The interrater reliability among physical therapists newly trained in a classification system for acute low back pain. Journal of Orthopaedic and Sports Physical Therapy, 34, 430-439.*
- Hestbaek, L., Leboeuf-Yde, C., & Manniche, C. (2003). *Low back pain: what is the long-term course? A review of studies of general patient populations. European Spine Journal, 12, 149-165.*
- Hickey, B. W., Milosavljevic, S., Bell, M. L., & Milburn, P. D. (2007). *Accuracy and reliability of observational motion analysis in identifying shoulder symptoms. Man.Ther., 12, 263-270.*

- Hides, J. A., Jull, G. A., & Richardson, C. A. (2001). Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine*, 26, E243-E248.
- Hides, J. A., Richardson, C. A., & Jull, G. A. (1996). Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. *Spine*, 21, 2763-2769.
- Hodges, P. W. & Moseley, G. L. (2003). Pain and motor control of the lumbopelvic region: effect and possible mechanisms. *Journal of Electromyography and Kinesiology*, 13, 361-370.
- Hodges, P. W., Moseley, G. L., Gabrielsson, A., & Gandevia, S. C. (2003). Experimental muscle pain changes feedforward postural responses of the trunk muscles. *Experimental Brain Research*, 151, 262-271.
- Hodges, P. W. & Richardson, C. A. (1996). Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. *Spine*, 21, 2640-2650.
- Hodges, P. W. & Richardson, C. A. (1999). Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. *Archives of Physical Medicine and Rehabilitation*, 80, 1005-1012.
- Houben, R. M., Ostelo, R. W., Vlaeyen, J. W., Wolters, P. M., Peters, M., & Stomp-van den Berg SG (2005). Health care providers' orientations towards common low back pain predict perceived harmfulness of physical activities and recommendations regarding return to normal activity. *European Journal of Pain: Ejp*, 9, 173-183.
- Indahl, A., Velund, L., & Reikeraas, O. (1995). Good prognosis for low back pain when left untampered. A randomized clinical trial. *Spine*, 20, 473-477.
- Jarvik, J. G. & Deyo, R. A. (2002). Diagnostic evaluation of low back pain with emphasis on imaging. *Annals of Internal Medicine*, 137, 586-597.
- Jensen, T. S., Sorensen, J. S., & Kjaer, P. (2007). Intra- and interobserver reproducibility of vertebral endplate signal (modic) changes in the lumbar spine: the Nordic Modic Consensus Group classification. *Acta Radiologica*, 48, 748-754.
- Karoly, P. & Jensen, M. P. (1987). Multimethod assessment of chronic pain.
- Kelly, B. T., Williams, R. J., Cordasco, F. A., Backus, S. I., Otis, J. C., Weiland, D. E. et al. (2005). Differential patterns of muscle activation in patients with symptomatic and asymptomatic rotator cuff tears. *Journal of Shoulder & Elbow Surgery*, 14, 165-171.
- Kent, P. & Keating, J. (2004). Do primary-care clinicians think that nonspecific low back pain is one condition? *Spine*, 29, 1022-1031.

- Kibler, W. B., Uhl, T. L., Maddux, J. W., Brooks, P. V., Zeller, B., & McMullen, J. (2002). *Qualitative clinical evaluation of scapular dysfunction: a reliability study*. *J.Shoulder.Elbow.Surg.*, 11, 550-556.
- Kilby, J., Stigant, M., & Roberts, A. (1990). *The reliability of back pain assessment by physiotherapists, using a "McKenzie algorithm"*. *Physiotherapy*, 76, 579-583.
- Kilpikoski, S., Airaksinen, O., Kankaanpaa, M., Leminen, P., Videman, T., & Alen, M. (2002). *Interexaminer reliability of low back pain assessment using the McKenzie method*. *Spine*, 27, E207-E214.
- Kjaer, P., Albert, H., Jensen, T. S., Leboeuf-Yde, C., Bendix, T., Wedderkopp, N. et al. (2006a). *[Back pain, radiology and end plate changes by means of Modic]*. *Ugeskrift for Laeger*, 168, 1668-1669.
- Kjaer, P., Korsholm, L., Bendix, T., Sorensen, J. S., & Leboeuf-Yde, C. (2006b). *Modic changes and their associations with clinical findings*. *European Spine Journal*, 15, 1312-1319.
- Koes, B. W., van Tulder, M. W., Ostelo, R., Kim, B. A., & Waddell, G. (2001). *Clinical guidelines for the management of low back pain in primary care: an international comparison*. *Spine*, 26, 2504-2513.
- Krebs, D. E., Edelstein, J. E., & Fishman, S. (1985). *Reliability of observational kinematic gait analysis*. *Physical Therapy*, 65, 1027-1033.
- Lærum, E., Brox, J. I., Espeland, A., Storheim, K., & og, h. (2007). *Korsryggsmerter med og uten nerverotaffeksjon [Computer software]*. Oslo: Formi.
- Landis, J. R. & Koch, G. G. (1977). *The measurement of observer agreement for categorical data*. *Biometrics*, 33, 159-174.
- Lin, J. J., Hanten, W. P., Olson, S. L., Roddey, T. S., Soto-quijano, D. A., Lim, H. K. et al. (2005). *Functional activity characteristics of individuals with shoulder dysfunctions*. *Journal of Electromyography and Kinesiology*, 15, 576-586.
- Loisel, P., Vachon, B., Lemaire, J., Durand, M. J., Poitras, S., Stock, S. et al. (2002). *Discriminative and predictive validity assessment of the quebec task force classification*. *Spine*, 27, 851-857.
- Long, D. M., BenDebba, M., Torgerson, W. S., Boyd, R. J., Dawson, E. G., Hardy, R. W. et al. (1996). *Persistent back pain and sciatica in the United States: patient characteristics*. *Journal of Spinal Disorders*, 9, 40-58.
- Machado, L. A., de Souza, M. S., Ferreira, P. H., & Ferreira, M. L. (2006). *The McKenzie method for low back pain: a systematic review of the literature with a meta-analysis approach*. *Spine*, 31, E254-E262.

- Maluf, K. S., Sahrman, S. A., & Van Dillen, L. R. (2000). *Use of a classification system to guide nonsurgical management of a patient with chronic low back pain.[see comment]. Physical Therapy, 80, 1097-1111.*
- Manniche, C. (1989). *[Back exercises for patients with chronic low back pain]. Ugeskrift for Laeger, 151, 2009-2012.*
- Manniche, C., Lundberg, E., Christensen, I., Hesselsoe, G., & Bentzen, L. (1989). *[Intensive dynamic back exercises for patients with chronic low back pain. A controlled study]. Ugeskrift for Laeger, 151, 2030-2034.*
- May, S., Littlewood, C., & Bishop, A. (2006). *Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review. Australian Journal of Physiotherapy, 52, 91-102.*
- McCarthy, C. J., Arnall, F. A., Strimpakos, N., Freemont, A., & Oldham, J. A. (2004). *The biopsychosocial classification of non-specific low back pain: a systematic review. Physical Therapy Reviews, 9, 17-30.*
- McDowell, I. & Newell, C. (1996). *Measuring health a guide to rating scales and questionnaires. (2nd ed ed.) New York: Oxford University Press.*
- Milette, P. C., Fontaine, S., Lepanto, L., Cardinal, E., & Breton, G. (1999). *Differentiating lumbar disc protrusions, disc bulges, and discs with normal contour but abnormal signal intensity. Magnetic resonance imaging with discographic correlations. Spine, 24, 44-53.*
- Moseley, G. L. & Hodges, P. W. (2005). *Are the changes in postural control associated with low back pain caused by pain interference? Clinical Journal of Pain, 21, 323-329.*
- Moseley, G. L., Hodges, P. W., & Gandevia, S. C. (2002). *Deep and superficial fibers of the lumbar multifidus muscle are differentially active during voluntary arm movements. Spine, 27, E29-E36.*
- Nachemson, A. (1999). *Back pain: delimiting the problem in the next millennium. International Journal of Law and Psychiatry, 22, 473-490.*
- NAV (2008). *Sykefraværstatistikk pr. 4. kvartal 2007. NAV [On-line]. Available: <http://www.nav.no/Om+NAV/Tall+og+analyse/Jobb+og+helse/Sykefrav%C3%A6r/Sykefrav%C3%A6rsstatistikk>*
- Niemisto, L., Lahtinen-Suopanki, T., Rissanen, P., Lindgren, K. A., Sarna, S., & Hurri, H. (2003). *A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. Spine, 28, 2185-2191.*

- O'Sullivan, P. (2005). *Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. [Review] [63 refs]. Manual Therapy, 10, 242-255.*
- O'Sullivan, P. B. (2000). *Lumbar segmental 'instability': clinical presentation and specific stabilizing exercise management. Man.Ther., 5, 2-12.*
- O'Sullivan, P. B. & Beales, D. J. (2007a). *Diagnosis and classification of pelvic girdle pain disorders, Part 2: illustration of the utility of a classification system via case studies. Man.Ther., 12, e1-12.*
- O'Sullivan, P. B. & Beales, D. J. (2007b). *Diagnosis and classification of pelvic girdle pain disorders--Part 1: a mechanism based approach within a biopsychosocial framework. Man.Ther., 12, 86-97.*
- O'Sullivan, P. B., Grahamslaw, K. M., Kendell, M., Lapenskie, S. C., Moller, N. E., & Richards, K. V. (2002). *The effect of different standing and sitting postures on trunk muscle activity in a pain-free population. Spine, 27, 1238-1244.*
- O'Sullivan, P. B., Phytty, G. D., Twomey, L. T., & Allison, G. T. (1997). *Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. Spine, 22, 2959-2967.*
- Panjabi, M., Abumi, K., Duranceau, J., & Oxland, T. (1989). *Spinal stability and intersegmental muscle forces. A biomechanical model. Spine, 14, 194-200.*
- Panjabi, M. M. (1992a). *The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. Journal of Spinal Disorders, 5, 383-389.*
- Panjabi, M. M. (1992b). *The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. Journal of Spinal Disorders, 5, 390-396.*
- Peat, J. K., Xuan, W., Mellis, C., & Williams, K. (2002). *Health science research a handbook of quantitative methods. London: Sage.*
- Peng, B., Hou, S., Wu, W., Zhang, C., & Yang, Y. (2006). *The pathogenesis and clinical significance of a high-intensity zone (HIZ) of lumbar intervertebral disc on MR imaging in the patient with discogenic low back pain. European Spine Journal, 15, 583-587.*
- Petersen, T., Laslett, M., Thorsen, H., Manniche, C., Ekdahl, C., & Jacobsen, S. (2003). *Diagnostic classification of non-specific low back pain. Physiotherapy Theory and Practice, 19, 213-237.*

- Petersen, T., Olsen, S., Laslett, M., Thorsen, H., Manniche, C., Ekdahl, C. et al. (2004). *Inter-tester reliability of a new diagnostic classification system for patients with non-specific low back pain. Australian Journal of Physiotherapy, 50, 85-94.*
- Petersen, T., Thorsen, H., Manniche, C., & Ekdahl, C. (1999). *Classification of non-specific low back pain: a review of the literature on classifications systems relevant to physiotherapy. Physical Therapy Reviews, 4, 265-281.*
- Poiraudeau, S., Rannou, F., Baron, G., Le, H. A., Coudeyre, E., Rozenberg, S. et al. (2006). *Fear-avoidance beliefs about back pain in patients with subacute low back pain.[see comment]. Pain, 124, 305-311.*
- Polit, D. F. & Beck, C. T. (2004). *Nursing research principles and methods. (7th ed ed.) Philadelphia, Pa: Lippincott Williams & Wilkins.*
- Rasmussen-Barr, E., Nilsson-Wikmar, L., & Arvidsson, I. (2003). *Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. Man.Ther., 8, 233-241.*
- Riddle, D. L. (1998). *Classification and low back pain: a review of the literature and critical analysis of selected systems. Physical Therapy, 78, 708-737.*
- Riddle, D. L. & Rothstein, J. M. (1993). *Intertester reliability of McKenzie's classifications of the syndrome types present in patients with low back pain. Spine, 18, 1333-1344.*
- Sahrmann, S. A. (1988). *Diagnosis by the physical therapist--a prerequisite for treatment. A special communication. Physical Therapy, 68, 1703-1706.*
- Sahrmann, S. (2002). *Diagnosis and treatment of movement impairment syndromes. St. Louis: Mosby.*
- Sandanger, I., Moum, T., Ingebrigtsen, G., Dalgard, O. S., Sorensen, T., & Bruusgaard, D. (1998). *Concordance between symptom screening and diagnostic procedure: the Hopkins Symptom Checklist-25 and the Composite International Diagnostic Interview I. Social Psychiatry and Psychiatric Epidemiology, 33, 345-354.*
- Schellhas, K. P., Pollei, S. R., Gundry, C. R., & Heithoff, K. B. (1996). *Lumbar disc high-intensity zone. Correlation of magnetic resonance imaging and discography. Spine, 21, 79-86.*
- Siepe, C. J., Mayer, H. M., Wiechert, K., & Korge, A. (2006). *Clinical results of total lumbar disc replacement with ProDisc II: three-year results for different indications. Spine, 31, 1923-1932.*
- Soukup, M. G. & Vollestad, N. K. (2001). *Classification of problems, clinical findings and treatment goals in patients with low back pain using the ICDH-2 beta-2. Disability and Rehabilitation, 23, 462-473.*

-
- Sterling, M., Jull, G., & Wright, A. (2001). The effect of musculoskeletal pain on motor activity and control. J.Pain, 2, 135-145.*
- Storheim, K., Brox, J. I., Holm, I., Koller, A. K., & Bo, K. (2003). Intensive group training versus cognitive intervention in sub-acute low back pain: short-term results of a single-blind randomized controlled trial. J.Rehabil.Med., 35, 132-140.*
- Streiner, D. L. & Norman, G. R. (2003). Health measurement scales a practical guide to their development and use. (3rd ed ed.) Oxford: Oxford University Press.*
- Strender, L. E., Sjoblom, A., Sundell, K., Ludwig, R., & Taube, A. (1997). Interexaminer reliability in physical examination of patients with low back pain. Spine, 22, 814-820.*
- Stuge, B., Veierod, M. B., Laerum, E., & Vollestad, N. (2004). The efficacy of a treatment program focusing on specific stabilizing exercises for pelvic girdle pain after pregnancy: a two-year follow-up of a randomized clinical trial. Spine, 29, E197-E203.*
- Svensson, E. (1984). Forskningsmetode grunnbok for helsearbeidere. Oslo: Aschehoug/Tanum-Norli.*
- Szeto, G. P., Straker, L. M., & O'Sullivan, P. B. (2005). EMG median frequency changes in the neck-shoulder stabilizers of symptomatic office workers when challenged by different physical stressors. Journal of Electromyography & Kinesiology, 15, 544-555.*
- Taylor, N., Evans, O., & Goldie, P. (2001). Reliability of measurement of angular movements of the pelvis and lumbar spine during treadmill walking. Physiotherapy Research International, 6, 205-223.*
- The Quebec Task Force on Spinal Disorders (1987). Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Spine, 12, S1-59.*
- Tropiano, P., Huang, R. C., Girardi, F. P., Cammisa, F. P., Jr., & Marnay, T. (2005). Lumbar total disc replacement. Seven to eleven-year follow-up. Journal of Bone and Joint Surgery, 87, 490-496.*
- Van Dillen, L. R., Sahrman, S. A., Norton, B. J., Caldwell, C. A., Fleming, D., McDonnell, M. K. et al. (2001). Effect of active limb movements on symptoms in patients with low back pain. Journal of Orthopaedic and Sports Physical Therapy, 31, 402-413.*

- Van Dillen, L. R., Sahrman, S. A., Norton, B. J., Caldwell, C. A., Fleming, D. A., McDonnell, M. K. et al. (1998). Reliability of physical examination items used for classification of patients with low back pain. *Physical Therapy*, 78, 979-988.
- Van Dillen, L. R., Sahrman, S. A., Norton, B. J., Caldwell, C. A., McDonnell, M. K., & Bloom, N. J. (2003). Movement system impairment-based categories for low back pain: stage 1 validation. *Journal of Orthopaedic & Sports Physical Therapy*, 33, 126-142.
- Vibe, F. K., O'Sullivan, P. B., Kvale, A., & Skouen, J. S. (2008). Inter-examiner reliability of a classification system for patients with non-specific low back pain. *Man.Ther.*.
- Vollenbroek-Hutten, M. M., Hermens, H. J., Wever, D., Gorter, M., Rinket, J., & Ijzerman, M. J. (2004). Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multi-axial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clinical Rehabilitation*, 18, 566-579.
- Vroomen, P. C., de Krom, M. C., & Knottnerus, J. A. (1999). Diagnostic value of history and physical examination in patients suspected of sciatica due to disc herniation: a systematic review. *Journal of Neurology*, 246, 899-906.
- Vroomen, P. C., de Krom, M. C., & Knottnerus, J. A. (2000). Consistency of history taking and physical examination in patients with suspected lumbar nerve root involvement. *Spine*, 25, 91-96.
- Waddell, G. (1993). How patients react to low back pain. *Acta Orthop.Scand.Suppl*, 251, 21-24.
- Waddell, G., Newton, M., Henderson, I., Somerville, D., & Main, C. J. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, 52, 157-168.
- Waddell, G. (2004). *The back pain revolution. (2nd ed ed.)* Edinburgh: Churchill Livingstone.
- Walker, B. F. (2000). The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *Journal of Spinal Disorders*, 13, 205-217.
- Werneke, M. & Hart, D. L. (2001). Centralization phenomenon as a prognostic factor for chronic low back pain and disability. *Spine*, 26, 758-764.
- Werner, E. L. & Indahl, A. (2005). [Knowledge, practice and attitudes to back pain among doctors, physiotherapists and chiropractors]. *Tidsskr.Nor Laegeforen.*, 125, 1794-1797.

White, L. J. & Thomas, J. S. (2002). *The rater reliability of assessments of symptom provocation in patients with low back pain. Journal of Back and Musculoskeletal Rehabilitation, 16, 83-90.*

World Health Organization (1980). *International classification of impairments, disabilities, and handicaps a manual of classification relating to the consequences of disease : published for trial purpose in accordance with resolution WHA29.35 of the Twenty-ninth World Health Assembly, May 1976. Geneva: WHO.*

World Health Organization (2001). *ICF International classification of functioning, disability and health. Geneva: World Health Organization.*

World Health Organization (2004). *ICD-10: International statistical classification of diseases and related health problems. (2nd ed ed.) Geneva: World Health Organization.*

Vedlegg 1



Ortopedisk Senter
FOU-enheten, NAR

Regional komité for medisinsk forskningsetikk, Øst-Norge (REK Øst)
Postboks 1130
Blindern, 0318 Oslo

Oslo, 27.04.04

Ang. protokolltilførsel i prosjektet "Sammenligning av skiveprotese og ikke-operativ behandling ved kroniske korsryggsmerter med degenerative skiveforandringer".

Vi viser til prosjektvurdering foretatt av REK 27.01.04 og Deres svarbrev datert 30.01.04 (ref. 43-04013). Vi takker for positiv vurdering av hovedprosjektet og på tilleggsøknad / orientering datert 17.02.04 angående helseøkonomisk vurdering og bruk av spørreskjemaet SF-36. Protokoll for hovedstudien følger vedlagt (vedlegg 1).

Denne norske multisenterstudien innebærer tilgang på et unikt pasientmateriale, noe som åpner for mange interessante og viktige problemstillinger. Pasientene som inkluderes i studien inkluderes i stor grad på grunnlag av røntgenologiske funn. Litteraturen beskriver at det per i dag ikke foreligger kliniske tester som kan identifisere pasientgruppen. En australsk forskergruppe har imidlertid lansert en klinisk test som de mener blant annet kan identifisere denne type pasienter. Testen brukes av fysioterapeuter, men den er ikke tidligere utprøvd på denne pasientgruppen. To fysioterapeuter knyttet til prosjektet "Sammenligning av skiveprotese og ikke-operativ behandling ved kroniske korsryggsmerter med degenerative skiveforandringer" ønsker å benytte denne testen på pasientene som inkluderes i studien for å se på følgende (fullstendig protokoll for metodestudien følger vedlagt (vedlegg 2)):

1. Gjøre en reliabilitetstest av et utvalg av pasientene
2. Undersøke testens validitet og prediktive verdi

Et annet moment vi også ønsker å studere i dette pasientmaterialet er seksualfunksjon og vannlatingsproblemer. Årsaken til dette er en hypotese i ryggforskningssmiljøer om at langvarige korsryggsmerter og / eller ryggkirurgi kan medføre endringer i seksualfunksjon og/eller vannlating. Litteraturen beskriver hyppig endringer i seksualliv som følge av smertetilstanden i seg selv eller som følge av eksempelvis depresjon og somatisering oppstått i kjølevannet av langvarige smerter. Så vidt vi har kunnet bringe på det rene har imidlertid ingen per i dag publisert data på mulige fysiologiske / nevromuskulære endringer som følge av langvarige korsryggsmerter eller som følge av kirurgisk inngrep i ryggen, noe vi anser som viktig, ikke minst med tanke på å oppdage mulige bivirkninger av nye behandlingsmetoder. Vedlagt følger derfor også et spørreskjema angående dette tema som vi ønsker å forelegge komiteen (vedlegg 3). Det

Ullevål universitetssykehus HF
0407 OSLO

Telefon:
22 11 74 64

Besøksadresse:
Kirkeveien 166

Bankgiro:
1644 06 05897

Foretaksnr.:
983 971 784

Telefaks:
23 01 52 70

HELSE ØST

er utarbeidet etter råd fra et svensk ryggforskingsmiljø som har samlet noen foreløpige data på dette tema men som ennå ikke er publisert (O Hägg ved Sahlgrenska og P Fritzell i Faldun). Spørsmål vedrørende eventuelle forstyrrelser i vannlating er satt opp i samarbeid med professor / fysioterapeut Kari Bø ved Norges idrettshøgskole og baserer seg på anbefalinger fra the International Continence Society. Selv om de fleste spørsmål ofte inngår i en vanlig klinisk undersøkelse og anamneseopptak er vi klar over at enkelte pasienter kan føle det belastende å svare på denne type spørsmål. For å gjøre situasjonen lettere for pasientene vil vi derfor trykke spørsmålene på et eget ark og forskningssykepleier tilknyttet prosjektet vil personlig forelegge dem pasientene.

Vi håper på positiv vurdering og beklager eventuelt ekstra bryderi vår tilleggsøknad måtte innebære.

På vegne av prosjektledelsen
Oliver Grundnes

Kjersti Storheim

Dr.med
Ortopedisk Senter
Ullevål Universitetssykehus
Kirkeveien 166
0407 Oslo

Fysioterapeut / dr.scient
Kmpetanses. for klin.forsk
Ullevål Universitetssykehus
Kirkeveien 166
0407 Oslo

Vedlegg 1- Protokoll for hovedstudien
 2- Protokoll for metodestudie
 3- Spørreskjema for kartlegging av seksualfunksjon og
 vannlatningsproblemer

REGIONAL KOMITE FOR MEDISINSK FORSKNINGSETIKK
Øst-Norge (REK I)

Seksjonsoverlege Oliver Grundnes
Ortopedisk senter
Ullevål universitetssykehus

Deres ref.:

Vår ref.: 276-04013

Dato: 10. mai 2004

**Sammenligning av skiveprotese og ikke-operativ behandling ved kroniske
korsryggmerter med degenerative skiveforandringer**
Tillegg til protokollen

Vi viser til Deres brev av 27.04.04, der det redegjøres for enkelte tilleggsundersøkelser som ønskes gjennomført.

Komiteen har ingen innvendinger mot at tilleggsundersøkelsene gjennomføres. Den har imidlertid merket seg at den ikke har fått svar på sitt brev av 10.03.04, der den ber om en nærmere redegjørelse for opplegget med bueleddsblokkade på et utvalg av pasientene.

Med vennlig hilsen



Ida Nyquist
sekretær

Vedlegg 2

Datatilsynet

Ullevål universitetssykehus HF
v/ Dr. med. Oliver Grundnes
Kirkevn. 1666
0407 OSLO

Deres ref

Vår ref (bes oppgitt ved svar)
2004/1079-2 MOF/-

Dato
29.09.2004

KONSESJON TIL Å BEHANDLE HELSEOPPLYSNINGER

Datatilsynet viser til Deres søknad av 02.07.2004 om konsesjon til å behandle helseopplysninger.

Datatilsynet har vurdert søknaden og gir Dem med hjemmel i helseregisterloven § 5, jf. personopplysningsloven § 33, jf. § 34, konsesjon til å behandle helseopplysninger i forbindelse med prosjektet: *"Effekt av skiveprotese versus konservativ behandling ved kroniske korsryggsmarter med degenerative skiveforandringer: en prospektiv randomisert multisenterstudie"*.

Databehandlingsansvarlig er Ullevål universitetssykehus ved øverste leder. Gjennomføringen av det daglige ansvaret kan delegeres.

Konsesjonen er gitt under forutsetning av at behandlingen foretas i henhold til søknaden og de bestemmelser som følger av helseregisterloven med forskrifter.

Dersom det skjer endringer i behandlingen i forhold til de opplysninger som er gitt i søknaden, må dette fremmes i ny konsesjonssøknad.

I medhold av helseregisterloven § 5, jf. § 36, jf. personopplysningsloven § 35, fastsettes i tillegg følgende vilkår for behandlingen:

1. Den databehandlingsansvarlige skal hvert tredje år sende Datatilsynet bekreftelse på at behandlingen skjer i overensstemmelse med søknaden og helseregisterlovens regler.
2. Den databehandlingsansvarlige må slette / anonymisere opplysninger som direkte eller indirekte kan identifisere enkeltpersoner, når de ikke lenger er nødvendig for formålet med behandlingen, senest innen 2015.

Postadresse:
Postboks 8177 Dep
0034 OSLO

Kontoradresse:
Tollbugt 3

Telefon:
22 39 69 00


Telefaks:
22 42 23 50


Org.nr:
974 761 467

Hjemmeside:
www.datatilsynet.no

Datatilsynet tar forbehold om at konsesjonen kan bli trukket tilbake eller at nye og endrede vilkår kan bli gitt dersom dette er nødvendig ut fra personvern hensyn.

Med hilsen


Hanne P. Gulbrandsen (e f)
rådgiver


Monica Fornes
rådgiver
(saksbehandler,
telefon 22 39 69 00)

kopi: Ullevål Universitetssykehus HF
v/ IKT-sikkerhetssjef
Heidi Thorsteinsen
Konsern IT-Administrasjonen
0407 OSLO

Søknad om konsesjon for behandling av personopplysninger i henhold til personopplysningsloven (pol.) § 33

Vennligst ikke heft eller stift blanketten sammen

Sendes til: Datatilsynet
Postboks 8177 Dep
0034 OSLO

E-postadresse: postkasse@datatilsynet.no

Telefaksnummer: 22 42 23 50

Hjemmeside: <http://www.datatilsynet.no>

Telefonnummer: 22 39 69 00

Informasjon om søkeren.			
A. Behandlingsansvarlig virksomhet.			
Fullstendig navn/firma (fylles alltid ut)			Organisasjonsnummer
Ullevål universitetssykehus HF			9 8 3 9 7 1 7 8 4
Postadresse	Postnr.	Sted	Land
Kirkeveien 166	0407	Oslo	Norge
Telefonnummer	Telefaksnummer	E-postadresse/hjemmeside	
22118080	22119950	www.uus.no	
B. Behandlingsansvarliges representant			
Fullstendig navn/firma			Organisasjonsnummer
Postadresse	Postnr.	Sted	Land
Telefonnummer	Telefaksnummer	E-postadresse/hjemmeside	
C. Daglig ansvar for oppfyllelse av den behandlingsansvarliges plikter er tillagt:			
Stillingsangivelse (fylles alltid ut)			
Administrerende direktør, kontaktperson: Oliver Grundnes, dr.med, Ortopedisk Senter.			
D. Benyttes databehandler? Nei <input type="checkbox"/> Ja <input checked="" type="checkbox"/> Hvis ja, legg ved avtale			
Fullstendig navn/firma på databehandleren			Organisasjonsnummer
Postadresse	Postnr.	Sted	Land
Telefonnummer	Telefaksnummer	E-postadresse/hjemmeside	
E. Gi en beskrivelse av hva slags virksomhet som behandlingsansvarlige driver i vid forstand.			
<p>UUS som behandlingsansvarlig, er et universitetssykehus. Studien er videre en multisenterstudie med følgende andre aktører: Universitetssykehuset i Nord Norge, Tromsø, St.Olavs hospital, Trondheim, Haukeland Universitetssykehus (avdelingene Kysthospitalet i Hagavik og Nakke-ryggklinikken på Nesttun), Bergen</p>			
F. I hvilken del av virksomheten skal behandlingen foretas?			
<p>For UUS: Ortopedisk senter og fysikalskmedisinsk avdeling, For UNN: Tverrfaglig ryggpoliklinikk, For St.Olavs hospital: Nasjonalt senter for spinale lidelser og nevrokirurgisk avdeling, For Haukeland sykehus: Kysthospitalet i Hagavik og Nakke/ryggpoliklinikken på Nesttun.</p>			

1 Om behandlingen.	
1.1 Behandlingen omfatter følgende typer personopplysninger:	
1.1.1 Ikke-sensitive personopplysninger <i>Identifikasjonsopplysninger</i> <input type="checkbox"/> Navn, adresse, fødselsdato <input type="checkbox"/> Fødselsnummer (11 siffer) <input type="checkbox"/> Fingeravtrykk, iris <input checked="" type="checkbox"/> Annet; elektronisk lagres opplysningene aidentifisert <i>Atferdsopplysninger</i> <input type="checkbox"/> Loggføring av adferd, herunder tidspunkt og sted for handlinger <input type="checkbox"/> Preferanser (ønsker, behov o.l.) <input type="checkbox"/> Detaljopplysninger tilknyttet omsetning av varer eller tjenester <input type="checkbox"/> Annet <i>Opplysninger om tredjepersoner</i> <input type="checkbox"/> Navn, adresse, fødselsdato <input type="checkbox"/> Fødselsnummer (11 siffer) <input type="checkbox"/> Annet <i>Fjernsynsovervåkning og lydopptak (sett eventuelt. flere kryss)</i> <input type="checkbox"/> Fjernsynsovervåkning <input checked="" type="checkbox"/> Billedopptak; videoopptak, ikke identifiserbart <input type="checkbox"/> Lydopptak <i>Opplysninger i forbindelse med</i> <input type="checkbox"/> Bankvirksomhet <input type="checkbox"/> Forsikringsvirksomhet <input type="checkbox"/> Telekommunikasjonsvirksomhet	1.1.2 Sensitive personopplysninger, jf. pol. § 2 nr. 8 <i>Behandlingen omfatter opplysninger om</i> <input type="checkbox"/> rasemessig eller etnisk bakgrunn, eller politisk, filosofisk eller religiøs oppfatning <input type="checkbox"/> at en person har vært mistenkt, siktet, tiltalt eller dømt for en straffbar handling <input checked="" type="checkbox"/> helseforhold <input type="checkbox"/> seksuelle forhold <input type="checkbox"/> medlemskap i fagforeninger. Presiser nærmere: (eventuelt i eget vedlegg)
	<div style="border: 1px solid black; padding: 5px;"> Pasientene lagres med nummer elektronisk og vil ha en manuell kryssreferanseliste som låses inn lokalt på hvert sykehus Data innhentes ved spørreskjema, klinisk undersøkelse og videoopptak. Se vedlagte protokoll for hovedstudien og tilleggsprotokoll for metodestudie som innebærer videoopptak. </div>
	<i>Behandling av sensitive personopplysninger:</i> <input type="checkbox"/> som skjer helt eller delvis skjer med elektroniske hjelpemidler <input type="checkbox"/> som inngår eller skal inngå i et manuelt personregister
1.2 Behandlingen omfatter opplysninger om:	
<input type="checkbox"/> Ansatte i egen virksomhet <input type="checkbox"/> Adgangskontrollerte <input type="checkbox"/> Annet: (skriv inn i feltet nedenfor)	<input type="checkbox"/> Elever/student/barnehagebarn <input type="checkbox"/> Medlemmer <input checked="" type="checkbox"/> Pasienter <input type="checkbox"/> Kunder/klienter/brukere <input type="checkbox"/> Tilfeldig utvalgte <input type="checkbox"/> Seleksjonsutvalg
<div style="border: 1px solid black; height: 40px; width: 100%;"></div>	
1.3 Hva er formålet med behandlingen?	
<p>Formålet med studien er å sammenligne effekt av innsetting av skiveprotese og ikke-operativ behandling hos pasienter med kroniske korsryggsmerter forårsaket av degenerative skiveforandringer. Studien er en nasjonal multisenterstudie og inkluderer Universitetssykehusene i Tromsø, Trondheim, Bergen og Oslo (Ullevål). Innsetting av skiveprotese er en forholdsvis ny behandlingsmetode overfor pasientgruppen. Behandlingen er imidlertid ikke tidligere sammenlignet med et strukturert konservativt behandlingsopplegg e. Dette anses både nasjonalt og internasjonalt som meget viktig før konklusjoner rundt behandlingseffekt av skiveprotese kan trekkes. Samtlige effektmål er validert for bruk på kroniske ryggpasienter og anbefales av internasjonale ekspertpanel brukt i forskningsprosjekter på målgruppen. Alle personer som er ansvarlig for datainnsamling er erfare forskere og er ansatt i forskningsmiljøer direkte underlagt de nevnte universitetssykehus.</p> <p>Videoopptak (se egen protokoll, gjelder bare Ullevål) skal innhentes for å validere og reliabilitetsteste en nyutviklet klinisk test benyttet av fysioterapeuter. Testen er utviklet av Australske forskere og det anses som viktig å videreutvikle og kvalitetssikre denne testen.</p>	
1.4 Hvordan samles personopplysningene inn?	
<input checked="" type="checkbox"/> Manuelt <input type="checkbox"/> Elektronisk <input checked="" type="checkbox"/> Videoopptak <input type="checkbox"/> Lydopptak	<input type="checkbox"/> Annet: (skriv inn i feltet nedenfor) <div style="border: 1px solid black; padding: 5px;"> Manuelt: spørreskjema og registreringsskjema (se vedlegg til protokoll) Videoopptak (gjelder bare Ullevål): opptak av korsryggen skjer bakfra og fra siden. Under opptak bakfra kan pasienten ikke identifiseres, ved opptak fra siden vil pasienten bære operasjonshette og laserbriller for å unngå identifisering (se egen protokoll) </div>

1.5 Hvordan skal opplysningene brukes?		
<input checked="" type="checkbox"/> Registreres <input checked="" type="checkbox"/> Lagres <input checked="" type="checkbox"/> Sammenstilles, med hva:	<p>I denne nasjonale multisenterstudien skal det etableres en database ved hvert senter som deltar i studien. Det vil skje en utveksling av data fra de ulike senter til Trondheim og til Oslo (stipendiater knyttet til studien er lokalisert her). Utgangspunktet for databasen som bygges i studien er den nasjonale ryggdatabasen som er lagt til Universitetssykehuset i Tromsø. Det foreligger egen godkjenning fra Datatilsynet for denne nasjonale ryggdatabasen.</p>	
<input type="checkbox"/> Utleveres, til hvem:	<p>Data innhentet på enkelte pasienter ved hvert sykehus vil bli utlevert til databasen som multisenterstudien skal etablere i Trondheim og i Oslo. Utleveringen skjer i form av aidentifiserte opplysninger, da hvert sykehus selv sitter på koblingsnøkkel for egne pasienter. Opplysningene overføres til Trondheim og Oslo på CD i rekommandert post.</p>	
<input checked="" type="checkbox"/> Annet: (skriv inn i feltet)	<p>Pasientene vil lagres med nummer elektronisk og vil ha en kryssreferanseliste manuelt som låses inn. Alle data som utveksles mellom databasene på det enkelte sykehus og til Trondheim og Oslo vil være aidentifiserte, da ingen andre enn prosjektleder og prosjektmedarbeidere på det enkelte sykehus vil ha tilgang til den innlåste kryssreferanselisten. Ansvarlig for databasen i Trondheim er Øystein Nygaard, dr.med, leder for Nasjonalt Senter for Spinale Lidelser.</p>	
1.6 Overføring av personopplysninger til utlandet.		
Overføres opplysningene til utlandet? <input type="checkbox"/> Nei		
<input checked="" type="checkbox"/> Ja Mottakers navn og adresse	<p>Peter Fritzell, overlege ved ortopedisk klinikk Falu lasarett, 791 82 Falun. Opplysningene som overføres er aidentifiserte. Opplysningene overføres til Falun / dr.Fritzell på CD i rekommandert post.</p>	
1.7 Personopplysningene innhentes fra:		
<input type="checkbox"/> den registrerte selv <input checked="" type="checkbox"/> andre enn den registrerte: (skriv inn fra hvem i feltet nedenfor)	<p>Vi benytter resultater fra discografi og bueleddsblokade for verifisering av diagnosen. Disse opplysningene kommer fra lege. I den kliniske testen utført av fysioterapeuter (videoopptak) vil fysioterapeuten foreta en klassifisering av pasientens symptomer.</p>	
1.8 Rettslig grunnlag for behandling av personopplysninger.		
1.8.1 Behandling av personopplysninger i henhold til personopplysningsloven § 8 <input checked="" type="checkbox"/> med den registrertes samtykke <input type="checkbox"/> ved at det er fastsatt i lov at det er adgang til slik behandling <input type="checkbox"/> for å oppfylle en avtale med den registrerte, eller utføre gjøremål etter den registrertes ønske før en slik avtale skal inngås <input type="checkbox"/> for at den behandlingsansvarlige skal kunne oppfylle en rettslig forpliktelse <input type="checkbox"/> for å utføre en oppgave av allmenn interesse <input type="checkbox"/> for å utøve offentlig myndighet <input type="checkbox"/> for at den behandlingsansvarlige eller tredjepersoner som opplysningene utleveres til kan vareta en berettiget interesse, og hensynet til den registrertes personvern ikke overstiger denne interessen	1.8.2 Behandling av personopplysninger i henhold til personopplysningsloven § 9 <input checked="" type="checkbox"/> med den registrertes samtykke <input type="checkbox"/> når det er fastsatt i lov at det er adgang til slik behandling <input type="checkbox"/> når behandlingen er nødvendig for å beskytte en persons vitale interesser, og den registrerte ikke er i stand til å samtykke <input type="checkbox"/> når det utelukkende behandles opplysninger som den registrerte selv frivillig har gjort alminnelig kjent <input type="checkbox"/> når behandlingen er nødvendig for å fastsette, gjøre gjeldende eller forsvare et rettskrav <input type="checkbox"/> når behandlingen er nødvendig for at den behandlingsansvarlige kan gjennomføre sine arbeidsrettslige plikter eller rettigheter <input type="checkbox"/> når behandlingen er nødvendig for forebyggende sykdomsbehandling, medisinsk diagnose, sykepleie eller pasientbehandling eller for forvaltning av helsetjenester, og opplysningene behandles av helsepersonell med taushetsplikt <input type="checkbox"/> når behandlingen er nødvendig for historiske, statistiske eller vitenskapelige formål, og samfunnets interesse i at behandlingen finner sted klart overstiger ulempene den kan medføre for den enkelte <input type="checkbox"/> i henhold til personopplysningsloven § 9 annet ledd	
1.9 Hvordan skal informasjonsplikten i pol. §§ 19 og 20 oppfylles?		
Når opplysninger samles inn fra den registrerte	<input checked="" type="checkbox"/> Muntlig <input checked="" type="checkbox"/> Skriftlig	Hvis skriftlig, legg ved kopi av dokumentasjonen,
Når opplysninger samles inn fra andre enn den registrerte	<input type="checkbox"/> Muntlig <input type="checkbox"/> Skriftlig	
1.10 Har virksomheten utarbeidet et interkontrollsystem? Nei <input type="checkbox"/> Ja <input checked="" type="checkbox"/>		

2 Om oppfyllelse enkelte bestemmelser i personopplysningsloven.	
2.1	Hvordan skal det sikres at opplysningene er tilstrekkelige og relevante i forhold til formålet med behandlingen? Legg ved en beskrivelse av rutinen. (Se nederst på skjema)
2.2	Hvordan skal det sikres at opplysningene er korrekte og oppdaterte i forhold til formålet med behandlingen. Legg ved en beskrivelse av rutinen. (Se nederst på skjema)
2.3	Hvordan skal det sikres at opplysningene slettes når formålet er oppfylt. Legg ved en beskrivelse av sletterutinen.
2.4	Er det meningen at opplysningene skal oppbevares etter at formålet er oppfylt? Nei <input type="checkbox"/> Ja x <input type="checkbox"/> Hvis ja; angi grunnlag:
	<input type="checkbox"/> i medhold av arkivloven
	<input type="checkbox"/> annen lovgivning, hvilken lov hjemmel _____
	<input type="checkbox"/> historiske eller statistiske formål
	X <input type="checkbox"/> vitenskapelige formål
	Dersom opplysningene behandles for historiske, statistiske eller vitenskapelige formål, legg ved en beskrivelse av de samfunnsinteresser som vil oppveie personvernulempene.
	Presiser nærmere: (eventuelt i eget vedlegg)
	Den aktuelle studien har 5 års oppfølging og planlegges avsluttet i 2015. Dersom ikke annet søkes innen da, anonymiseres opplysningene ved at kodelistene slettes. Det vil i henhold til Datatilsynets krav bli sendt melding hvert 3. år. Videre bruk av data eller sletting søkes å endelig avgjøres innen utgangen av studien det her søkes konsesjon for. Dette begrunnes ut fra at man først ved avslutning av denne studien vil ha grunnlag for å mene om innsamlede data er av type kvalitet og kvantitet som gjør det interessant å benytte det i videre studier. Nytt samtykke vil innhentes før slik videre bruk, noe som også er opplyst om i samtykkeerklæringen. Ved eventuelle nye problemstillinger vil ny konsesjon bli innhentet.
3 Informasjonssikkerhet	
3.1	Hvordan og til hva skal informasjonsteknologi brukes? Legg ved sikkerhetsmål fra virksomheten. VEDLEGG 1
3.2	Hva er de overordnede valg i sikkerhetsarbeidet? Legg ved sikkerhetsstrategi. VEDLEGG 1
3.3	Hvordan har behandlingsansvarlige klarlagt sannsynligheten for og konsekvensen av eventuelt sikkerhetsbrudd? Legg ved de krav som virksomheten stiller til tilgjengelighet, integritet og konfidensialitet for personopplysningene. Legg også ved resultatene fra risikovurderingen. Gjennomføring av risikovurdering er benyttet for å vurdere sannsynlighet og konsekvenser av eventuelt sikkerhetsbrudd. Se VEDLEGG 2
3.4	Legg ved organisasjonskart med ansvars og myndighetsforhold i forhold til sikkerhet, IT-drift og behandlingen. VEDLEGG 1
3.5	Legg ved oversikt over informasjonssystemets konfigurering. Dette omfatter hvordan utstyr og program er sammenkoplet med eventuelt bruk av sikkerhetsbarrierer og ulike nettverk. VEDLEGG 3
3.6	Er nettet som inneholder personopplysninger tilknyttet et eksternt nett som for eksempel Internett? Ja <input checked="" type="checkbox"/> Nei <input type="checkbox"/>
3.7	Overføres personopplysningene i eksternt datanett? Ja <input type="checkbox"/> Nei <input checked="" type="checkbox"/>
3.8	Utføres behandlingen i henhold til annet regelverk som regulerer sikkerheten. Hvis ja; hvilket regelverk: _____
Underskrift	
Sted og dato	Underskrift
02.07.04	

2.1 og 2.2: Det henvises til svar på vurdering av etisk komité og til protokollen for studien. Samtlige benyttede målemetoder er validert for bruk på kroniske ryggpasienter og anbefales av internasjonale ekspertpanel brukt i forskningsprosjekter på målgruppen.

Vedlegg 3

Rand.nr.....

SAMTYKKE ERKLÆRING

Ullevål universitetssykehus ved Avdeling for Fysikalsk Medisin og Rehabilitering og Ortopedisk senter har som en av sine oppgaver å drive forskningsarbeide innefor ryggbehandling og - rehabilitering. Du er nå inkludert i en studie som skal sammenligne skiveprotese versus ikke kirurgisk ryggrehabilitering. Som en del av utredningen, før behandlingen starter opp, skal du til bevegelsesanalyse hos fysioterapeut. I analysen inngår intervju og video opptak.

Det samtykkes med dette til at fysioterapeuter ved Avdelning for Fys. Med og Rehab. gjør videoopptak av(navn på pasient) med henblikk på bruk til følgende formål:

- Etterfølgende analyse av videofilmet ryggfunksjon som en del av ovennevnte forsknings studie.
- Intern undervisning (UUS) av helsepersonell.

Undertegnede aksepterer at aktuelle produksjon og etterfølgende nevnte formidling skal kunne skje uten krav på kompensasjon eller annen form for vederlag.

Pasientens signatur

Dato (Pasienten skriver selv dato)

Som prosjekt tilknyttet fysioterapeut bekrefter jeg at pasienten har fått informasjon om funksjonsanalysen ved videoopptak og at denne blir brukt bare til ovennevnte foremål.

Fysioterapeutens signatur

Vedlegg 4

SAMMENLIGNING AV OPERASJON MED SKIVEPROTESE OG TVERRFAGLIG RYGGREHABILITERING VED LANGVARIGE KORSRYGGSMERTER MED DEGENERATIVE SKIVEFORANDRINGER

En prospektiv randomisert multisenter studie

Om lag 70 – 85 % av den voksne befolkning har ryggsmarter i løpet av sitt voksne liv og pasienter med langvarige ryggsmarter, dvs. smerter i mer enn 3 mnd, har et høyere forbruk av helsetjenester enn de fleste andre pasientgrupper (Hart et al 1995, Von Korff et al. 1991). Det naturlige forløp hos disse pasientene er preget av langsom og ofte usikker spontan bedring. Etter at ryggpasienter har vært ute av arbeid i to år, er det tilnærmet ingen som kommer tilbake i aktivt inntektsgivende arbeid (Hagen et al 2000). Flere av de konservative behandlingsformene har i randomiserte studier vist effekt. Dette gjelder både konvensjonell treningsbehandling og ved bruk av kognitiv adferdsterapeutiske prinsipp (van Tulder 1997, Furlan 2001).

Spørsmålet om operativ behandling melder seg når pasientene opplever vedvarende smerter tross gjennomført konservativ behandling. Den kirurgiske behandlingen er imidlertid kontroversiell. En rekke forfattere har pekt på at ryggsmarter har multifaktorielle årsaker og at tradisjonell kirurgisk behandling ikke er bedre enn det naturlige forløp. Mange nyere studier, som hovedsakelig har vært retrospektive med heterogene pasientmaterialer uten kontrollgrupper, har også vært inkonsistente. Den svenske ryggstudien viste i en randomisert sammenlignende studie mellom konservativ og operativ behandling at pasienter som ble operert med en avstivingsoperasjon hadde mindre smerter og bedre funksjon etter 1 og 2 år (Fritzell et al. 2001). I en norsk studie hvor et systematisk konservativt opplegg med trening og kognitiv

behandling ble sammenlignet med avstiving har ikke vist noen forskjell mellom de to gruppene (Brox et al. 2003).

Hos pasienter med langvarige korsryggsmerter og lokaliserte degenerative røntgenforandringer kan man de siste 10 – 15 årene ved enkelte ryggklinikker satt inn proteser som erstatning for mellomvirvelskiven (Cinotti et al 1996) . De foreløpige resultatene fra retrospektive ikke randomiserte studier uten avhengig observatør viser en suksessrate på 70 – 90 %, altså om lag det samme som avstivningsoperasjoner i denne typen studiedesign (Greenough and Peterson 1998, Bartagnoli et al 2002). Det foregår randomiserte studier i Europa og USA mellom skiveprotese og avstivning hos denne pasientkategorien.

Skiveprotese har en del teoretiske fordeler framfor tradisjonell avstivning:

- Ingen tilheling i ben eller skiver, kun sårtilheling.
- Lite kirurgisk traume på muskulatur og annet støtteapparat i ryggen
- Ingen øket belastning på naboskiver som følge av avstivning i et eller flere nivåer.

Skiveprotese har også noen teoretiske ulemper:

- Implantat slitasje med dannelse av wear products
- Vanskelig revisjonskirurgi ved implantat svikt
- Teknisk krevende kirurgi
- Langtidsresultater ikke kjent
- Potensielt alvorlige komplikasjoner

Til tross for at flere former for skiveproteser har vært på markedet i mange år og flere tusen pasienter har mottatt denne behandlingen foreligger det ingen studier hvor den operative behandlingen med skiveproteser sammenlignes med en ikke-operativ behandling.

Målet med denne studien er å sammenligne effekten av et strukturert tverrfaglig rehabiliteringsopplegg mot operasjon med skiveprotese hos

pasienter som har lokaliserte degenerative forandringer og smerter i korsryggen.

Hypotese:

Resultatene ved lumbale skiveproteser målt med smerte- og funksjonsscore er lik ved ikke operativ behandling etter 1 og 2 år

Tilleggshypotese:

Pasienter som opereres med skiveprotese får en raskere bedring og kommer raskere tilbake i arbeid enn pasienter som behandles ikke-operativt

Studievariabler:

Hovedeffektvariabel:

Funksjon (Oswestry disability index)

Tilleggsvariabler:

Arbeidsstatus, smerte, livskvalitet og fear-avoidance.

MATERIALE OG METODE

Studien er en multisenterstudie som foregår ved sykehusene:

- Universitetssykehuset i Nord Norge, Tromsø
- St Olav Hospital, Trondheim
- Kysthospitalet i Hagavik/Nakke-ryggklinikken på Nesttun, Bergen
- Ullevål Universitetssykehus, Oslo

Studiedesign

Prospektiv randomisert enkeltblindet studie. Pasientene randomiseres ved blokk randomisering til. Dette innebærer at det randomiseres for hvert studiested.

Den operative behandlingen består i innsetting av en skiveprotese i et av nivåene L4 – L5 eller L5 – S1. Det benyttes ingen ortose postoperativt. Den ikke-kirurgiske behandlingen vil foregå ved de respektive sykehusene etter en nærmere definert protokoll (vedlegg bakerst i protokollen).

Ved avstivningsoperasjon i ryggen har en rekke tidligere studier vist at pasienter som er tidligere ryggoperert har en dårligere prognose. Det er derfor viktig å skille disse pasientene ved inklusjon. Man har derfor besluttet å lage to separate grupper, en gruppe for tidligere ikke opererte i studie A og en for tidligere opererte i studie B. Behandlingen i de to studiene vil ellers være lik.

Inklusjonskriterier

- Alder 25 – 55 år
- Korsryggsmerter som hovedsymptom med varighet i mer enn 1 år.
- Gjennomgått strukturert fysioterapi eller kiropraktorbehandling
- Oswestry score lik eller over 30 % poeng
- Degenerative forandringer i mellomvirvelskiven i en av de to nederste nivåer. Degenerasjon bedømmes på MR etter følgende kriterier:
 - Høydereduksjon av mellomvirvelskiven med minst 40 %.
 - Modic forandringer type I og/eller II.
 - High Intensity Zone i skiven.
 - Morfologiske forandringer vurdert som endret signaltetthet i skiven

For at man skal kunne klassifisere skiven som degenerativ må enten det første kriteriet alene eller minst to av de andre kriterier være til stede. MR bildene vurderes av to uavhengige observatører.

Eksklusjonskriterier

- Kronisk generalisert smertesyndrom

- Degenerasjon påvist ved bildediagnostikk i mer enn et nivå. For at et skivenivå skal klassifiseres som normalt må ikke skivehøyden være redusert med mer enn 40 % og ikke ha mer enn et av de øvrige kriterier for degenerativ skivepatologi. Bedømmingen av en normal skive skjer ved to uavhengige observatører. Ved uenighet bedømmes bildene av en tredje observatør og utfallet avgjøres ved simpelt flertall.
- Symptomgivende spinal stenose
- Skiveprolaps eller rescesstenose med nerverotsaffeksjon
- Spondylolyse
- Spondylolisthese, isthmisk
- Artritt
- Tidligere fraktur L1 – S1
- Psykisk eller somatisk sykdom som gjør at pasienten ikke egner seg for et av behandlingsalternativene
- Forstår ikke norsk skriftlig og muntlig
- Medikamentmisbruker
- Osteoporose
- Medfødt eller ervervet deformitet

For studie A gjelder i tillegg følgende eksklusjonskriterie:

- Tidligere ryggoperert

For studie B gjelder i tillegg følgende eksklusjonskriterier:

- Tidligere laminectomert
- Tidligere operert for prolaps og to residiver, dvs. ikke mer enn tre tidligere ryggoperasjoner

Utvelgelse/inkludering

Pasientene velges ut ved de enkelte sykehusene. Andre utrednings og behandlingssentra kan tilsluttes i studien etter oppstart dersom dette skjer i tråd med de behandlingsprinsipper som gjelder for denne protokollen.

Dersom en pasient kan inkluderes skal pasienten bli forespurt om medvirkning i studien. Informert samtykkeerklæring skal undertegnes

samtidig av pasient og den lege som inkluderer pasienten. Pasienten er ikke inkludert i studien før samtykkeerklæring er undertegnet.

Pasienter som ikke ønsker å delta i studien får tilbud om behandling i samsvar med dagens praksis.

Klinisk undersøkelse

Demografiske og sosiale variabler; kjønn, alder, arbeidsstatus, sosial status, røyking.

Nevrologisk undersøkelse i underekstremitetene

- Sensibilitet
- Motorikk
- Reflekser
- Lasegue

Røntgenologiske prosedyrer

- Skjelettrøntgen stående front og side
- Funksjonsrøntgen med maksimal bøy og strekk med pasienten i sittende posisjon
- MR lumbalcolumna

Spørreskjema

Utdanning og yrke, trygdestatus, smertetegning (Öhlund et al 1996), medikamentforbruk de siste 2 uker, Oswestry (Fairbank et al 1980), egen vurdering av arbeidsevne, tid ute av arbeid, self-efficacy for pain sub-scale (Lorig et al 1989), fear avoidance beliefs (Waddell et al 1993) , Hopkins symptoms check list (Derogatis et al 1974), tidligere behandling, EQ 5D (EuroQol 1990), seksual- og vannlatingsfunksjon (seksualfunksjon: ref. O Hägg, vannlating: Abrams et al 2002).

Kostnadseffektivitet

Det er etablert samarbeid med Peter Fritzell og medarbeidere i Falun, Sverige, for å gjøre en fullstendig økonomisk analyse av kostnader. Dette omfatter i tillegg til selve behandlingskostnadene også utgifter til blant annet sykmelding, trygdeytelser, tilleggsbehandling og pasienter og pårørendes utgifter (Fritzell et al 2004). Data innhentes fra trygdekontor, pasienten selv og økonomiansvarlig ved de enkelte sykehus.

Registreringsskjema for pasienter som opereres i ryggen

For pasientene som randomiseres til operasjon fylles i tillegg ut registreringsskjema (pre- og postoperativt) utarbeidet for bruk i ryggdatabasen ved Universitetssykehuset i Nord-Norge (vedlagt).

BEHANDLINGSPROSEDYRER

Operasjon

Skiveprotesen settes inn via en retroperitoneal tilgang. Kirurgen velger om det benyttes et lavt tverrsnitt (Phannensteal) eller paramedialt langsgående snitt i huden. Det benyttes gjennomlysning under prosedyren for å sikre at protesen settes i midtlinjen og tilstrekkelig baktil i skiven. Standard innsetningsutstyr fra produsent. Pasienten har restriksjoner mot tunge løft og eksessive bevegelser de første 6 ukene. Det benyttes samme protese ved alle behandlingssteder.

Ikke-kirurgisk behandling

Behandlingen tar utgangspunkt i intervensjonen til Brox et al., en tverrfaglig intervensjon som kombinerer veiledet fysisk trening og kognitiv tilnærming. (Se vedlagte protokoll for ikke-kirurgisk behandling)

Evaluering

Kontroll ved blindet unbiased lege 1 år og 2 år etter behandling.

Hovedeffektvariabel:

- Funksjon og smerte målt ved Oswestry

Andre effektvariable:

- Smerte (VAS skala og smertetegning)
- Arbeidsstatus
- Livskvalitet (EQ 5D)
- Fear-avoidance (FABQ)
- Følelser og symptomer (HSCL-25)
- Smertemestring (self-efficacy for pain)

Evalueringstidspunkt

Alle effektvariabler registreres ved behandlingsstart og etter 6 uker, 3, 6, og 12 måneder og 2 år. Det planlegges en spørreundersøkelse etter 5 år. I tillegg til spørreskjema gjøres røntgenundersøkelse.

Registrering av komplikasjoner

Komplikasjoner i forbindelse med operasjon og kontrollene registreres.

Endring av gruppe

Pasienter som av ulike grunner ikke fullfører behandlingen i henhold til randomiseringen undersøkes etter 1 og 2 år. Resultatet fra pasienter som endrer gruppe under studien skal inkluderes i de statistiske analyser etter "intention-to-treat" prinsippet, det vil si deres resultat legges i den gruppen de ble randomisert til og ikke den gruppen de ble behandlet i.

Drop-outs

Pasienter som har blitt randomisert til studien men som av ulike grunner ikke kan følges opp ved kontrollene.

Databehandling

Primærdata legges inn i SPSS, behandlingsgrupper må være ukjent for den som puncher. Data skal være koblet på registreringsnummer. På en uavhengig datafil skal registreringsnummer og navn kobles.

Ryggdatabasen ved Universitetssykehuset i Nord-Norge benyttes til registreringen. Skjema som benyttes i denne studien følger også stort sett malene fra denne databasen. Databasen kan bygges ut ved de enkelte behandlingssteder.

Beregning av utvalgsstørrelse

Man ønsker å oppdage forskjeller på 10 prosentpoeng på Oswestry score.

Basert på tidligere studier antas baseline standard deviasjon å ligge på 12. Resultatet kan også måles som enten en "suksess" eller "ingen suksess" etter 2 år ut fra pasientenes egen evaluering hvor "mye bedre" og "bedre" er "suksess", mens "uendret" eller "verre" er "ingen suksess". Suksess antas å forekomme hos 5 % av pasientene i den konservative gruppen.

Sannsynligheten for å oppdage en statistisk forskjell (power) mellom gruppen skal være minst 80 %. Man ønsker et signifikansnivå på 0.05. Ut fra dette blir $n=62$ i hver gruppe. Ved multisenterstudier regner man 20 % økning i utvalgsstørrelsen hvilket gjør at man ønsker å inkludere 75 pasienter i hver gruppe.

Statistiske metoder

Parametrisk eller non-parametrisk test avhengig om materialet er normalfordelt. T-test eller Wilcoxon's sign rank test for vurdering av effekt. Flerveis Anova eller regresjon for kontroll av cofoundere/skjevfordeling eller for vurdering av flere effektvariable.

Multiple regresjon eller logistisk regresjon for vurdering av prediktorer av respondere/non-respondere på henholdsvis operasjon eller trening.

Responder defineres som positiv effekt etter pasientens svar på spørreskjema: "Hvordan vurderer du ryggen din nå?"

Delprosjekter

- Røntgenanalyser
- RSA studie vedrørende migrasjon og bevegelse
- Discografi (egen protokoll vedlagt)
- DCRA studie ved biomekanikk
- Fysiske tester
- Prediktorer for respondere vs non-responders i de ulike behandlingsgrupper

Referanser:

Abrams P, Cardozo L, Fall M, et al. The standardisation of terminology of lower urinary tract function: report from the standardisation sub-committee of the international continence society. *Neurourology and Urodynamics* 2002;21:167-178.

Bertagnoli R, Kumar S. Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. *Eur Spine J* 2002;11:S131-S136

Brox JI, Sørensen R, Friis A et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913-1921

Cinotti G, David T, Postacchini F. Results of disc prosthesis after a minimum follow-up of 2 years. *Spine* 1996;21:995-1000

Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974; 19(1):1-15.

EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy* 1990; 16(3):199-208.

Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980; 66(8):271-273.

Fritzell P, Hägg O, Wessberg P. 2001 Volvo award winner in clinical studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain. A multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001;26:2521-2534

Fritzell P, Hagg O, Jonsson D, Nordwall A and the Swedish Lumbar Spine Study Group. Cost-Effectiveness of Lumbar Fusion and Nonsurgical Treatment for Chronic Low Back Pain in the Swedish Lumbar Spine Study: A Multicenter, Randomized, Controlled Trial From the Swedish Lumbar Spine Study Group. *Spine* 2004;29(4):421-434.

Furlan AD, Clarke J, Esmail R et al. A critical review of reviews on the treatment of chronic low back pain. *Spine* 2001;26:E155-162

Greenough CG, Petersom MD, Hadlow S, Fraser RD. Instrumented posterolateral lumbar fusion. *Spine* 1998;23:479-486

Hagen KB, Thune O. Work incapacity from low back pain in the general population. *Spine* 1998;23:2480-2487

Hart LG, Deyo RA, Cherkin DC. Physicians office visits for low back pain: frequency, clinical evaluation and treatment patterns from a US national survey. *Spine* 1995;20:11-19

Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. *Arthritis Rheum* 1989; 32(1):37-44.

Öhlund C, Eek C, Palmbald S, Areskoug B, and Nachemson A. Quantified pain drawing in subacute low back pain. Validation in a nonselected outpatient industrial sample. *Spine* 1996; 21: 1021-1030

van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain: a systematic review of randomized controlled trials of the most common interventions. *Spine* 1997;22:2128-2156

Von Korff M, Wagner EH, Dworkin SF et al. Chronic pain and use of ambulatory health care. *Psychosom Med* 1991;53:61-79

Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear- avoidance beliefs in chronic low back pain and disability. *Pain* 1993; 52(2):157-16

Vedlegg 5

Klassifikasjonssystem for uspesifikke kroniske ryggsmarter ad modum O'Sullivan

Teori bak klassifikasjonssystemet og behandlingsprinsipper

Behandlingen som knyttes til O'Sullivans klassifikasjonssystemet bygger på nyere behandlings prinsipper innen fysioterapi hvor man i den senere tid har fokusert på nevro-muskulær kontroll. Her vektlegges samspillet mellom passive strukturer (vertebra, intervertebral disk, zygapophyseal ledd og ligamenter), aktive strukturer (sener og muskler) og nervesystemet (perifere- og sentrale nervesystem) (Panjabi et al., 1989; Panjabi, 1992a). Målet for disse systemene er sammen å opprettholde den intervertebrale neutrale sonen (Panjabi, 1992b). Panjabi viste i sin studie fra 1989 at den nøytrale sonen øker ved intersegmentale skader og diskdegenerasjon og minsker ved muskelstimulering rundt et bevegelsessegment. Størrelsen på den nøytrale sonen er et viktig mål på den segmentale stabiliteten.

Teorigrunnlaget bygger videre på Bergmarks inndeling av muskulatur i to systemer: det globale muskelsystemet som produserer kraft og hvor muskulaturen ikke har direkte feste på columna, og det lokale muskelsystemet bestående av muskulatur som har feste direkte på lumbale vertebra og som sørger for segmental stabilitet og utøver en direkte kontroll på lumbale segmenter. Ved smerte blir det lokale muskelsystemet svekket, hvilket kroppen kompenserer med økt global muskelaktivitet. Ubalanse i disse muskelsystemene som oppstår ved muskelskjelettsmerte mener man kan føre til nedsatt neuromuskulær kontroll, økning av den segmentale nøytrale sonen og en endring av bevegelsesmønster (Bergmark, 1989; O'Sullivan et al., 1997; Sterling et al., 2001; Sahrman, 2002). Studier har vist at normalisering av dyp stabiliserende muskulatur som transversus abdominis, diaphragms og lumbale multifider leder til økt segmental kontroll av den nøytrale sonen (Hodges & Richardson, 1996).

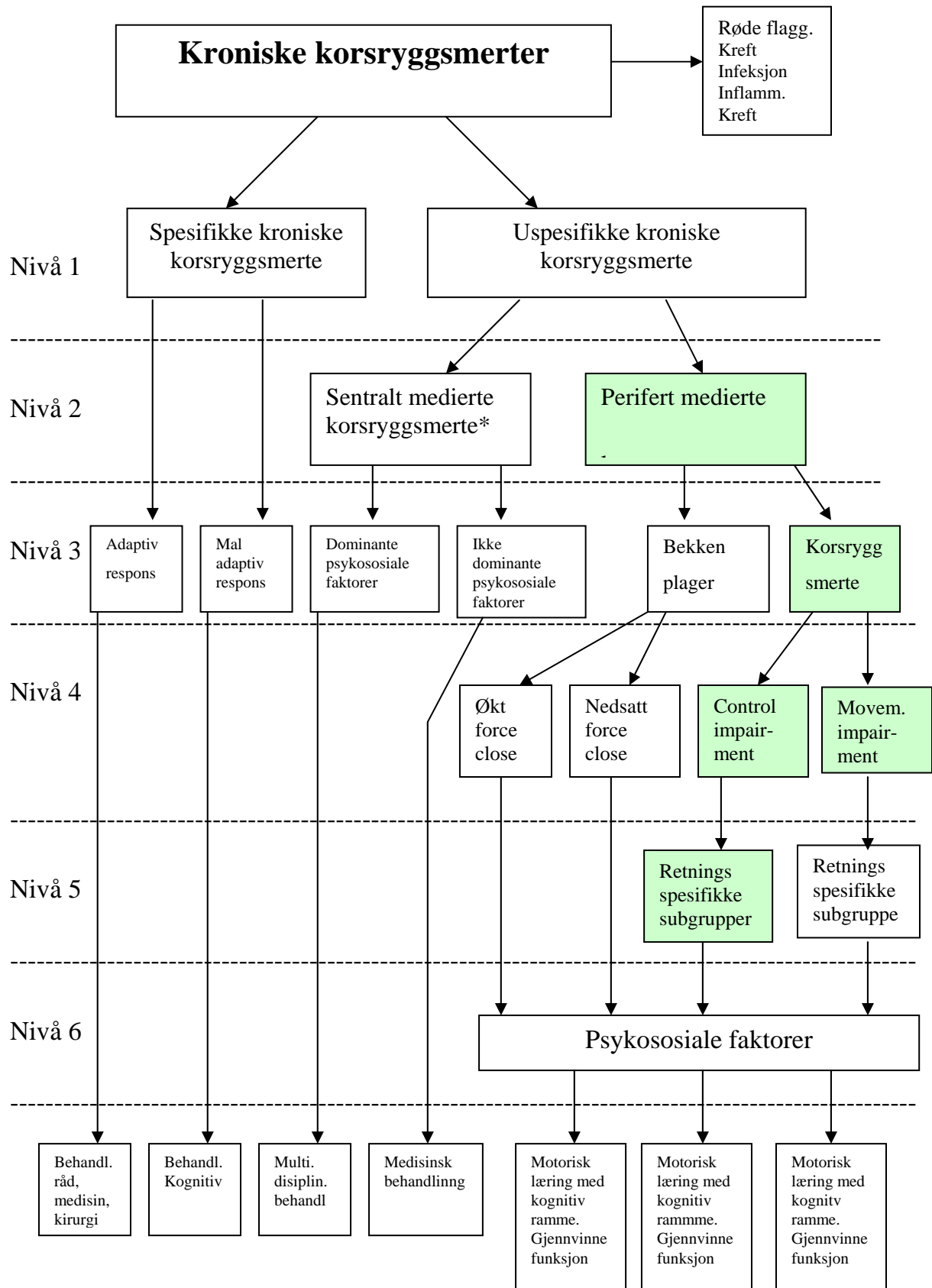
Randomiserte kontrollerte studier utført av fysioterapeuter basert på dette

teorigrunnlaget har vist positiv effekt både overfor pasienter med akutte korsryggsmerter (Hides et al., 1996; Hides et al., 2001), kroniske korsryggsmerter (Niemisto et al., 2003; Rasmussen-Barr et al., 2003), spondylolisthese (O'Sullivan et al., 1997) og pasienter med bekkensmerter etter fødsel (Stuge et al., 2004).

Behandlingsprinsippene for pasienter med nedsatt motorisk segmental kontroll baseres på en modell for motorisk læring der man identifiserer de uhensiktsmessige bevegelsesmønsteret eller mønstrene (Sahrmann, 2002) Bevegelsens ulike momenter isoleres og trenes på nytt for å settes sammen i funksjonelle øvelser som er tilpasset pasientens individuelle behov. I dette behandlingssystemet inngår mye kognitive komponenter som riktig forståelse av problemet og hva som vedlikeholder smerten. Pasienten må medvirke aktivt i behandlingsprosessen og ny motorisk læring må implementeres i dagliglivet funksjoner (O'Sullivan, 2005).

Subgruppering ad modum O'Sullivan

Kjartan Vibe Fersum som er doktorgradstipendiat ved Universitetet i Bergen har nylig publisert (article in press) en reliabilitetsstudie på klassifikasjonssystemet til O'Sullivan og i sin artikkel presenterer han klassifikasjonssystemet som det fremstår i 2008 (Vibe et al., 2008). Klassifikasjonen skjer ut fra anamnese, kliniske undersøkelser, billeddiagnostikk, pasientskjemaer for smerte, fysisk funksjon, fear avoidance beliefs, angst og depresjon (figur A).



* Sentralt medierte korsryggsmerter: Gule flagg; smertene sekundære psykososiale faktorer (ikke organiske)

Figur A. Klassifikasjonsprosess av O'Sullivan hentet fra Vibe Fersum's artikkel 2008 (O'Sullivan, 2005; O'Sullivan & Beales, 2007b; O'Sullivan & Beales, 2007a; Vibe et al., 2008)

Ved subgruppering til *control impairment* (MCI) eller *movement impairment* (MI) ser men etter følgende kriterier:

Kontroll dysfunksjon

- Lokalisert LBP som provoseres mekanisk
- Ingen nedsatt bevegelse i smerteretningen
- Smerteprovokasjon: i nøytral sone ved belastning, repeterte yterstillings strekk, retningsbasert og smerten øker ofte gradvis
- Nedsatt kontroll i nøytral sone
- Liten kjennskap om smertemekanismer
- Repetert bevegelse i smertefull retning øker smerten
- Økt kontroll i smertefull retning reduserer smerten
- Varierende kognitive/fear avoidance/psykososiale faktorer

Bevegelses dysfunksjon

- Lokalisert mekanisk årsak til smerte
- Nedsatt spinal bevegelse i smerteretningen: segmental /multisegmental/retningsbestemt/ ved belastning
- Både aktiv og passiv bevegelses nedsettelse
- Muskulær beskyttelse/ smerte som kommer direkte ved provokasjon
- Tilbakekommende smerterespons i smerteretningen
- Smerten forverres av stabiliserende øvelser
- Patologisk redsel for bevegelse
- Ekstremt påpasselig
- Redsel og uro

Kontroll dysfunksjon og bevegelses dysfunksjon subgrupperes videre etter smerteretning og bevegelsesmønster. Spesielt ved kontroll dysfunksjon er det satt opp kriterier for å klassifiseres til de ulike smerte mønstrene (vedlegg 2).

Vedlegg 6

Hovedkriterier for skåring til de ulike subgruppene ut fra anamnese og video

Movement impairment

Pasienten har innskrenket bevegelighet og klarer ikke ta ut full bevegelse i en retning eller flere. Stunt stopp ofte med plutselig smerte. Vanlig forbundet med fear avoidance beliefs. Høg beskyttende muskelaktivitet med tendens til å låse hele lumbal - bekken når man går inn i smertefull retning.

Control impairment

Ingen nedsatt bevegelighet i smerteretningen, men nedsatt kontroll av det symptomatiske spinale segmentet. Pasienten har gradvis økende smerte etter repeterte og vedvarende strekk av vevet. Ofte kompensatoriske spenninger i ovenfor - og nedenforliggende ledd. Nedsatt kontroll av lumbal - bekken posisjonering.

1) Fleksjonsmønster

Ved holdnings og bevegelsesanalyse:

- i stående ses avflatet lumballordose og bakover rotert bekken
- ved fremover bøyning ses økt fleksjon ved det symptomatiske segmentet med kompensatorisk ekstensjon i thorakal - lumbal delen, bekkenet blir stående og hoftelddet brukes i liten grad
- ved bakover bøyning ses nedsatt ekstensjon i det symptomatiske segmentet med økt ekstensjon i thorakal - lumbal overgangen
- ved knebøy ses økt fleksjon i det symptomatiske segmentet og bakover rotert bekken
- ved sitting ses økt fleksjon i nedre lumbal columna, bakover rotert bekken og ekstendert thoraco-lumbal overgang.

- ved oppreising ses økt fleksjon i det symptomatiske segmentet og ekstensjon i thorakal - lumbal overgangen med bakover rotert bekken
- problem å rotere bekkenet fremover og ekstendere nedre lumbal isolert fra øvre lumbal og thorakal

Anamnese:

- sentral ryggsmerte ofte årsaks relatert til en "kink" ved fremverbøying / rotasjon eller flere gjentatte fremover bøyninger/rotasjoner
- problem ved fremover bøyning og det å være i semiflekterte posisjoner med vanskelighet å komme opp igjen og må ofte bruke hendene som støtte
- smerteprovokasjon ved fleksjonsrelaterte holdninger og aktiviteter
- smertelettelse ved ekstensjons aktiviteter og posisjoner der lumbalcolumna er lordosert

2) *Fleksjonsmønster med shift*

Ved holdnings og bevegelses analyse:

- i stående som ved fleksjonsmønster, men med en lateral shift ved de symptomatiske segmentet
- ved fremover bøyning ses økt fleksjon i symptomatisk segment med lateral deviasjon av trunkus ovenfor dette segmentet
- ved oppreising som ved fleksjonsmønster men med deviasjon ved det symptomatiske segmentet
- ved bakover bøyning som ved fleksjonsmønster men med lateral deviasjon
- ved ett ben stående ses lateral shift av thoraks i forhold til bekkenet, kan gi positiv Trendelenburg
- ved knebøy som ved fleksjonsmønster, men med lateral deviasjon
- ved sitting som ved fleksjonsmønster, men med deviasjon
- ved oppreising økt fleksjon og lateraldeviasjon i problematisk segment
- problem å rotere bekkenet fremover og ekstendere nedre lumbal isolert fra øvre lumbal og thorakal

Anamnese:

- unilaterale korsryggsmerter

- smerteprovokasjon ved rotasjon eller ved å strekke seg samtidig som holdningen er fremover bøyet, samme posisjon som forårsaket ryggsmerte første gang
- smertelettelse ved ekstensjonsaktiviteter og posisjoner, strekke til motsatt side av shiften eller korrigere shiften

3) Aktiv ekstensjonsmønster

Ved holdnings og bevegelses analyse:

- i stående ses en økt segmental lordose ved det symptomatiske segmentet
- thoraks står anteriørt for bekkenet
- ved fremover bøyning ses hyperekstensjon av symptomatisk segment og problem med å slippe lordosen, overdreven fremover rotasjon av bekkenet
- ved oppreising hyperekstenderes det symptomatiske segmentet tidlig
- ved bakover bøyning ses økt ekstensjon i det symptomatiske segmentet med fremover rotasjon av bekken
- ved et ben stående skyves bekkenet bak og hoften på standbenet roteres innover
- ved knebøy ses økt ekstensjon i det symptomatiske segmentet og fremover rotert bekken
- ved sitting holdes lordosen og bekkenet er fremover rotert
- ved oppreising holdes lordosen hele veien opp

Anamnese

- sentrale korsrygg smerter ofte årsaksrelatert til smertedebut i forbindelse med et ekstensjon/rotasjons traume eller repeterte belastninger i ekstensjon som for eksempel sportsaktiviteter eller arbeide med mye ekstensjonskomponenter
- smerte provokasjon ved ekstensjons relaterte holdninger og aktiviteter, men også ved fremover bøyning da man holder lordosen i hyperekstensjon
- smertelettelse fleksjons holdninger og aktiviteter der lumbalcolumna er flektert

4) Passivt ekstensjonsmønster

Ved holdnings og bevegelses analyse:

- ved stående holdning ses økt lordose i det symptomatiske segmentet med en nedsatt lordose oversegmentet som gir en lang thorakal kyfose som får tyngdepunkt posteriovert for bekkenet en så kalledt "sway holdning"
- ingen problem med fremover bøyning
- ved oppreising ses tendens til hyperekstensjon i det symptomatiske segmentet
- ved bakover bøyning ses økt ekstensjon i det symptomatiske segmentet med redusert ekstensjon over det instabile segmentet og overdrevet fremskutt bekken
- ved et ben stående skyves bekkenet anteriort og hoften på standbenet roterer innover
- ingen problem ved knebøy
- ved sitting sitter pasienten i "slump posisjon" der hele columna er flektert
- ved oppreising ses en hyperekstensjon av det symptomatiske segmentet og overdreven fremskyvning av bekkenet ved overgang til oppreiset posisjon

Anamnese

- sentrale korsryggsmerter ofte årsaksrelatert til et traume eller repeterte overbelastninger i ekstensjon
- smerteprovokasjon samme som ved aktivt ekstensjons mønster der ekstensjons holdninger og aktiviteter provoserer
- smertelettelse ved fleksjons holdninger og aktiviteter der lumbalcolumna er i nøytral stilling eller flektert

5) Multidireksjonalt mønster

Ved holdnings og bevegelses analyse:

- ved stående holdning ses variasjon i alle mønster
- ved fremover bøyning ses økt fleksjon i det symptomatiske segmentet
- ved oppreising ses variasjon og alturnering i strategier
- ved bakover bøyning ses økt ekstensjon i det symptomatiske segmentet
- ved et ben stående varieres strategiene
- ved knebøy varieres strategiene
- ved sitting varieres strategiene

- ved oppreising enten fleksjon eller ekstensjon i det ustabile segmentet

Anamnese:

- store smerter og funksjonstap
- vanlig å låse ryggen for å plutselig gå over til fleksjon, ekstensjon, shift eller å henge i det symptomatiske segmentet
- alle vekt bærende posisjoner er smertefulle og provoserer

Vedlegg 7

SJEKKLISTE VED VIDEOFILMNING

Fysiot: _____ Randnr: _____

Bakfra		Klarer	Klarer ikke	Smerter
1. Stående holdning	krum - svai			
2. Fremoverbøy				
3. Oppreising				
4. Bakoverbøy				
5. Sidebøy bilat.				
6. Et ben stående x 2				
7. Knebøy				
8. Krakk, sittende holdning				
9. Sittende krum-svai				
10. Sittende til stående				
Fra siden		Klarer	Klarer ikke	Smerter
11. Stående holdning	krum - svai			
12. Fremoverbøy				
13. Oppreising				
14. Bakoverbøy				
15. Knebøy				
16. Krakk, sittende holdning				
17. Sittende krum-svai				
18. Sittende til stående				

Vedlegg 8

Fysiot:_____ Rand nr:_____

Anamnese

Alder:

Kjønn:

Arbeid:

Sport/fritids aktiviteter:

Hovedproblem:

Varighet LBP:

Skademekanisme:

Smerte:

Smerte provokasjon:

- **Bevegelse:**

- **Aktiviteter:**

- **Holdning:**

- **Behandling:**

- **Trening:**

Smerte lettelse:

- **Bevegelse:**

- **Aktiviteter:**

- **Holdning:**

- **Behandling:**

- **Trening:**

Vedlegg 9

Observatør: _____ Randnr. _____

Movement impairment Control impairment

Klassifikasjon av bevegelses mønster

Fleksjon			
Fleksjon + Lateral "shift"			
Retning "Shift"	V		
	H		
Passiv Ekstensjon			
Aktiv Ekstensjon			
Multi- directional			
Ekt.	Lat. Shift V		Lat. Shift R
Fleks.	Lat. Shift V		Lat. Shift R
Begrunnelse for klassifikasjon:			
Sikker <input type="checkbox"/> Usikker <input type="checkbox"/>			

Vedlegg 10

OSWESTRY LISTE FOR FUNKSJONSBEGRENSNINGER, versjon 2.0

Vennligst les: Dette spørreskjemaet er utformet for å gi behandleren opplysninger om hvordan ryggsmertene dine har påvirket din evne til å klare deg i dagliglivet. Vennligst svar på hvert avsnitt, og marker bare **det ene feltet** i hvert avsnitt som gjelder for deg. Vi forstår at du kanskje synes at to av utsagnene i hvert avsnitt gjelder deg, men vennligst **marker bare feltet som best beskriver ditt nåværende problem.**

Del 1 – Smerteintensitet

- 1. Jeg har ingen smerter for øyeblikket
- 2. Smertene er veldig svake for øyeblikket
- 3. Smertene er moderate for øyeblikket
- 4. Smertene er temmelig sterke for øyeblikket
- 5. Smertene er veldig sterke for øyeblikket
- 6. Smertene er de verste jeg kan tenke meg for øyeblikket

Del 2 – Personlig stell (vaske seg, kle på seg, osv.)

- 1. Jeg kan stelle meg selv på vanlig måte uten at det forårsaker ekstra smerter
- 2. Jeg kan stelle meg selv på vanlig måte, men det er veldig smertefullt
- 3. Det er smertefullt å stelle meg selv, og jeg gjør det langsomt og forsiktig
- 4. Jeg trenger noe hjelp, men klarer det meste av mitt personlige stell
- 5. Jeg trenger hjelp hver dag til det meste av eget stell
- 6. Jeg kler ikke på meg, har vanskeligheter med å vaske meg, og holder sengen

Del 3 – Løfte

- 1. Jeg kan løfte tunge ting uten å få mer smerter
- 2. Jeg kan løfte tunge ting, men får mer smerter
- 3. Smertene hindrer meg i å løfte tunge ting opp fra gulvet, men jeg greier det hvis det som skal løftes er gunstig plassert, f.eks. på et bord
- 4. Smertene hindrer meg i å løfte tunge ting, men jeg kan klare lette eller middels tunge ting, hvis det er gunstig plassert
- 5. Jeg kan bare løfte noe som er veldig lett
- 6. Jeg kan ikke løfte eller bære noe i det hele tatt

Del 4 – Gå

- 1. Smerter hindrer meg ikke i å gå i det hele tatt
- 2. Smerter hindrer meg i å gå mer enn 1 ½ km
- 3. Smerter hindrer meg i å gå mer enn ¾ km
- 4. Smerter hindrer meg i å gå mer enn 100 m
- 5. Jeg kan bare gå med stokk eller krykker
- 6. Jeg ligger for det meste i sengen og jeg må krabbe til toalettet

Del 5 – Sitte

- 1. Jeg kan sitte så lenge jeg vil i en hvilken som helst stol
- 2. Jeg kan sitte så lenge jeg vil i min favorittstol
- 3. Smerter hindrer meg i å sitte i mer enn en time
- 4. Smerter hindrer meg i å sitte i mer enn en halv time
- 5. Smerter hindrer meg i å sitte i mer enn ti minutter
- 6. Smerter hindrer meg i å sitte i det hele tatt

Forts. OSWESTRY LISTE FOR FUNKSJONSBEGRENSNINGER

Del 6 – Stå

- 1. Jeg kan stå så lenge jeg vil uten å få mer smerter
- 2. Jeg kan stå så lenge jeg vil, men får mer smerter
- 3. Smerter hindrer meg i å stå i mer enn en time
- 4. Smerter hindrer meg i å stå i mer enn en halv time
- 5. Smerter hindrer meg i å stå i mer enn ti minutter
- 6. Smerter hindrer meg i å stå i det hele tatt

Del 7 – Sove

- 1. Søvnene min forstyrres aldri av smerter
- 2. Søvnene min forstyrres av og til av smerter
- 3. På grunn av smerter får jeg mindre enn seks timers søvn
- 4. På grunn av smerter får jeg mindre enn fire timers søvn
- 5. På grunn av smerter får jeg mindre enn to timers søvn
- 6. Smerter hindrer all søvn

Del 8 – Seksualliv

- 1. Seksuallivet mitt er normalt og forårsaker ikke mer smerter
- 2. Seksuallivet mitt er normalt, men forårsaker noe mer smerter
- 3. Seksuallivet mitt er normalt, men svært smertefullt
- 4. Seksuallivet mitt er svært begrenset av smerter
- 5. Seksuallivet mitt er nesten borte på grunn av smerter
- 6. Smerter forhindrer alt seksualliv

Del 9 – Sosialt liv

- 1. Det sosiale livet mitt er normalt og forårsaker ikke mer smerter
- 2. Det sosiale livet mitt er normalt, men øker graden av smerter
- 3. Smerter har ingen betydelig innvirkning på mitt sosiale liv, bortsett fra at de begrenser mine mer fysiske aktive sider, som sport osv.
- 4. Smerter har begrenset mitt sosiale liv og jeg går ikke så ofte ut
- 5. Smerter har begrenset mitt sosiale liv til hjemmet
- 6. På grunn av smerter har jeg ikke noe sosialt liv

Del 10 – Reising

- 1. Jeg kan reise hvor som helst uten smerter
- 2. Jeg kan reise hvor som helst, men det gir mer smerter
- 3. Smertene er ille, men jeg klarer reiser på to timer
- 4. Smerter begrenser meg til korte reiser på under en time
- 5. Smerter begrenser meg til korte, nødvendige reiser på under 30 minutter
- 6. Smerter forhindrer meg fra å reise, unntatt for å få behandling

Skåring; kode om spørsmålene til 0-5 (1=0...6=5). Summer hvert spørsmål, deles med antall besvarte spørsmål, multipliseres med 0,2 og 100 for å få en prosentskår.

The Modified Oswestry Disability Index (Baker et al 1990)
Oversatt av Margreth Grotle og Nina K-Vøllestad 2001,
Seksjon for Helsefag, Universitetet i Oslo

Vedlegg 11

SMERTE, FYSISK AKTIVITET OG JOBB
(Fear-Avoidance Beliefs Questionnaire,
Waddell et al 1993)

Her er noe av det som andre har fortalt oss om ryggsmertene sine. Kryss av for ett tall fra 0 (*helt uenig*) til 6 (*helt enig*) for hvert utsagn for å si hvor mye fysiske aktiviteter som å bøye seg, løfte, gå eller kjøre vil påvirke ryggen *din*.

		HELT UENIG		USIKKER		HELT ENIG		
		0	1	2	3	4	5	6
1	Smertene mine ble forårsaket av fysisk aktivitet							
2	Fysisk aktivitet forverrer smertene mine							
3	Fysisk aktivitet kan skade ryggen min							
4	Jeg burde ikke utføre fysiske aktiviteter som (kan) forverre smertene mine							
5	Jeg kan ikke utføre fysiske aktiviteter som (kan) forverre smertene mine							

Følgende utsagn handler om hvordan det vanlige arbeidet ditt påvirker eller kan påvirke ryggsmertene dine

		HELT UENIG		USIKKER		HELT ENIG		
		0	1	2	3	4	5	6
6	Smertene mine ble forårsaket av arbeidet mitt eller et uhell på jobben							
7	Arbeidet mitt forverrer smertene mine							
8	Jeg har framsatt erstatningskrav for smertene mine							
9	Arbeidet mitt er for tungt for meg							
10	Arbeidet mitt forverrer eller kan forverre smertene mine							
11	Arbeidet mitt kan skade ryggen min							
12	Jeg burde ikke utføre det vanlige arbeidet mitt med mine nåværende smerter							
13	Jeg kan ikke utføre det vanlige arbeidet mitt med mine nåværende smerter							
14	Jeg kan ikke utføre det vanlige arbeidet mitt før smertene er behandlet							
15	Jeg tror ikke jeg vil være tilbake på det vanlige arbeidet mitt innen tre måneder							
16	Jeg tror ikke jeg noen gang vil være i stand til å komme tilbake til det arbeidet							

The Fear-Avoidance Beliefs Questionnaire (FABQ) (Waddell et al 1993)
 Oversatt av Margreth Grotle og Nina K:Vøllestad 2001,
 Seksjon for Helsefag, Universitetet i Oslo,

Vedlegg 12

Hvordan har du det?

Når smerter og andre plager har vart en tid, blir en gjerne sliten og oppgitt. Dette gir ofte slike plager som nevnt nedenfor. Samlet blir disse her brukt som mål på at en er legemlig og psykisk presset. Vurder hvor mye hvert symptom har vært til plage eller ulempe for deg de siste 14 dagene (til og med i dag). Sett ring rundt tallet som passer best. Husk å sette en ring rundt aktuelt tall for hver plage/hvert symptom.

(sett ring rundt tallet)	Ikke i det hele tatt	Litt	En god del	Svært mye
1. Plutselig skremt uten grunn.	1	2	3	4
2. Føler du deg engstelig.	1	2	3	4
3. Føler du deg svimmel eller kraftløs.	1	2	3	4
4. Nervøs eller urolig.	1	2	3	4
5. Hjerterbank.	1	2	3	4
6. Skjelving.	1	2	3	4
7. Føler deg ansent eller opphisset.	1	2	3	4
8. Hodepine.	1	2	3	4
9. Anfall av redsel eller panikk	1	2	3	4
10. Rastløshet, kan ikke sitte rolig	1	2	3	4
11. Føler deg slapp og uten energi.	1	2	3	4
12. Anklager deg selv for ting.	1	2	3	4
13. Har lett for å gråte.	1	2	3	4
14. Tap av seksuell interesse/opplevelse.	1	2	3	4
15. Dårlig appetitt.	1	2	3	4
16. Vanskelig for å sove.	1	2	3	4
17. Følelse av håpløshet mht. framtiden.	1	2	3	4
18. Føler deg nedfor.	1	2	3	4
19. Føler deg ensom.	1	2	3	4
20. Har tanker om å ta ditt eget liv.	1	2	3	4
21. Følelse av å være fanget.	1	2	3	4
22. Bekymrer deg for mye.	1	2	3	4
23. Føler ikke interesse for noe.	1	2	3	4
24. Føler at alt krever stor anstrengelse.	1	2	3	4
25. Føler at du ikke er noe verd.	1	2	3	4

HSCL-25