

# **Multidimensjonal stratifisert behandling av uspesifikke korsryggsmerter: en systematisk oversikt**

Alexander Emil Berge



Masteroppgave i helsefag

Studieretning Klinisk masterstudium i manuellterapi for fysioterapeuter

Institutt for global helse og samfunnsmedisin

Universitetet i Bergen

Høst 2016

Antall ord: 10075

## Forord

Som fysioterapeut møter jeg en stor andel pasienter med langvarige muskel-skjelettplager. Blant disse er særlig uspesifikke korsryggsmerter hyppige, og erfaring tilsier at disse er vanskelig å behandle. Min oppfatning er at mange fysioterapeuter behandler disse pasientene høyst ulikt, med ulike tilnærminger og forklaringsmodeller. Samtidig opplever jeg et stadig større fokus på evidensbasert praksis. Som et forsøk på å hjelpe klinikere i behandling av uspesifikke korsryggsmerter, er det utviklet en rekke tilnærminger som forsøker å gi målrettet behandling til sub-grupper av denne pasientgruppen, såkalt stratifisert behandling. Selv om jeg ikke har personlig erfaring med denne type behandling, synes jeg dette virker spennende og har tro på at det kan bidra til å hjelpe klinikere til mer målrettet og effektiv behandling av korsryggsmerter. Jeg ønsket derfor å finne mer ut av evidensen for stratifisert behandling av korsryggsmerter.

Arbeidet med denne studien har vært en interessant og berikende prosess som har utviklet meg som kliniker i forståelsen av korsryggsmerter. Jeg ønsker å takke de som har bidratt og hjulpet meg gjennom denne prosessen. Først en stor takk til min veileder Kjartan Vibe Fersum for enestående hjelp og veiledning. Jeg vil også takke bibliotekar Regina Küfner Lein for uvurderlig hjelp i søkeprosessen. I tillegg vil jeg takke for gode råd fra studiekolleger. Sist men ikke minst vil jeg takke min nærmeste familie for oppmuntring og støtte gjennom hele prosessen.

Alexander Emil Berge

Oktober 2016

## DEFINISJONER AV FORKORTELSER

|        |                                                                    |
|--------|--------------------------------------------------------------------|
| CFT    | Cognitive functional therapy                                       |
| CI     | Konfidensintervall                                                 |
| MDT    | McKenzies Mechanical Diagnosis and Therapy                         |
| MSI    | Movement System Impairment                                         |
| MUS    | Medisinske uforklarlige symptomer                                  |
| NSLBP  | Uspesifikke korsryggsmerter                                        |
| NSCLBP | Uspesifikke kroniske korsryggsmerter                               |
| ODI    | Oswestry Disability Questionnaire                                  |
| PNRS   | 11-punkts numerisk smerteskala                                     |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analysis |
| RCT    | Randomiserte kontrollerte studier                                  |
| RMDQ   | Roland and Morris Disability Questionnaire                         |
| SB     | Start Back stratifiseringsmodell                                   |
| SF-12  | 12-item Short Form Health Survey                                   |

# Innholdsfortegnelse

|                                                                              |           |
|------------------------------------------------------------------------------|-----------|
| <b>SAMMENDRAG</b> .....                                                      | <b>5</b>  |
| <b>ABSTRACT</b> .....                                                        | <b>6</b>  |
| <b>1. INTRODUKSJON</b> .....                                                 | <b>7</b>  |
| 1.1. BAKGRUNN .....                                                          | 7         |
| 1.2. HENSIKT .....                                                           | 8         |
| <b>2. TEORI</b> .....                                                        | <b>9</b>  |
| 2.1. HVA ER USPESIFIKKE KORSRYGGSMERTER .....                                | 9         |
| 2.2. STRATIFISERT BEHANDLING.....                                            | 11        |
| 2.2.1. <i>Vurdering av metodologiske faktorer ved klassifisering</i> .....   | 13        |
| 2.3. UTFORDRINGER MED MEDISINSKE UFORKLARLIGE SYMPTOMER.....                 | 14        |
| 2.4. SYSTEMATISKE OVERSIKTER – STYRKER OG SVAKHETER.....                     | 15        |
| <b>3. METODE</b> .....                                                       | <b>17</b> |
| 3.1. UTVALG .....                                                            | 17        |
| 3.1.1. <i>Inklusjonskriterier</i> .....                                      | 17        |
| 3.1.2. <i>Eksklusjonskriterier</i> .....                                     | 18        |
| 3.2. SØKEMETODER FOR IDENTIFISERING AV STUDIER .....                         | 18        |
| 3.3. DATAPRODUKSJON OG ANALYSE .....                                         | 19        |
| 3.3.1. <i>Seleksjon av studier</i> .....                                     | 19        |
| 3.3.2. <i>Ekstraksjon av data</i> .....                                      | 19        |
| 3.3.3. <i>Vurdering av metodologisk kvalitet av inkluderte studier</i> ..... | 20        |
| <b>4. RESULTATER</b> .....                                                   | <b>21</b> |
| 4.1. SELEKSJON AV STUDIER .....                                              | 21        |
| 4.2. EKSKLUDERTE STUDIER .....                                               | 22        |
| 4.3. KARAKTERISTIKKER AV INKLUDERTE STUDIER .....                            | 22        |
| 4.4. METODOLOGISK KVALITET AV INKLUDERTE STUDIER.....                        | 22        |
| 4.5. EFFEKT AV INTERVENSJONER.....                                           | 22        |
| <b>5. DISKUSJON</b> .....                                                    | <b>26</b> |
| 5.1. OPPSUMMERING AV RESULTATER.....                                         | 26        |
| 5.2. METODOLOGISK KVALITET AV INKLUDERTE STUDIER.....                        | 28        |
| 5.3. STYRKER OG SVAKHETER VED DE TO KLASIFISERINGSSYSTEMENE .....            | 29        |
| 5.4. BETRAKTNINGER AV STRATIFISERT BEHANDLING .....                          | 33        |
| 5.5. BEGRENSENINGER MED DENNE SYSTEMATISKE OVERSIKTEN .....                  | 35        |
| <b>6. KONKLUSJON</b> .....                                                   | <b>36</b> |
| <b>LITTERATURLISTE</b> .....                                                 | <b>37</b> |
| <b>VEDLEGG</b> .....                                                         | <b>45</b> |

# SAMMENDRAG

**Studiedesign:** Systematisk oversikt

**Bakgrunn:** Uspesifikke korsryggsmerter (NSLBP) er en heterogen tilstand som har vist seg vanskelig å behandle. For å øke behandlingseffekter er stratifisert behandling ansett som en lovende tilnærming. Det innebærer at pasienter med like karakteristikk klassifiseres i sub-grupper som får målrettet behandling. Spesielt er det anbefalt at stratifisert behandling bør baseres på en biopsykososial modell hvor alle dimensjoner blir inkludert.

**Hensikt:** Å systematisk gjennomgå litteraturen for effekt av multidimensjonal stratifisert behandling av NSLBP sammenlignet med ikke-stratifisert behandling.

**Metode:** Systematiske søk i databasene EMBASE, MEDLINE, CINAHL, PEDRO, og PsycINFO ble gjennomført sammen med en erfaren bibliotekar frem til februar 2016, supplert med søk for hånd i referanselister til relevante studier. Kun randomiserte kontrollerte studier (RCT) av multidimensjonal stratifisert behandling ble inkludert. Utfallsmålene var smerte, funksjon og livskvalitet.

**Resultat:** 2 RCT med ulike klassifiseringssystemer ble inkludert (antall deltakere = 972). Begge studiene hadde høy metodisk kvalitet. Den ene studien testet effekten av et prognosebasert klassifiseringssystem (Start Back) og fant beskjedne men signifikante effekter sammenlignet med standard fysioterapi ved 4 og 12 måneder. I den andre studien fant de betydelige og klinisk viktige effekter av et mekanismebasert klassifiseringssystem (Cognitive functional therapy) sammenlignet med manuell terapi og øvelser ved 3 og 12 måneder.

**Konklusjon:** Denne studien viser at multidimensjonal stratifisert behandling av NSLBP er mer effektivt enn standard konservativ behandling. På grunn av få studier, stor variasjon i effektstørrelser samt ulike studiekarakteristikk, må resultatene tolkes med varsomhet. Evidensgrunnlaget for multidimensjonal stratifisert behandling er derfor beskjedent og mer forskning kreves for å kunne gi spesifikke anbefalinger.

**Nøkkelord:** uspesifikke korsryggsmerter, stratifisert behandling, klassifisering, sub-grupper, biopsykososial, systematisk oversikt

## ABSTRACT

**Studiedesign:** Systematic review

**Background:** Non-specific low back pain (NSLBP) is a heterogeneous condition that has proven difficult to treat. To increase treatment effects stratified care is regarded as a promising approach. This means that patients with similar characteristics are classified into subgroups receiving targeted therapy. In particular, it is recommended that stratified care should be based on a biopsychosocial model where all dimensions are included.

**Objective:** To review the effect of multidimensional stratified care of NSLBP compared with non-stratified care.

**Method:** Systematic searches of the databases EMBASE, MEDLINE, CINAHL, PEDRO, and PsycINFO were conducted with an experienced librarian up to February 2016, supplemented by hand searching of reference lists in relevant studies. Only randomized controlled trials (RCT) of multidimensional stratified care was included. The outcome measures were pain, function and quality of life.

**Results:** 2 RCT with different classification systems were included (number of participants = 972). Both studies had a low risk of bias. One study tested the effect of a classification system based on prognostic factors (Start Back) and found modest but significant effects compared to standard physiotherapy at 4 and 12 months. In the second study they found substantial and clinically relevant effects of a mechanism-based classification system (Cognitive functional therapy) compared with manual therapy and exercises at 3 and 12 months.

**Conclusion:** This study shows that multidimensional stratified care of NSLBP is more effective than standard conservative treatment. Because of few studies, large variation in effect sizes and different study characteristics, the results must be interpreted with caution. The evidence base for multidimensional stratified treatment is therefore modest and more research is required to provide specific recommendations.

**Keywords:** non-specific low back pain, stratified care, classification, subgroups, biopsychosocial, systematic review

# 1. INTRODUKSJON

## 1.1. Bakgrunn

Ryggplager er et av de største helseproblemene i den vestlige verden (Ma et al., 2014, Hoy et al., 2014). Blant smerter fra muskel-og skjelettsystemet er korsryggsmerter en av de hyppigst forekommende i Norge, og opptil 90 % av disse omtales som uspesifikke korsryggsmerter (NSLBP) (Freburger et al., 2009, Lærum, 2013). Halvparten av befolkningen har hatt smerter det siste året og 15-20 % opplever ryggsmerter til enhver tid (Lærum, 2013). Det er rapportert at opptil 84 % av befolkningen får korsryggsmerter en eller flere ganger i løpet av livet. Selv om de fleste med akutte korsryggsmerter blir bra i løpet av kort tid, er det vist at korsryggsmerter karakteriseres av tilbakevendende episoder med smerter (Hoy et al., 2010). I vestlige populasjoner er det også observert en økende forekomst av kroniske korsryggsmerter (Freburger et al., 2009). Korsryggsmerter utgjør en av de største helseutgiftene i landet, og er i global sammenheng den viktigste enkeltårsak til redusert funksjonsevne (Lærum, 2013, Hoy et al., 2014). For å optimalisere behandling og redusere byrden på samfunnet og det enkelte individ er derfor korsryggsmerter et prioritert forskningsområde (Balague et al., 2012, Haldeman et al., 2012).

I behandling av korsryggsmerter ser man en trend i økende bruk av fysioterapi, medisiner og kirurgi (Bagnall, 2010, Mafi et al., 2013). Til tross for dette viser ingen intervensjoner betydelig bedre effekt enn andre, og med kun relativt små effekter særlig i det lange løp. Dette gjelder ofte brukte behandlingsmetoder som akupunktur og dry needling, øvelser, manuell terapi, atferdsterapi og kirurgi (Furlan et al., 2005, Hayden et al., 2005, Henschke et al., 2010, Rubinstein et al., 2011, Jacobs et al., 2012). Studier hvor ulike intervensjoner levert av fysioterapeuter har blitt sammenlignet direkte, viser heller ingen store forskjeller i effekt. Dette gjelder for eksempel tradisjonell fysioterapi sammenlignet med spesifikke øvelser (Cairns et al., 2006), manuell terapi sammenlignet med kortvarig atferdsrelatert smertemestringsprogram (Hay et al., 2005), samt kombinert utholdenhets- og styrketrening sammenlignet med gradert aktivitet og trening på problemløsning (Smeets et al., 2008).

Til tross for iherdig forskning på området ser det ut til å være en manglende effekt av ryggbehandling. Flere potensielle årsaker til dette har blitt foreslått, for eksempel kan resultatene påvirkes av et naturlig forløp med spontan bedring, variabilitet i behandlingseffekt på grunn av heterogenitet av studiedeltakere, underestimering av uspesifikke

behandlingseffekter som terapeutens oppmerksomhet og empati, samt pasientens preferanser og forventninger (Foster et al., 2011). Den primære hensikten med intervensjonsstudier er å måle den gjennomsnittlige behandlingseffekten til de inkluderte pasientene. I en heterogen pasientgruppe vil derfor pasienter i varierende grad enten kunne bidra med positiv eller negativ effekt av behandlingen. Den gjennomsnittlige behandlingseffekten står dermed i fare for å bli begrenset av de pasientene som ikke responderer på behandlingen (Ibid). NSLBP representerer ofte en heterogen pasientgruppe (O'Sullivan, 2012). Det har lenge vært antatt at klassifisering av ryggsmarter i mer homogene grupper vil kunne gi mer målrettet behandling og bedre behandlingseffekt (Borkan et al., 2002). Slik klassifisering, hvor intervensjoner blir utformet for å treffe en spesifikk sub-gruppe, har blitt omtalt som stratifisert behandling (Foster et al., 2013).

Et mangfold av ulike klassifiseringssystemer har blitt utviklet og testet, med ulike teorier og metoder (Karayannis et al., 2012, Foster et al., 2013). På bakgrunn av den multidimensjonale natur av korsryggsmarter hvor biologiske, psykologiske og sosiale faktorer interagerer med hverandre, er det spesielt foreslått at stratifisert behandling bør baseres på en biopsykososial modell hvor alle dimensjoner blir ivaretatt i diagnostisering og behandling (O'Sullivan, 2005). Ut i fra vår kjennskap foreligger det ingen systematiske oversikter som har undersøkt effekten av multidimensjonal stratifisert behandling.

Uspesifikke korsryggsmarter er et betydelig helseproblem både for den enkelte pasient og samfunnet. For å bedre adressere alle dimensjonene i den heterogene pasientgruppen som NSLBP representerer, er det antatt at multidimensjonal stratifisert behandling vil gi bedre resultater.

## **1.2. Hensikt**

Hovedformålet med studiet er å systematisk gjennomgå litteraturen for å vurdere effekten av multidimensjonal stratifisert behandling for NSLBP. Mer spesifikt betyr det å undersøke effekten av klassifiseringssystemer som inkluderer både biologiske, psykologiske og sosiale dimensjoner, og hvor matchende behandling til sub-grupper er integrert.



## 2. TEORI

### 2.1. Hva er uspesifikke korsryggsmerter

Korsryggsmerter beskriver smerter i området mellom 12. ribben og glutealfoldene, med eller uten utstråling til underekstremitetene (Lærum et al., 2007). Den mest aksepterte metoden for å klassifisere ryggsmerter er å benytte en diagnostisk tredeling, hvor man differensierer mellom uspesifikke korsryggsmerter (NSLBP), spesifikke korsryggsmerter, og mulig underliggende alvorlig patologi indikert med røde flagg (Lærum et al., 2007, Koes et al., 2010). Alvorlig underliggende patologi forekommer sjeldent og utgjør bare 1-2 % av alle korsryggsmerter. Eksempler er kreft, inflammatorisk sykdom, infeksjoner eller frakturer. Spesifikke korsryggsmerter er primært radikulopater med kompresjon av nerverøtter eller ryggmargen og hvor ryggsmertene er assosiert med nevrologiske utfall og utstrålende smerter. Disse står for 5-10 % av alle korsryggsmerter (Haldeman et al., 2012). Det store flertallet av alle ryggplager faller under begrepet NSLBP og står for omtrent 90 % av alle ryggplager (Koes et al., 2006, Freburger et al., 2009, Balague et al., 2012). NSLBP inkluderer alle korsryggsmerter som ikke kan tilskrives en kjent spesifikk patologi (f.eks. infeksjon, tumor, fraktur, nerverotsaffeksjon eller cauda equina syndrom), og er i hovedsak en diagnose basert på eksklusjon av annen patologi (Koes et al., 2006, Balague et al., 2012).

Selv om NSLBP per definisjon mangler en kjent årsak, er det blitt identifisert mange potensielle årsaker til utvikling og opprettholdelse av smertene. Spesielt har biomedisinske faktorer lenge vært antatt å spille en kausal rolle (Balague et al., 2012). Strukturelle og biomekaniske dysfunksjoner samt svekkelser i motorisk kontroll har blitt fremsatt som potensielle årsaksforklaringer, men populasjonsstudier har til nå vist manglende evidens (O'Sullivan, 2012). Degenerative forandringer påvist med billeddiagnostikk er vist å korrelere dårlig med ryggsmerter. Slike funn er vanlige blant asymptomatiske og ikke sammenfallende med utvikling av NSLBP. Det er derfor ikke holdepunkter for å tilskrive slike funn årsaken til smertene (Endean et al., 2011). Det er også manglende evidens for at mekaniske faktorer relatert til arbeid hvor holdningsmønster, sitting, løfting, bøying, ståing og manuelt arbeid, kan være en uavhengig årsak til ryggsmertene. Heller ikke inaktivitet og fysisk dekondisjonering er direkte assosiert med NSLBP (Balague et al., 2012). Selv om det ikke kan identifiseres biomedisinske kausale mekanismer på populasjonsnivå, vil likevel fysiske faktorer som holdning og endret motorisk kontroll, dekondisjonering og smerteatferd kunne påvirke symptombildet for den enkelte pasient (O'Sullivan, 2005).

I mangel på tilstrekkelige biomedisinske forklaringsmodeller har den biopsykososiale modellen i økende grad blitt anerkjent i forståelsen av NSLBP. Allerede i 1987 introduserte Waddell (1987) den biopsykososiale modellen. Han anså psykologiske og sosiale faktorer som like viktig som de biologiske i forklaringen av korsryggsmerter og hvordan det påvirker sykdomsforløpet. Et like stort problem som selve opplevelsen av smerte, mente Waddell var knyttet til pasientens og samfunnets oppfatninger og reaksjoner på smertene.

Uhensiktsmessige reaksjoner er for eksempel unngåelse av fysisk aktivitet, arbeid eller i det sosiale liv (Ibid). Det er i dag godt dokumentert at pasienter med NSLBP ofte opplever angst og depresjon samt negative påvirkninger i sosiale sammenhenger, både i arbeidslivet og fritid (Koes et al., 2006, Balague et al., 2012). Inntil nå er det ikke funnet noen klar sammenheng mellom ryggsmerte, funksjonsnedsettelse, degenerative prosesser og spinal mobilitet (Jarvik et al., 2005, Lederman, 2011). I stedet ser man at psykososiale faktorer som depresjon, stress, uhensiktsmessige oppfatninger av smertene, frykt-unngåelsesatferd og arbeidsrelaterte problemer i større grad predikerer funksjonsnedsettelse og kronifisering av ryggplager (Grotle et al., 2006, Linton, 2001, Crook et al., 2002). Samtidig er det kjent at funksjonsnedsettelse er sterkt assosiert med kognitive og atferdsmessige faktorer sammenlignet med biomekaniske og strukturelle faktorer (O'Sullivan, 2012). I dag er det bred enighet om at NSLBP må forstås i lys av en biopsykososial ramme hvor mange årsaksfaktorer interagerer med hverandre og påvirker hvert enkelt individ i varierende grad. Disse inkluderer patoanatomiske, fysiske, nevrofysiologiske, psykologiske, og sosiale faktorer (O'Sullivan, 2005).

Noe av grunnen til den åpenbare behandlingsresistensen av NSLBP antas å skyldes at det representerer en heterogen pasientgruppe, hvor individualisert behandling vanskeliggjøres på grunn av et komplekst årsaksforhold (Balague et al., 2012, O'Sullivan, 2012). Til tross for at den biopsykososiale modellen de siste tiårene har fått en dominerende rolle i konseptualiseringen av NSLBP, ser det ikke ut til man har kommet nærmere noen løsning på problemet. En viktig grunn til dette kan være at intervensjoner sjelden integrerer alle komponentene av den biopsykososiale modellen, samt en manglende evne til å innlemme og anvende den i klinisk praksis (Pincus et al., 2013). Behandlingen er ofte endimensjonalt rettet og tar ikke høyde for alle de faktorer som kan påvirke smertene (Mafi et al., 2013, O'Sullivan, 2012). I tillegg kan det se ut som mye av ryggbehandling fortsatt opprettholdes av en biomedisinsk modell med overdrevent fokus på patoanatomiske årsaksforklaringer (Bagnall,

2010). I behandling av NSLBP er det derfor flere indikasjoner på at en bedre integrering og adressering av de ulike dimensjonene vil være fordelaktig.

## **2.2. Stratifisert behandling**

De store variasjonene i presentasjonen av NSLBP, mangel på en klar diagnose som sier noe om smertemekanismene, en kompleks interaksjon av biopsykososiale faktorer samt manglende effekt av tradisjonelle intervensjoner, har gjort stratifisert behandling av korsryggsmerter til et hovedanliggende innen forskning og klinisk praksis (Waddell, 2005, Karayannis et al., 2012, Balague et al., 2012, O'Sullivan, 2012). Målsetningen er å identifisere sub-grupper av pasienter med felles karakteristikk, hvor behandlingen er målrettet mot aktuelle populasjon. Det vil potensielt kunne hjelpe klinikere i diagnostisering, maksimere behandlingsresultat, redusere bivirkninger og effektivisere helsetjenesten (Foster et al., 2013).

Det er gjort mange forsøk på å utvikle stratifisert behandling for korsryggsmerter, med stor variasjon i filosofi og metoder. Karayannis og medarbeidere (2012) har identifisert en rekke bevegelsesbaserte klassifiseringssystemer. Dette inkluderer klassifisering hvor behandling rettes mot underliggende mekanisk dysfunksjon; behandlingsbasert klassifisering hvor man identifiserer pasientkarakteristikk som kan predikere respons på behandling; patoanatomisk klassifisering hvor patologiske strukturer avdekket av ortopediske tester styrer behandling; identifisering av dysfunksjonelle bevegelser og posturale mønster; samt klassifisering basert på underliggende smertemekanismer. Selv om flere av de bevegelsesbaserte systemene i ulik grad vurderer psykologiske faktorer ved vedvarende korsryggsmerter, er det få som spesifiserer hvordan disse blir vurdert og integrert i behandlingen. De fleste benytter dessuten en biomekanisk eller patoanatomisk forklaring på ryggsmertene, og få tar hensyn til nevrofysiologiske faktorer for å skille mellom ulike smertemekanismer (Karayannis et al., 2012).

For kroniske uspesifikke korsryggsmerter fant Fairbank et al. (2011) fem klassifikasjonssystemer med behandlingstrategier som var rettet mot aktuelle subgrupper. Bare to av systemene hadde gjennomført effektstudier, hvor intervensjon ble sammenlignet med en kontrollgruppe som fikk standard behandling uten klassifisering. Forfatterne konkluderer med at det foreligger utilstrekkelig evidens til å anbefale et klassifiseringssystem for kroniske uspesifikke korsryggsmerter (Ibid). Flere klassifiseringssystemer har også blitt

utarbeidet for å identifisere subgrupper basert på prognose, med stor vekt på psykologiske og sosiale faktorer (Fairbank et al., 2011, Foster et al., 2013). Selv om de adresserer de psykososiale faktorene i større grad enn mange bevegelsesbaserte systemer, mangler de fleste av disse målrettede behandlingsstrategier (Fairbank et al., 2011).

Stratifisert behandling kan samlet inndeles i tre hovedkategorier, som alle har vist lovende bidrag til ryggbehandling. En kategori er klassifisering basert på prognostiske faktorer, som medfører at behandling styres av pasientens risiko for vedvarende plager (Foster et al., 2013). Eksempler er tilnærminger som fokuserer på en enkelt prognostisk faktor, for eksempel frykt for bevegelse (George et al., 2008), eller kombinasjon av flere psykososiale faktorer (Bergstrom et al., 2001). Det har også blitt utviklet en modell som forsøker å inkludere både psykososiale og fysiske prognostiske faktorer (Hill et al., 2011). Klassifisering basert på underliggende mekanismer er en annen kategori (Foster et al., 2013). Her guides behandlingen av hvilke mekanismer som antas å forårsake smertene. Instabilitet (Fritz et al., 2005), smertemekanismer (Scholz et al., 2009) og patoanatomiske årsaker (Petersen T, 2003) er eksempler på mekanisme-baserte systemer. En tredje kategori er klassifisering basert på behandlingsrespons, hvor kliniske funn benyttes for å avgjøre om en pasient vil respondere på en spesifikk behandling (Foster et al., 2013). Dette gjelder for eksempel kliniske prediksjonsregler for behandling som manipulasjon (Childs et al., 2004) og stabiliserende øvelser (Hicks et al., 2005).

På tross av mye forskning på stratifisert behandling det siste tiåret har systematiske oversikter vist manglende evidens for hvilke systemer som bør anbefales (Stanton et al., 2010, Haskins et al., 2012, Foster et al., 2013). Dette skyldes trolig stor variasjon og uenighet i begrunnelsen for hvilke faktorer som bør inkluderes (Haskins et al., 2012), samt få studier av høy kvalitet (Stanton et al., 2010). Spesielt har det blitt argumentert for en manglende integrering av psykososiale faktorer (Ford et al., 2007). Mange klassifiseringssystemer har dessuten hatt et endimensjonalt fokus og mangel på en fullstendig vurdering av korsryggsmerter innen en biopsykososial ramme (Fersum et al., 2010, Dankaerts and O'Sullivan, 2011). De siste årene har det tilkommet flere effektstudier vedrørende stratifisert behandling hvor flere dimensjoner av den biopsykososiale modellen blir inkludert (Hill et al., 2011, Apeldoorn et al., 2012, Vibe Fersum et al., 2013).

### **2.2.1. Vurdering av metodologiske faktorer ved klassifisering**

I utviklingen av klassifiseringssystemer foreligger det store metodologiske variasjoner, med lite konsensus angående hvilke symptomer og tegn som skal karakterisere sub-grupper (Ford et al., 2007, Kent et al., 2010). På grunn av mange potensielle metodologiske fallgruver bør fremstilling, vurdering og rapportering av klassifiseringssystemer tolkes med varsomhet (Kent et al., 2010). Til nå finnes det ikke noe kjent standard metodologisk rammeverk for å utvikle klassifiseringssystemer og identifisere sub-grupper (Kent et al., 2010, Foster et al., 2011).

Det er fremstilt flere metodologiske vurderinger som anses vesentlig i utviklingen av klassifiseringssystemer. For det første bør kravene til en robust modell være oppfylt. Det innebærer at den er plausibel, har distinkte sub-grupper som utelukker hverandre, er reliabel, klinisk nyttig og enkel i bruk. For det andre bør forskningsdesignet og analyser være hensiktsmessig for å studere den aktuelle modellen (Foster et al., 2011). Valg av studiedesign avhenger av om man ønsker å studere generelle prognostiske faktorer eller modifieringseffekter av behandling. Prognostiske faktorer er symptomer og tegn som indikerer sannsynlig utfall hos pasienter uavhengig av behandling. Modifieringseffekter av behandling er symptomer og tegn som identifiserer pasientgrupper som responderer ulikt på en spesifikk behandling (Kent et al., 2010). Optimalt for utforskning av prognostiske faktorer er observasjonsstudier som longitudinelle kohortstudier. For å studere modifieringseffekter av behandling kreves det kontrollerte forsøk med minst to grupper (Foster et al., 2011).

Det er utviklet flere prinsipper for å vurdere validiteten og generaliserbarheten ved studier som forsøker å identifisere sub-grupper av respondere på spesifikke intervensjoner (Hancock et al., 2009). Først og fremst kreves det en randomisert kontrollert studie for å bestemme om en sub-gruppe er assosiert med respons på en intervensjon. Et annet prinsipp er å forhindre multiple sub-gruppe-analyser i håp om å finne en potensiell sub-gruppe ved tilfeldigheter (Ibid). Et tredje prinsipp består i å demonstrere at effekten av en spesifikk behandling for en sub-gruppe er større enn hos pasienter som ikke tilhører den aktuelle sub-gruppen. I undersøkelser av sub-grupper kreves også et betraktelig større utvalg enn å måle vanlig effekt av to-armete intervensjonsstudier (Brookes et al., 2004). Et viktig prinsipp i siste fase av valideringen av stratifisert behandling er å utføre gjentatte eksperimentelle studier på andre populasjoner. Dette gjøres gjennom smale valideringsstudier med tilsvarende pasienter og miljø, noe som vil si noe om variabiliteten av effekter innen en sub-gruppe. Deretter vil brede valideringsstudier, hvor ulike miljø, karakteristikker av behandlere og bredere grupper av

pasienter blir inkludert, kunne skape generaliserbarhet utover den opprinnelige profilen til de som ble studert (Kent et al., 2010). Til sist er impact-analyser indisert for å etablere gjennomførbarhet og forbedre resultater i klinisk praksis. Selv om sub-grupper har vist gode resultater i valideringsstudier vil det ikke nødvendigvis bety at de kan bli effektivt implementert i vanlig klinisk praksis (Ibid).

### **2.3. utfordringer med medisinske uforklarlige symptomer**

NSLBP har blitt ansett for å være et eksempel på såkalt medisinske uforklarlige symptomer (MUS). Felles for disse er fraværet av forklarende patologi. Ingen enkelt psykologisk eller organisk patologi kan identifiseres og forklare symptomene i sin helhet (Eriksen et al., 2013). Klassifisering av slike symptomer har vist seg vanskelig å gjennomføre. Spørsmålet er om problemet kan løses ved å gjøre mer av det samme? Det vil si å bruke de samme vitenskapelige metodene som har gjort det mulig å klassifisere andre sykdommer, slik at en mer detaljert forståelse av delene vil kunne avdekke årsaken til helheten. Tradisjonelt har det vært en forventning om at kausale mekanismer til nå ikke har blitt avdekket men at mer forskning vil kunne gi svar, for eksempel ved å gjøre flere RCT (Ibid). Til tross for flere tiårs iherdig forskning er det fortsatt stor usikkerhet knyttet til de kausale mekanismene bak NSLBP (Balague et al., 2012, O'Sullivan, 2012). Selv om man ikke finner spesifikke årsaksforklaringer på populasjonsnivå vil det i individuelle tilfeller likevel kunne være tydelige indikasjoner på at NSLBP er forårsaket av en spesifikk årsak, for eksempel et løft. Det kan derfor se ut til at populasjonsstudier gir begrenset informasjon om årsakssammenhenger ved NSLBP (Eriksen et al., 2013).

Innføringen av den biopsykosiale modellen har vært et forsøk på å skape en mer helhetlig forståelse av pasienten (Pincus et al., 2013). Selv om modellen inkluderer psykososiale faktorer ser det ut til å foreligge en kropp-sinn dualisme med atskillelse mellom psykologiske og somatiske mekanismer (Butler et al., 2004). Selv om korsryggsmerter sees i lys av en biopsykososial modell, er det fortsatt godt forankret i et vitenskapelig ontologisk rammeverk basert på reduksjonisme, hvor de ulike komponentene er separert og behandles hver for seg. Det impliserer at årsakssammenhenger går nedover og oppover, hvor sosiale fenomener har psykologiske årsaker, som igjen kan forklares av biologiske årsaker. Det er ikke nødvendigvis bevisbart at mekanismene bak NSLBP kan forklares ved å forstå enkeltdelene, selv om man tar høyde for psykologiske og sosiale aspekter (Kerry et al., 2012).

Det kan se ut til at MUS representerer en begrensning ved dagens vitenskapelige paradigme, hvor mange av de forskningsmetoder som benyttes er basert på en reduksjonistisk og mekanistisk tilnærming. Det antas at de ulike delene og dimensjonene av helse kan forstås og behandles som separate prosesser, med ingen direkte interaksjon mellom delene. Kausalitet er knyttet til robuste korrelasjoner, med RCT som anerkjent gullstandard, og en antagelse om at samme årsak vil gi samme effekt. Generell kunnskap basert på statistiske gjennomsnitt vektlegges fremfor individuelle karakteristikk (Kerry et al., 2012). For MUS er individuelle variasjoner det normale og like årsaker/intervensjoner ser ut til å gi vidt forskjellige symptomer/resultater. En kan få ryggsmarter av å løfte en ting mens andre løfter hver dag uten å få det. Populasjonsdata ser derfor ut til å være en begrensning i forklaringen av MUS (Eriksen et al., 2013). Manglende årsaksforklaringer, meningsfulle diagnostiske beskrivelser og målrettede intervensjoner har ført til at MUS representerer en utfordring for vitenskapelig forskning med en mekanistisk og reduksjonistisk tilnærming (Ibid).

#### **2.4. Systematiske oversikter – styrker og svakheter**

En systematisk oversikt har til hensikt å oppsummere all forskning som er gjort innenfor et spesifikt emne. Den syntetiserer resultatene fra primærstudier ved bruk av metoder som reduserer bias og tilfeldige feil. Avhengig om studiene er tilstrekkelig like til å kombinere resultatene, kan systematiske oversikter inkludere en statistisk syntetisering, såkalt meta-analyse (Gopalakrishnan and Ganeshkumar, 2013). Sammenlignet med et litteraturstudie vil en systematisk oversikt benytte mest mulig reproduserbare og verifiserbare prosedyrer. Ved å utvikle en protokoll på forhånd som nøyaktig beskriver utvalget og hvordan datainnhenting skal foregå, vil prosessen bli mer transparent og objektiv. Målsetningen er å komme med konklusjoner som i størst mulig grad er korrekte og ikke er påvirket av en forutinntatt prosess eller utvalg av studier (Polit and Beck, 2012).

Når det gjelder syntetisering av evidens innen helse er systematiske oversikter blitt en referanse standard. De er en viktig del i utviklingen av kliniske retningslinjer og påvirker kliniske beslutninger (Moher et al., 2015). På grunn av et strengt vitenskapelig design basert på forhåndsbestemte kriterier og reproduserbare metoder, kan systematiske oversikter gi reliable vurderinger av effekter av intervensjoner og mer holdbare konklusjoner. Systematiske oversikter kan også identifisere mangel på evidens, og dermed bidra til å guide fremtidig

forskning (Gopalakrishnan and Ganeshkumar, 2013). Andre fordeler med systematiske oversikter er å gi forskere, helsearbeidere og makthaverne lett tilgang til informasjon, forbedre generaliserbarhet av resultater, generere nye forskningshypoteser, og generelt øke presisjon av resultater (Ibid).

Det er likevel påpekt svakheter ved slike studier. Faktorer som vil kunne medføre feil i systematiske oversikter er lokalisering og seleksjon av studier, heterogenitet, mangelfull informasjon på viktige utfallsmål, uhensiktsmessige sub-gruppe analyser og duplisering av publisering. Verdien av systematiske oversikter avhenger av kvaliteten på rapportering. Utilstrekkelig rapportering vil derfor begrense leserens evne til å vurdere styrker og svakheter (Gopalakrishnan and Ganeshkumar, 2013). På grunn av at primærstudier kan inneholde systematiske feil, vil inkludering av slike studier i systematiske oversikter kunne føre til misvisende konklusjoner (Roberts et al., 2015). Publiseringbias er ikke uvanlig, og vil si at studier med positive resultater tenderer til å inkluderes i større grad på grunn av utilstrekkelig sortering av studier (Gopalakrishnan and Ganeshkumar, 2013). Det er anslått at så mye som halvparten av alle eksperimenter er upubliserede. Selv om studier blir identifisert er det også en fare for selektiv rapportering av utfallsmål. Til og med studier utført av The Cochrane Collaboration, som anses for å være av høy kvalitet, kan ofte være forutinntatt. For eksempel har de blitt kritisert for seleksjon av studier og for å ta med for mange lavkvalitetsstudier (Roberts et al., 2015). De fleste systematiske oversikter og meta-analyser består dessuten av få deltakere, noe som kan føre til at man ikke oppdager viktige kliniske behandlingseffekter, eller at man får falske positive resultater. Det er også en fare for at sjekklister og verktøy for risikovurdering av bias ikke er i stand til å oppdage subtile systematiske feil som i økende grad forekommer i industrielt sponsede studier (Greenhalgh et al., 2014). På grunn av at systematiske oversikter integrerer og oppsummerer forskning kan de øke generaliserbarheten sammenlignet med resultater fra et enkelt studie. Det er likevel usikkerhet blant klinikere til hvilken grad resultater kan overføres til individuelle pasienter. Det er derfor foreslått at forfattere av systematiske oversikter bør ha økt fokus på å vurdere faktorer som kan påvirke generaliserbarheten i systematiske oversikter (Nasser et al., 2012).

For å øke gjennomsiktigheten, tilregneligheten og forskernes integritet anbefales det å følge en protokoll som beskriver rasjonale, hypotese og den planlagte metoden for den systematiske oversikten. På den måten sikrer man at studien blir tilstrekkelig planlagt og at planen er dokumentert før datainnsamlingen starter (Moher et al., 2015).



### **3. METODE**

Denne systematiske oversikten følger nylig oppdaterte retningslinjer for systematiske oversikter i Cochrane Back and Neck Group (Furlan et al., 2015), og Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) (Moher et al., 2015).

#### **3.1. Utvalg**

Her gjennomgås kriterier som ligger til grunn for inkluderte studier.

##### **3.1.1. Inklusjonskriterier**

###### **Studie design**

Det vil kun bli inkludert RCT som har brukt en anerkjent randomiserings-prosedyre (Furlan et al., 2015).

###### **Deltakere**

- Voksne som er 18 år eller eldre, med NSLBP med eller uten bensmerter.
- Alle deltakere skal ved baseline klassifiseres i sub-grupper til en matchende intervensjon.

###### **Intervensjon**

Bare studier med et klassifiseringssystem basert på alle dimensjonene i den biopsykososiale modellen, hvor intervensjoner er målrettet mot de aktuelle sub-grupper, inkluderes. Dette innebærer at biologiske, psykologiske og sosiale faktorer både ligger til grunn for identifisering av sub-grupper og adresseres i behandling.

###### **Kontrollgruppe**

Det er ingen begrensinger på kontrollgruppe.

###### **Utfallsmål**

Bare pasient-rapporterte utfallsmål blir inkludert. Effekten blir evaluert på bakgrunn av smerte, diagnose-spesifikk funksjon og livskvalitet.

- Mål på smerte inkluderer ulike smerteskalaer (f.eks. visual analogue scale (VAS), pain numerical rating scale (PNRS) og McGill pain score).

- Mål på funksjon inkluderer spørreskjema for rygg-spesifikk funksjon (f.eks. Roland-Morris disability questionnaire (RMDQ), Oswestry Disability Index (ODI))
- Mål på livskvalitet inkluderer spørreskjema for opplevd helsestatus eller livskvalitet (f.eks. SF-36, EuroQol thermometer)

### **Språk**

Bare engelsk og skandinavisk litteratur blir inkludert.

### **Etiske hensyn**

Bare studier som har beskrevet om de har fulgt Helsinkideklarasjonen blir inkludert

### **3.1.2. Eksklusjonskriterier**

#### **Deltakere**

Følgende lidelser og karakteristikk blir ekskludert:

- Spinale lidelser med alvorlig underliggende patologi (eks: tumor, infeksjoner, systemisk inflammasjon, alvorlige traumer, cauda equina syndrom)
- Spesifikke spinale lidelser på grunn av fraktur, spinal stenose eller radikulopati med nevrologiske utfall.
- Bekkensmerter
- Gravide

### **3.2. Søkemetoder for identifisering av studier**

For å identifisere relevante studier på tvers av fagfelt, samt følge anbefalte retningslinjer har litteratursøk blitt gjennomført i databasene EMBASE, MEDLINE, CINAHL, PEDRO, og PsycINFO (Furlan et al., 2015). Elektroniske søk har blitt utført med assistanse av en erfaren bibliotekar. For å identifisere studiene ble det brukt en kombinasjon av både frie tekstord og emneord. Søkord ble deretter satt sammen ved hjelp av boolske operatører som AND, OR, og NOT til en sensitiv søkestrategi. Som et hjelpemiddel i søkeprosessen ble det anvendt et PICO-skjema (Kirkehei and Ormstad, 2013). Eksempel på søkeord og kombinasjoner som ble brukt er low back pain AND stratified care, low back pain AND classification, low back pain AND subgrouping. I tillegg til det systematiske søket er artikler som er sitert i inkluderte

studier samt relevante systematiske oversikter blitt vurdert (Polit and Beck, 2012). En fullstendig oversikt over søkeord og søkestrategier er beskrevet i Vedlegg 1.

**Figur 1: PICO-skjema for strukturering av potensielle søkeord**

|                     | Norsk                                                                                                | Engelsk                                                                                    |
|---------------------|------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| <b>Population</b>   | ryggsmerter, korsryggsmerter                                                                         | back pain, low back pain                                                                   |
| <b>Intervention</b> | stratifisert behandling, randomisert kontrollert studie, klassifisering, subgruppering, biopsykosial | stratified care, randomized controlled trial, classification, subgrouping, biopsychosocial |
| <b>Comparison</b>   | vanlig behandling                                                                                    | treatment as usual                                                                         |
| <b>Outcome</b>      | smerte, funksjonsnedsettelse, livskvalitet                                                           | pain, disability, quality of life                                                          |

Identifiserte studier ble nedlastet til et elektronisk referanseprogram (EndNote version X7), og duplikater ble fjernet.

### 3.3. Dataproduksjon og analyse

#### 3.3.1. Seleksjon av studier

Utvelgelse av studier ble gjort av førsteforfatter. I første fase av sortering ble tittel og sammendrag anvendt for å møte inklusjons – og eksklusjonskriteriene. Utvalgte studier fra første fase samt studier identifisert ved søk for hånd, ble deretter vurdert for valgbarhet. Fulltekst av identifiserte studier for valgbarhet ble innhentet og undersøkt (Figur 2). Usikkerhet vedrørende endelig inklusjon av studier ble løst i diskusjon med veileder.

#### 3.3.2. Ekstraksjon av data

Ifølge anbefalinger ble det brukt en standardisert form for å trekke ut følgende data fra fulltekst artikler: data som beskriver karakteristikk ved studiet (f.eks. i hvilket land studiet er utført, rekruttering av deltakere) karakteristikk ved deltakere (f.eks. antall deltakere, kjønn, alder), beskrivelse av eksperimentell og kontroll-intervensjon, varighet på oppfølging, utfallsmål og resultater (Furlan et al., 2015).

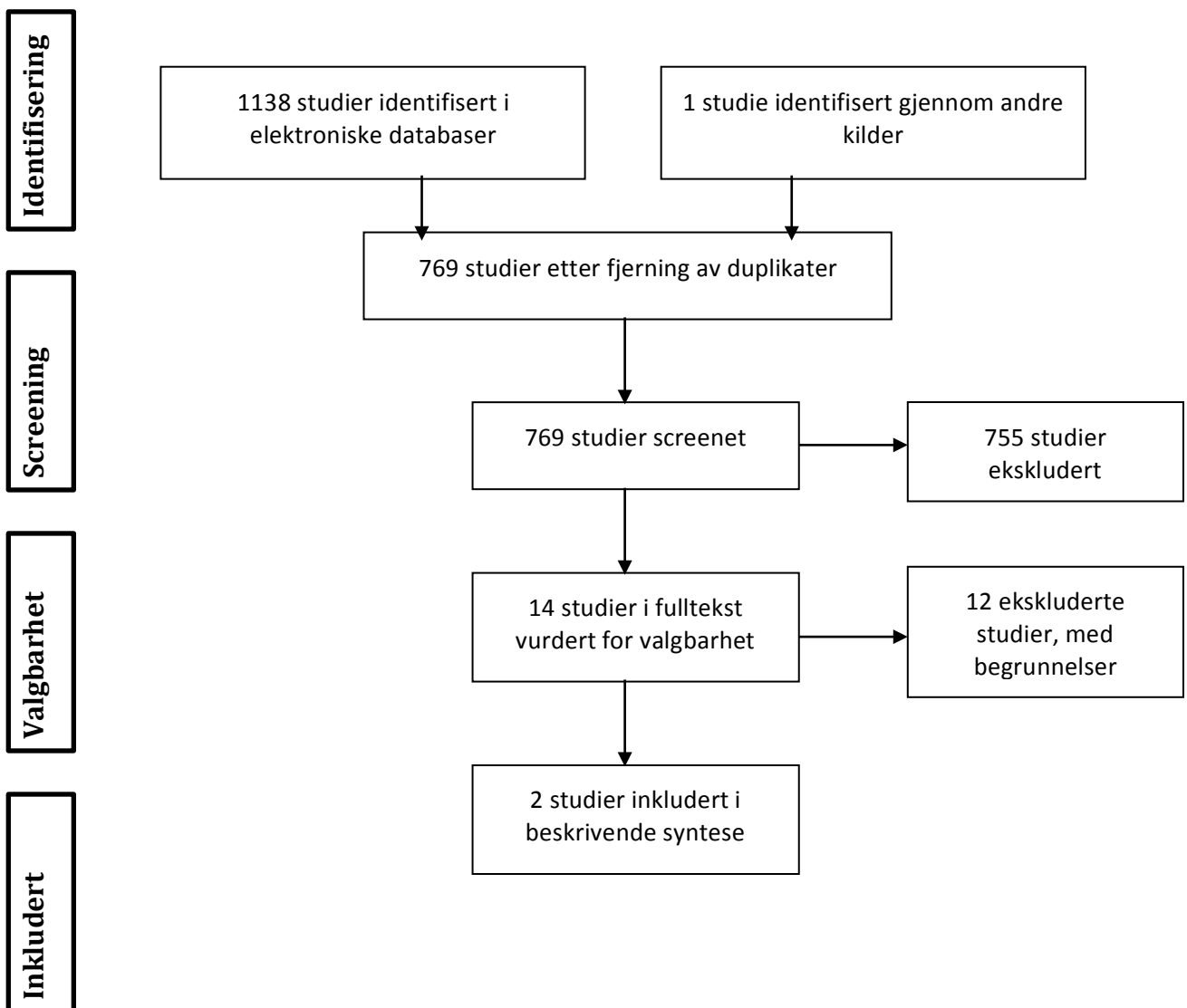
### **3.3.3. Vurdering av metodologisk kvalitet av inkluderte studier**

Vurdering av risiko for systematiske feil ble gjort på bakgrunn av 13 kriterier anbefalt av Cochrane back and Neck Group (Furlan et al., 2015). Usikkerheter ble løst i diskusjon med veileder. Den metodologiske kvaliteten på studiene ble deretter diskutert og vurdert ut i fra helhetsinntrykket av systematiske feil, da det ikke anbefales å bruke cut-off verdier for å bestemme kvaliteten (Ibid).

## 4. RESULTATER

### 4.1. Seleksjon av studier

1138 studier ble funnet ved systematisk søk. I tillegg ble 1 studie identifisert via gjennomgang av referanser i relevante studier og systematiske oversikter om emnet. Etter at duplikater ble fjernet, ble 769 studier screenet med hensyn til inklusjons- og eksklusjonskriterier på bakgrunn av tittel og sammendrag. 14 studier ble vurdert som randomiserte kontrollerte studier av klassifiseringssystemer og ble videre vurdert for valgbarhet på bakgrunn av artikkel i fulltekst. Av disse ble 2 studier inkludert i det endelige utvalget.



Figur 2: Flyttdiagram. Oppsummering av seleksjonsprosessen.

## **4.2. Ekskluderte studier**

Etter første screening ble 755 studier ekskludert, på grunn av manglende klassifiseringssystem, populasjon eller adekvat design. 12 studier ble deretter ekskludert etter vurdering av valgbarhet. Hovedårsaken til eksklusjon av RCT var utilstrekkelig klassifiseringssystem. En studie manglet klassifiseringssystem og så kun på sub-gruppeeffekter av multidisiplinær behandling (Vollenbroek-Hutten et al., 2004). De fleste klassifiseringssystemene manglet en fullstendig biopsykososial vurdering med adressering av psykologiske og sosiale faktorer. Dette gjaldt behandlingsbasert klassifisering basert på Delittos klassifiseringssystem (Fritz, 1998, Fritz et al., 2003, Brennan et al., 2006, Apeldoorn et al., 2012), klassifisering basert på McKenzies Mechanical Diagnosis and Therapy (MDT) (Petersen et al., 2002, Miller et al., 2005, Petersen et al., 2011, Hallidy et al., 2015), og klassifisering basert på Movement System Impairment (MSI) (Henry et al., 2014). Et annet klassifiseringssystem inkluderte psykososiale faktorer i vurderingen av sub-grupper, men behandling var kun rettet mot sub-grupper med posturale og biomekaniske dysfunksjoner (Sheeran et al., 2013). Et klassifiseringssystem baserte sub-grupper på patoanatomiske kausale mekanismer (Chan et al., 2013). Selv om det inkluderte en multifaktoriell sub-gruppe med potensielle underliggende psykososiale eller nevrofysiologiske mekanismer, er denne kun basert på fravær av en tydelig patoanatomisk klassifisering. De inkluderte dessuten en sub-gruppe med radikulopater (Ibid).

## **4.3. Karakteristikker av inkluderte studier**

To studier ble inkludert. En studie ble utført i England og publisert i 2011 (Hill et al., 2011), mens den andre fant sted i Norge og ble publisert i 2013 (Vibe Fersum et al., 2013). Begge var randomiserte kontrollerte studier som undersøkte effekten av multidimensjonal stratifisert behandling. Hill og medarbeidere (2011) undersøkte effekten av Start Back stratifiseringsmodell (SB), et prognosebasert system som klassifiserte deltakerne i enten lav, middels eller høy risiko for vedvarende plager. Behandlingen var rettet mot aktuelle sub-gruppe, hvor lav-risiko pasienter fikk kun en konsultasjon med råd og undervisning mens medium og høy-risiko pasienter i tillegg fikk standardisert fysioterapi rettet mot fysiske symptomer og funksjon. Høy-risiko pasienter fikk fysioterapi med psykologisk tilnærming for å adressere psykologiske faktorer i tillegg til fysiske symptomer og funksjon (Ibid). Vibe Fersum og medarbeidere (2013) undersøkte effekten av Cognitive functional therapy (CFT). Det er et mekanismebasert klassifiseringssystem hvor patoanatomiske, nevrofysiologiske,

psykososiale, fysiske og livstils – faktorer integreres i klassifiseringen. På bakgrunn av klinisk undersøkelse kombinert med psykososial screening og eventuelle radiologiske funn, ble pasienter klassifisert i en av fem sub-grupper. En gruppe hvor smerter og bevegelsesforstyrrelser er sekundære til underliggende alvorlig patologi eller spesifikke korsryggsmerter, en gruppe hvor hovedsakelig psykologiske og/eller sosiale faktorer driver smertene, og en stor gruppe hvor forstyrret bevegelsesmønster anses som underliggende smertemekanisme. Denne siste gruppen ble videre klassifisert i to sub-grupper karakterisert med henholdsvis redusert bevegelse og unngåelsesatferd, eller svekket bevegelseskontroll og uhensiktsmessig provokasjon av smerter (O'Sullivan, 2005). Avhengig av klassifisering mottok hver pasient en målrettet intervensjon rettet mot å endre uhensiktsmessige atferd relatert til kognisjoner, bevegelsesmønster og livsstil (Vibe Fersum et al., 2013).

Til sammen inkluderte studiene 972 deltakere av begge kjønn. Det var i midlertidig en stor forskjell i utvalgsstørrelse, med omtrent 7 ganger så mange deltakere i studien til Hill og medarbeidere (2011). CFT inkluderte kun pasienter med kroniske NSLBP, mens SB inkluderte korsryggsmerter med eller uten bensmerter og uavhengig av varighet. Begge studiene målte smerte ved bruk av en numerisk rating skala fra 0-10 hvor 0 tilsvarte ingen smerte og 10 tilsvarte verst tenkelige smerte. Vibe Fersum og medarbeidere (2013) målte funksjon med Oswestry Disability Questionnaire (ODI) mens Hill og medarbeidere (2011) brukte Roland and Morris Disability Questionnaire (RMDQ). Helserelatert livskvalitet ble målt med spørreskjemaene EuroQol EQ-5D og Short Form 12 i studien med SB, mens i studien med CFT hadde de ingen primære mål på livskvalitet. Oppfølging av utfallsmål var omtrent like i begge studier med målinger etter intervensjonen og 12 måneder etter. Detaljert beskrivelse av karakteristikk for inkluderte studier finnes i Tabell 1 og mer detaljert i Vedlegg 2 og 3.

**Tabell 1: Oversikt over karakteristikk av inkluderte studier**

| Forfatter / land                  | Design | Setting                                                                                                                          | Deltakere                                                                         | Intervensjons-gruppe                                                                                                           | Kontrollgruppe            | Resultater                                                                                                            |
|-----------------------------------|--------|----------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Hill et al (2011)<br>England      | RCT    | Primærhelsetjenesten: 10 almenlegepraksiser i Keele General Practice Research Partnership                                        | 851 pasienter med NSLBP <sup>1</sup> , gjennomsnittlig alder 50 år, 59 % kvinner  | SB <sup>3</sup> : klassifisering i enten lav, middels eller høy risiko, med målrettet behandling                               | Standard fysioterapi      | SB <sup>3</sup> viste signifikant bedring i funksjon og livskvalitet ved 4 mnd og 12 mnd, og for smerte kun ved 4 mnd |
| Vibe Fersum et al (2013)<br>Norge | RCT    | Primær og sekundær helsetjenesten: Private fysioterapiklinikk i Bergen og en poliklinisk avdeling for rygg ved Haukeland sykehus | 121 pasienter med NSCLBP <sup>2</sup> , gjennomsnittlig alder 42 år, 51 % kvinner | CFT <sup>4</sup> : klassifisering til en av fem sub-grupper basert på underliggende smertemekanismer, med målrettet behandling | Manuell terapi og øvelser | CFT <sup>4</sup> viste signifikant bedring for smerter og funksjon ved 3 mnd og 12 mnd                                |

<sup>1</sup>Uspesifikke korsrygg smerter, <sup>2</sup>Uspesifikke kroniske korsrygg smerter, <sup>3</sup>Start Back stratifiseringsmodell (SB), <sup>4</sup>Cognitive functional therapy



#### 4.4. Metodologisk kvalitet av inkluderte studier

Samlet risiko for systematiske feil vurderes å være lav for inkluderte studier for alle utfallsmål. Et studie hadde beskjedent høyere kvalitet og inkluderte intention-to-treat analyse (Hill et al., 2011). Bortsett fra en høyere frafallsrisiko i Vibe Fersum og kolleger (2013) scoret begge studiene likt på de ulike kriteriene. Fullstendig oversikt over vurdering av risiko for systematiske feil er presentert i tabell 2 og mer detaljert i Vedlegg 2 og 3.

**Tabell 2: Oversikt over risiko for systematiske feil. Vurdering av hver enkelt feilkilde for inkluderte studier.**

| Systematiske feil / Forfatter                                     | Hill et al 2011 | Vibe Fersum et al 2013 |
|-------------------------------------------------------------------|-----------------|------------------------|
| 1. Adekvat randomiseringsmetode?                                  | +               | +                      |
| 2. Skjult tildeling av intervensjonsgruppe? Alle utfallsmål       | +               | +                      |
| 3. Blinding av pasient? Alle utfallsmål                           | -               | -                      |
| 4. Blinding av helse tilbyder? Alle utfallsmål                    | -               | -                      |
| 5. Blinding av assessor? Alle utfallsmål?                         | +               | +                      |
| 6. Akseptabel frafallsprosent? Alle utfallsmål                    | +               | +                      |
| 7. Alle deltakere analysert - intention-to-treat? Alle utfallsmål | +               | -                      |
| 8. Fri for selektiv rapportering av utfallsmål?                   | +               | +                      |
| 9. Tilstrekkelig likhet av baseline karakteristikker?             | +               | +                      |
| 10. Var ko-intervensjon unngått eller lik?                        | +               | +                      |
| 11. Var compliance akseptabel for alle grupper?                   | +               | +                      |
| 12. Likt tidspunkt for vurdering av utfallsmål for alle grupper?  | +               | +                      |
| 13. Er andre kilder til bias usannsynlig?                         | +               | +                      |
| Total risiko for systematiske feil (-/total)                      | 2/13            | 3/13                   |

+ = Ja, - = Nei, ? = Usikker.

#### 4.5. Effekt av intervensjoner

På grunn av at bare Hill og medarbeidere (2011) har analysert og sammenlignet resultatene for de enkelte sub-gruppene, er det i dette studiet kun benyttet resultater for

intervensjonsgruppen som helhet som grunnlag for vurdering av effekt. I Hill og medarbeidere (2011) ble smerte, funksjon og livskvalitet målt, men kun funksjon som primærutfallsmål. I Vibe Fersum og medarbeidere (2013) ble både smerte og funksjon brukt som primærutfallsmål, mens livskvalitet ble ikke målt.

Begge studiene målte smerte med en 11-punkts numerisk smerteskala (PNRS). Resultatene fra Hill og medarbeidere (Hill et al., 2011) viste en statistisk signifikant effekt for intervensjonsgruppen ved 4 måneder men ikke ved 12 måneder. Det ble ikke rapportert effektstørrelser eller deltakere med klinisk meningsfulle endringer. Vibe Fersum og medarbeidere (2013) fant signifikant reduksjon i smerte for intervensjonsgruppen sammenlignet med kontrollgruppen både ved 3 og 12 måneders oppfølging. I tillegg ble det rapportert endringer i PNRS score fra baseline. Ved 3 måneder hadde 76 % (38) av deltakerne i intervensjonsgruppen mer enn 2 poengs bedring i PNRS sammenlignet med 45,2 % (19) i kontrollgruppen, og ved 12 måneder 73,5 % (36) i intervensjonsgruppen sammenlignet med 30,7 % (12) i kontrollgruppen. En gjennomsnittlig forbedring i PNRS score var på 3,2 (95% CI: 2,5 – 3,9;  $p < 0,001$ ) for intervensjonsgruppen ved 12 måneders oppfølging, og 1,5 (95% CI: 0,7 – 2,2;  $p < 0,001$ ) for kontrollgruppen.

**Tabell 3: Resultater for smerte som utfallsmål.**

| Forfatter              | Utfallsmål        | Tidspunkt for måling | Stratifisert behandling Gj.snitt (SD) | Kontroll Gj.snitt (SD) | Gjennomsnittsforskjell (95 % CI) | P-verdi |
|------------------------|-------------------|----------------------|---------------------------------------|------------------------|----------------------------------|---------|
| Hill et al 2011        | PNRS <sup>1</sup> | 4 mnd                | 3,2 (2,5)                             | 2,6 (2,4)              | 0,55 (0,23 til 0,86)             | 0,0005  |
|                        | PNRS <sup>1</sup> | 12 mnd               | 3 (2,8)                               | 2,8 (2,6)              | 0,11 (0,21 til 0,43)             | 0,5002  |
| Vibe Fersum et al 2013 | PNRS <sup>1</sup> | 3 mnd                | 1,7 (1,7)                             | 3,8 (1,9)              | -2,1 (-2,7 til -1,4)             | <0,001  |
|                        | PNRS <sup>1</sup> | 12 mnd               | 2,3 (2)                               | 3,8 (2,1)              | -1,3 (-2,1 til -0,5)             | <0,001  |

<sup>1</sup>11-punkts numerisk smerteskala, 0 = ingen smerte, 10 = verst tenkelige smerte.

<sup>a</sup>Hill et al (2011) bruker smerte som sekundærutfallsmål, positive verdier for gjennomsnittsforskjell favoriserer stratifisert behandling.

<sup>b</sup>Vibe Fersum et al (2013) bruker smerte som primærutfallsmål, negative verdier for gjennomsnittsforskjell favoriserer stratifisert behandling.

Funksjon, målt med Roland Morris Disability Questionnaire (RMDQ) i Hill og medarbeidere (2011), viste samlet statistisk signifikant effekt for intervensjonsgruppen sammenlignet med kontrollgruppen. Dette gjaldt også for konservative sensitivitetsanalyser ved bruk av ikke-imputerte datasett. Gjennomsnittsforskjellen mellom intervensjon – og kontrollgruppen var statistisk signifikant både ved 4 og 12 måneders oppfølging. Dette tilsvarte effektstørrelser på henholdsvis 0,32 og 0,19. Deltakere med godt resultat ble definert som minst 30 % endring i RMDQ score fra baseline, og var ved 4 måneder 69 % (392) for intervensjonsgruppen og 56 % (158) for kontrollgruppen. Deltakere med godt resultat ved 12 måneder var 65 % (367) for intervensjonsgruppen og 57 % (160) for kontrollgruppen. Numbers needed to treat ble også rapportert og var henholdsvis 7 og 10,8 ved 4 og 12 måneder.

Vibe Fersum og medarbeidere (2013) målte funksjon med Oswestry Disability Index (ODI). Intervensjonsgruppen viste statistisk signifikant bedring sammenlignet med kontrollgruppen, både umiddelbart etter intervensjon ved 3 måneder og ved 12 måneders oppfølging. I tillegg ble det rapportert endringer i ODI score fra baseline. Ved 3 måneder viste 72 % (36) av deltakerne i intervensjonsgruppen en bedring på over 10 poeng sammenlignet med 33,4 % (14) i kontrollgruppen, og ved 12 måneder 63,3 % (31) i intervensjonsgruppen og 30,7 % (12) i kontrollgruppen. Den gjennomsnittlige forbedringen i ODI score var på 13,7 poeng (95% CI: 11,4 – 16,1;  $p < 0,001$ ) for intervensjonsgruppen og 5,5 poeng (95% CI: 2,8 – 8,3;  $p < 0,001$ ) for kontrollgruppen ved 12 måneder.

**Tabell 4: Resultater for funksjon som utfallsmål.**

| Forfatter              | Utfallsmål        | Tidspunkt for måling | Stratifisert behandling Gj.snitt (SD) | Kontroll Gj.snitt (SD) | Gjennomsnittsforskjell (95 % CI) <sup>a</sup> | P-verdi |
|------------------------|-------------------|----------------------|---------------------------------------|------------------------|-----------------------------------------------|---------|
| Hill et al 2011        | RMDQ <sup>1</sup> | 4 mnd                | 4,7 (5,9)                             | 3 (5,9)                | 1,81 (1,06 til 2,57)                          | <0,0001 |
|                        | RMDQ              | 12 mnd               | 4,3 (6,4)                             | 3,3 (6,2)              | 1,06 (0,25 til 1,86)                          | 0,0095  |
| Vibe Fersum et al 2013 | ODI <sup>2</sup>  | 3 mnd                | 7,6 (6,7)                             | 18,5 (8,1)             | -9,7 (-12,7 til -6,7)                         | <0,001  |
|                        | ODI               | 12 mnd               | 9,9 (9,8)                             | 19,7 (11,7)            | -8,2 (-12,6 til -3,8)                         | <0,001  |

<sup>1</sup>Roland Morris Disability Questionnaire (0-24), høy score indikerer alvorlig funksjonsnedsettelse.

<sup>2</sup>Oswestry Disability Index (0-100), høy score indikerer alvorlig funksjonsnedsettelse.

<sup>a</sup>Hill et al (2011): positive verdier favoriserer stratifisert behandling. Vibe Fersum et al (2013): negative verdier favoriserer stratifisert behandling.

Bare Hill og medarbeidere (2011) målte helserelatert livskvalitet. Stratifisert behandling resulterte i større gjennomsnittlig helseeffekter målt med EuroQol-5D. For 12-item Short Form Health Survey (SF-12) var det en statistisk signifikant forskjell mellom gruppene for den fysiske komponenten men ikke den mentale.

## 5. DISKUSJON

Hensikten med studien var å systematisk undersøke litteraturen for å vurdere effekten av multidimensjonal stratifisert behandling for NSLBP.

### 5.1. Oppsummering av resultater

To studier oppfylte inklusjonskriteriene og ble inkludert i denne systematiske oversikten (Hill et al., 2011, Vibe Fersum et al., 2013). De benyttet ulike klassifiseringssystem basert på ulike teorier; et prognosebasert klassifiseringssystem og et system basert på underliggende smertemekanismer. Kontrollgruppen var også forskjellig for de to studiene, med henholdsvis standard fysioterapi (Hill et al., 2011) og manuell terapi og øvelser (Vibe Fersum et al., 2013). I tillegg skilte studiene seg fra hverandre med hensyn til utvalg og utvalgsstørrelse. Begge studiene fant positive effekter av stratifisert behandling for NSLBP. Hill og medarbeidere (2011) viste at stratifisert behandling basert på prognostiske kriterier ga signifikant bedring i funksjon og helse relatert livskvalitet sammenlignet med standard fysioterapi, ved oppfølging etter 4 og 12 måneder. For smerte var det derimot kun signifikant bedring ved 4 måneders oppfølging. I den andre studien (Vibe Fersum et al., 2013) fant de signifikante effekter av stratifisert behandling basert på underliggende smertemekanismer sammenlignet med manuell terapi og øvelser, for både smerte og funksjon ved 3 og 12 måneders oppfølging. I dette studiet ble ikke helse relatert livskvalitet målt spesifikt.

Selv om begge studier fant signifikante effekter kan det stilles spørsmål om dette er klinisk relevante effekter og overførbart til pasienter i klinisk praksis. For primærutfallsmålet funksjon rapporterte Hill og medarbeidere (2011) relativt små effektstørrelser ved slutten av behandling (0,32) og enda mindre ved 12 måneders oppfølging (0,19). En klinisk betydningsfull endring er vist å tilsvare en reduksjon på minst 30 % i RMDQ score (Jordan et al., 2006). Basert på numbers needed to treat vil det si at omtrent for hver ellefte pasient som gis stratifisert behandling, er det kun en pasient som oppnår en klinisk betydningsfull bedring i funksjon ett år etter behandling. Den klinisk relevante effekten må derfor anses som relativt beskjeden. Sammenlignet med kontrollgruppen fant Vibe Fersum og medarbeidere (2013) gjennomsnittlig forbedringer i funksjon og smerte for intervensjonsgruppen som de anser som klinisk relevante effekter. Dette er basert på evaluering av utfallsmålene hvor endring i ODI på mer enn 10 poeng og PNRS på mer enn 2 poeng tilsvarer en klinisk betydningsfull endring (Ostelo et al., 2008). Resultatene viser at for både smerte og funksjon var det betydelig flere

med klinisk betydningsfulle endringer i intervensjonsgruppen sammenlignet med kontrollgruppen både ved kort- og langtids oppfølging. Stratifisert behandling med CFT viser dermed betydelig større og klinisk viktige effekter sammenlignet med SB.

Utvalgsstørrelsen var betydelig større i Hill og medarbeidere (2011). Begge studiene må likevel anses å ha relativt store utvalgsstørrelser sammenlignet med hva som er vanlig for to-armete intervensjonsstudier for korsryggsmerter (Henschke et al., 2012). Ved analyse av sub-grupper innen en arm kreves det flere ganger så stort utvalg som ved sammenligning av to enkeltintervensjoner (Brookes et al., 2004). Dette kan forklare den store forskjellen i utvalgsstørrelser, da Hill og medarbeidere (2011) i tillegg til å måle total effekt av stratifisert behandling foretok sub-gruppeanalyser.

En faktor som vil kunne påvirke tolkningen av resultatene er forskjeller i utvalg. CFT inkluderer kun langvarige korsryggsmerter, og kan derfor ikke si noe om hvordan klassifiseringssystemet håndterer akutte og subakutte korsryggsmerter. På den andre siden vil inkludering av akutte og subakutte korsryggsmerter i større grad kunne medføre effekter i form av naturlig forløp og regresjon til middelveiden, som ikke nødvendigvis kan attribueres til behandlingen (Menezes Costa et al., 2012). Deltakerne i CFT hadde mindre grad av funksjonstap ved baseline sammenlignet med deltakerne i SB. Selv om større grad av funksjonstap kan være prediktivt for vedvarende funksjonstap (Wilkens et al., 2013), er det også vist at større grad av funksjonstap ved baseline ga større klinisk betydningsfulle forbedringer hos pasienter som fikk multidisiplinær behandling, selv om forskjellene var minimale (Verkerk et al., 2013). På grunn av motstridende evidens vedrørende prognostiske faktorer for langvarige korsryggsmerter, er det usikkert hvilken betydning graden av funksjonstap og smerteintensitet har hatt på effekten av behandling (Verkerk et al., 2012). For å bedre kunne generalisere resultatene bør fremtidige studier derfor inkludere populasjoner med ulik grad av funksjonstap og varighet.

Studiene har benyttet ulike kontrollgrupper, noe som potensielt kan påvirke resultatene. Standard fysioterapibehandling og manuell terapi og øvelser gjenspeiler begge vanlig klinisk praksis, og baseres på kliniske retningslinjer og beste tilgjengelige evidens. De anses derfor for å være relativt like, og sammenlignbare med hensyn til vurdering av resultatene.

## 5.2. Metodologisk kvalitet av inkluderte studier

Begge studiene er vurdert til å ha lav risiko for systematiske feil. Studiekvaliteten i Hill og medarbeidere (2011) var i midlertidig noe høyere hovedsakelig fordi det ble inkludert intention-to-treat analyse som sikrer at alle klassifiserte deltakere ble analysert uavhengig av frafall. Vibe Fersum og medarbeidere (2013) foretok ikke intention-to-treat analyse. I samme studie var det også betydelig flere som ikke startet behandling i kontrollgruppen (13,5 %) sammenlignet med intervensjonsgruppen (1,6 %). Selv om intervensjons- og kontrollgruppen har en akseptabel og sammenlignbar frafallsprosent kan det medføre en ytterligere begrensning ved studiet. Forfatterne påpeker at 7 av de som ikke startet behandling i kontrollgruppen hadde negativ erfaring med manuell terapi og derfor kunne økt sannsynligheten for et dårligere resultat i kontrollgruppen. I tillegg ble det ikke funnet statistiske forskjeller i baseline karakteristikk mellom de som fullførte og de som ikke fullførte intervensjonen. Det er derfor lite sannsynlig at de estimerte behandlingseffektene er påvirket av forskjellene i frafall (Vibe Fersum et al., 2013).

Risiko for systematiske feil vedrørende blinding er vurdert som likt for begge studiene. På grunn av at begge studiene benytter fysioterapi som intervensjoner med interaksjon mellom terapeut og pasient samt aktiv deltakelse av pasienten, er verken helsetilbyder eller pasient blindet for intervensjonen. Blinding av pasienter anses å være spesielt viktig når utfallsmål er subjektive og rapportert av pasienten (Day and Altman, 2000). Mangel på blinding i intervensjonsstudier vil kunne føre til at pasienter som er klar over at de er i en intervensjonsgruppe forventer større effekter og dermed oppnår bedre resultater enn kontrollgruppen (Feys et al., 2014). Karakteristikk av de stratifiserte intervensjonene er i stor grad sammenlignbare med de ikke-stratifiserte og det anses derfor som lite sannsynlig at deltakerne visste hvilken intervensjonsgruppe de tilhørte. Basert på kriteriene for vurdering av systematiske feil, bør assessor av utfallsmål betraktes som ikke-blindet når utfallsmålene er selvrapporterte og hvor pasienter potensielt kan vite hvilken behandlingsgruppe de tilhører (Furlan et al., 2015). Begge studiene er derfor vurdert til å ha blindet assessor av utfallsmålene, selv om kun selvrapporterte utfallsmål er inkludert i denne oversikten.

For de resterende kriteriene av risiko for systematiske feil scorer begge studiene likt og ingen påfallende årsaker til svakheter er observert.

### 5.3. Styrker og svakheter ved de to klassifiseringssystemene

NSLBP representerer en heterogen sykdomsgruppe hvor kjente årsaks mekanismer ikke har latt seg avdekke og adressering av patoanatomiske faktorer ikke har bidratt til mer effektive behandlinger (Balague et al., 2012, O'Sullivan, 2012). Stratifisert behandling hvor sub-grupper av NSLBP identifiseres og gis målrettet behandling er av mange ansett som den fremtidige løsningen på problemet. Den biopsykososiale modellen har i stadig større grad fått fotfeste i forståelsen av korsryggsmerter og anbefales å implementere i klassifiseringssystemer (Foster and Delitto, 2011, Pincus et al., 2013, Ford and Hahne, 2013). Multidimensjonal stratifisert behandling viser i denne systematiske oversikten gode resultater sammenlignet med standard behandling. Det er i midlertidig viktige forskjeller på de to systemene som sannsynligvis vil ha implikasjoner for klinisk praksis og pasientutkomme.

Start Back stratifiseringsmodell (SB), består av klassifiseringsverktøyet Startback Screening Tool og matchende fysioterapibehandling. Klassifiseringsverktøyet er et kort pasientrapportert spørreskjema som klassifiserer alle med uspesifikke korsryggsmerter, basert på prognostiske faktorer som potensielt kan endres med behandling. Det består av 9 spørsmål som vurderer fysiske (bensmerter, komorbid smerte, funksjonsnedsettelse) og psykososiale faktorer (plagsomhet, katastrofetanker, frykt, angst og depresjon). Hensikten er å assistere klinikerens beslutningstaking vedrørende behandlingstiltak, og identifisere pasienter med potensielle modifiserbare prognostiske faktorer (Hill et al., 2008). Det er godt dokumentert at slike faktorer er av stor betydning for pasientens prognose og kan derfor anses som viktige i et prognosebasert klassifiseringssystem (Hayden et al., 2010, Campbell et al., 2013). I flere studier har klassifiseringsverktøyet vist seg å predikere prognose ved korsryggsmerter. I tillegg er det vist å være fordelaktig sammenlignet med unidimensjonale screeningverktøy for depressive symptomer, funksjonsnedsettelse, frykt-unngåelsesatferd og katastrofetanker (Wideman et al., 2012, Beneciuk et al., 2015, Page et al., 2015). I tillegg til bruk i primærhelsetjenesten, har klassifiseringsverktøyet også vist god validitet i andre settinger som fysioterapipraksiser (Beneciuk et al., 2015), kiropraktorklinikker (Kongsted et al., 2011) og spesialisthelsetjenesten (Morso et al., 2011). Den prediktive evnen til verktøyet kan i midlertidig være redusert i behandlingssituasjoner hvor enkelte modifiserbare prognostiske faktorer er godt ivaretatt eller brukt i andre settinger. Bruk av klassifiseringsverktøyet i en kiropraktorsetting indikerte at det var små og ikke signifikante forskjeller i prognose mellom de ulike sub-gruppene som alle fikk kiropraktorbehandling (Field and Newell, 2012). Dette kan skyldes at pasienter som oppsøker kiropraktorbehandling har mindre grad av negative



psykologiske faktorer, som dermed kan påvirke prognosen. Forfatterne spekulerer også i at kiropraktorbehandlingen kan ha ivare tatt psykologiske faktorer for aktuelle pasienter og dermed redusert den prediktive evnen til klassifiseringsverktøyet (Ibid). Det har også blitt gjort et impact-studie hvor SB har blitt sammenlignet med standard behandling blant fastleger i primærhelsetjenesten (Foster et al., 2014). Studiet viser at stratifisert behandling gir moderate forbedringer i utfallsmål, er kostnadseffektivt og bekrefter at modellen er gjennomførbar i primærhelsetjenesten. SB er tilrettelagt for primærhelsetjenesten og gir dermed mulighet for tidlig intervensjon, endring av kurs for pasienter med dårlig prognose samt redusere kostnader med hensyn til overbehandling (Robinson and George, 2012). En stor fordel er derfor å luke ut pasienter med lav risiko som ikke trenger videre behandling (Hill et al., 2011).

Utviklingen av SB er basert på identifisering og validering av viktige prognostiske faktorer for vedvarende korsryggsmerter ved å bruke data fra to epidemiologiske kohortstudier fra primærhelsetjenesten (Dunn and Croft, 2005, Foster et al., 2008) og en RCT (Hill et al., 2008). De kombinerte statistiske og kliniske metoder, noe som anses som styrker (Ibid). Selv om den matchede behandlingen er utviklet på bakgrunn av ekspertkonsensus og gjeldende ryggforskning foreligger det mindre omfattende arbeid i utarbeidelse og validering av den sammenlignet med klassifiseringsverktøyet. Behandlingsdelen kan derfor anses å være på et tidligere stadium i utvikling og validering (Hay et al., 2008, Main et al., 2012).

Cognitive functional therapy (CFT) forsøker å kartlegge den underliggende smertemekanismen ved å integrere patoanatomiske, nevrofysiologiske, psykososiale, fysiske og livstils – faktorer i klassifisering av pasienter med NSLBP (Vibe Fersum et al., 2013). I de ulike sub-gruppene integreres behandling av uhensiktsmessige bevegelsesmønstre med kognitiv tilnærming til psykologiske og sosiale begrensninger. På den måten forsøker CFT å inkludere mange ulike aspekter i behandlingen av den enkelte pasient, som hver for seg har dokumentert gevinster for pasienter med korsryggsmerter. Dette gjelder trening av bevegelsesmønstre og motorisk kontroll, atferdsbehandling med kognitive tilnærminger, smerteundervisning og adressering av psykososiale risikofaktorer (Ibid). I tillegg er det godt dokumentert at kommunikasjon, delt beslutningstaking og en pasientsentrert tilnærming er viktig for et godt pasient-terapeut forhold, kan øke compliance og effekt av behandling (Linton, 2005, O'Sullivan, 2012, Franklin et al., 2016).

CFT er utviklet på bakgrunn av tidligere studier som har vist bred støtte for klassifisering av de ulike sub-gruppene. Smerter karakterisert av mekanisk provokative bevegelser og smerter uten mekaniske karakteristikk er vist å ha ulike nevrofysiologiske mekanismer (Smart et al., 2011). Bruk av et validert psykososialt screeningverktøy for å identifisere de med høy psykologisk risiko har vist seg nyttig med hensyn til målrettet behandling (Boersma and Linton, 2005). I tillegg er mekaniske korsryggsmerter assosiert med ulike typer holdningsmønster, bevegelsesmønster og motoriske responser som kan bidra til å differensiere mellom ulike sub-grupper (Dankaerts et al., 2009). CFT har dessuten vist god inter-tester reliabilitet i klassifiseringen av sub-grupper, med identifisering av provoserende retningsbestemte bevegelsesmønster og assosierte psykososiale faktorer for pasienter med NSLBP (Dankaerts et al., 2006, Vibe Fersum et al., 2009).

En stor fordel med klassifiseringsverktøy til SB er at det er tidseffektivt og enkelt å bruke. Kun et selvrapporteringskjema med ni spørsmål fylles ut av pasienten og scores umiddelbart (Hill et al., 2008). Klassifiseringen i CFT er mer omfattende og baseres på en fullstendig klinisk undersøkelse (Vibe Fersum et al., 2013). Det er også viktige forskjeller i krav til utdanning og ferdigheter ved de to klassifiseringssystemene. I lav-risikogruppen til SB gis det ingen behandling utover første konsultasjon. Behandling av sub-gruppene medium- og høy risiko omfatter i tillegg evidensbasert fysioterapi hvor fysiske dysfunksjoner adresseres samt edukasjon og anbefalinger i samsvar med kliniske retningslinjer. Kun høy-risikogruppen får ivaretatt psykososiale risikofaktorer i behandlingen (Hay et al., 2008), og omfatter totalt 6 dager opplæring for terapeuter i tillegg til noe praksis (Main et al., 2012). For å benytte CFT av fysioterapeuter kreves det gjennomsnittlig 100 timer med opplæring (Vibe Fersum et al., 2009). Sammenlignet med SB stiller derfor CFT vesentlig større krav til utdanning av terapeuter. I tillegg krever CFT i større grad ferdigheter over et bredt spekter av domener, noe som kan være en barriere for implementering i klinisk praksis. Evne til effektiv kommunikasjon, strategier for avspenning, normalisere funksjonelle bevegelsesmønster, adressere uhensiktsmessig smerteadferd samt benytte motiverende prinsipper er elementer som vektlegges (Ibid).

På den andre siden vil et mer omfattende og individualisert klassifiseringssystem hvor den biopsykososiale modellen integreres på tvers av sub-grupper kunne antas å være fordelaktig. I CFT klassifiseres pasienten på bakgrunn av en fullstendig klinisk undersøkelse som omfatter både et prognostisk spørreskjema, anamnese og kliniske funn. Identifisering av

smertemekanismer for den enkelte pasient sikrer i større grad en individualisering av behandlingsopplegg (Vibe Fersum et al., 2013). Sammenlignet med det relativt enkle klassifiseringsverktøy til SB vil derfor subjektiv klinisk undersøkelse kunne inkludere flere aspekter i vurderingen av pasienten. Det er for eksempel vist at SB ikke nødvendigvis fanger opp alle faktorer som er viktige for prognose og behandlingsforløpet. For eksempel gjelder dette evnen til å identifisere psykososialt stress som ikke er relatert til smertene, som potensielt kan påvirke varigheten av symptomer og dermed prognose (Hill et al., 2010b). Klassifiseringsverktøyet i SB tar heller ikke hensyn til pasientens preferanser, forventninger og tidligere behandling, til tross for at slik informasjon åpenbart er viktig for kliniske beslutninger (Verbeek et al., 2004, Franklin et al., 2016). Pasientens preferanser er også ett av fundamentene i evidensbasert praksis (Sackett et al., 1996).

CFT kan anses for å være mer pasientsentrert hvor individuelle fysiske og psykososiale faktorer ivaretas uavhengig av sub-grupper. I behandling vektlegger CFT reflekterende kommunikasjon, gi pasienten en forståelse av smertemekanismene, korrigering av uhensiktsmessige oppfatninger samt delt beslutningstaking (O'Sullivan, 2005, Vibe Fersum et al., 2013). Det er dessuten vist at pasientens endring av perspektiv og rekonseptualisering av plagene har mye å si for pasientutkomme. De som har god effekt av CFT ser ut til å avhenge av pasientens tilegnelse av en biopsykososial forståelse av ryggsmertene, samt pasientens opplevelse av egenmestring og hvor mye de føler de kan håndtere plagene sine selv (Bunzli et al., 2016).

I SB klassifiseres medium-risikogruppen på bakgrunn av både fysiske og psykososiale variabler men skiller seg fra høy-risikogruppen med psykososial score under 4. Medium-risikogruppen får i midlertidig ikke behandling spesielt rettet mot aktuelle psykososiale faktorer (Hill et al., 2008), noe som potensielt kan føre til begrenset effekt av intervensjonen for pasienter hvor psykososiale faktorer er avgjørende for bedring. Beneciuk og medarbeidere (2015) påpeker at SB kun gir informasjon om en samlet risiko og ikke detaljert informasjon om spesifikke psykologiske faktorer, noe som potensielt kan føre til feilklassifisering. Noen pasienter kan ha behov for vurdering av spesifikke psykologiske faktorer. Høy-risikogruppen er vist å være sterkest assosiert med depressive symptomer, noe som kan tyde på at de vil ha nytte av en mer detaljert vurdering av slike faktorer, for eksempel i avgjørelse om de trenger videre henvisning til psykologisk helsehjelp (Ibid). Selv om klassifisering til høy-risikogruppen hovedsakelig er drevet av psykososiale variabler, scorer disse pasientene også

høyest på fysiske variabler, noe som illustrerer kompleksiteten i denne gruppen (Main et al., 2012). Slike mangler i integrering av fysiske og psykososiale faktorer samt forståelse av hvordan de ulike prognostiske faktorene påvirker utkomme, kan derfor tenkes å føre til mindre målrettet behandling ved SB sammenlignet med CFT.

Både SB og CFT er bygget på omfattende forskning og evidens fra tidligere studier, og de metodologiske faktorene angående utvikling av klassifiseringssystemene anses som solide. SB må likevel betraktes å ha kommet lenger i validering av klassifiseringssystemet enn CFT på bakgrunn av større og bredere valideringsstudier. Sammenlignet med CFT har SB tydeligere rammer for tiltak innen de ulike sub-gruppene og kan anses som mer standardiserte. SB er også i større grad styrende med hensyn til varighet og omfang av behandling for de ulike sub-gruppene. SB leverer derfor en mer generisk tilnærming som er enkel å ta i bruk og kan utføres på store populasjoner med mindre krav til omfattende ressurser og utdanning. CFT inkluderer på sin side en fullstendig klinisk vurdering, samt vektlegger delt beslutningstaking og pasientens preferanser. CFT anses som mer pasientsentrert og integrerer bedre fysiske og psykososiale aspekter i både klassifisering og behandling.

#### **5.4. Betragtninger av stratifisert behandling**

Stratifisert behandling handler til syvende og sist om den individuelle pasienten blir bedre i møte med en behandler, ved å klassifiseres på en bestemt måte. Begge klassifiseringssystemene ser derfor ut til å tilføre viktige elementer i behandling av korsryggpasienter, og trenger ikke nødvendigvis å utelukke bruk av det ene. For eksempel kan prognose benyttes som en første screening for å grovt bestemme behandlingsalternativer, mens underliggende smertemekanismer kan styre videre beslutningstaking om mer individuelle og spesifikke tiltak (Foster et al., 2013). De er i midlertidig utviklet for å guide konservativ behandling, og sier derfor ikke noe om hvilke pasienter som kan ha nytte av kirurgisk behandling. Å inkorporere kirurgi som behandlingsalternativ har blitt antydnet å kunne bidra til mer omfattende klassifiseringssystem som potensielt er mer effektive (Fairbank et al., 2011).

Både blant forskere og klinikere blir stratifisert behandling oppfattet som et viktig bidrag til å øke pasientsentrert behandling, bedre effektivitet og bedre bruk av ressurser (Sheeran et al., 2015). Fremtidig forskning anbefales å rettes mot stratifisert behandling som både inkluderer

prognostisk risikoprofilering, underliggende kausale mekanismer og behandlingsrespons (Foster et al., 2013, Abbott, 2016). Det er publisert flere retningslinjer for å forbedre kvaliteten på forskning innen stratifisert behandling, og RCT av høy kvalitet er ansett som det foretrukne designet. Det finnes likevel ingen metodisk gullstandard for å utvikle og validere klassifiseringssystemer, noe som representerer en begrensning i forskningen (Ford and Hahne, 2013). I hvor stor grad klinikere og pasienter er villige til å basere behandling utelukkende på stratifisert behandling er også usikkert. Selv om mange klinikere og helseaktører er positive til stratifisert behandling påpekes det i midlertidig en del barrierer og ulemper med stratifisert behandling. Dette gjelder oppfatninger om klassifisering som en trussel mot profesjonell selvstendighet og klinisk resonnering, krav til betydelig utdanning, og lite klinisk relevans (Sheeran et al., 2015).

Selv om mange advokerer stratifisert behandling av korsryggsmerter som den videre veien å gå, har det også blitt påpekt tydelige begrensninger med å klassifisere NSLBP med dagens vitenskapelige metoder. Evidens basert på en reduksjonistisk og mekanistisk tilnærming har ikke kunnet gi oss noe tydelig svar med hensyn til årsaksmekanismer for NSLBP (Balague et al., 2012, Eriksen et al., 2013). Generell kunnskap basert på statistiske gjennomsnitt vektlegges fremfor individuelle karakteristikk, noe som kan gå på bekostning av en helhetlig forståelse av pasienten og pasientsentrert behandling (Kerry et al., 2012). Det kan også stilles spørsmål om den biopsykososiale modellen medfører visse begrensninger, da den fortsatt er forankret i et reduksjonistisk rammeverk med et klart skille mellom det fysiologiske og psykologiske. I stedet for å prøve å forstå pasientens helhetlige opplevelse er det fare for at pasienten reduseres til fysiske, psykologiske og sosiale faktorer som man antar kan behandles hver for seg (Butler et al., 2004).

Det er grunn til å anta at en kompleks interaksjon av patoanatomiske, fysiske, neurofysiologiske, psykologiske og sosiale faktorer påvirker hvert enkelt individ i varierende grad (O'Sullivan, 2005, Balague et al., 2012). Ved NSLBP representerer derfor hver pasient en unik kombinasjon av ulike medvirkende faktorer og symptomer (Eriksen et al., 2013). Alt fra oppfatninger og erfaringer til betingete bevegelsesrespons og neurofysiologiske mekanismer utgjør en myriade av potensielle kausale faktorer for den enkelte pasient. Det betyr at pasienter innen en sub-gruppe kan respondere høyst ulikt på samme behandling fordi de tilfører situasjonen unike kausale faktorer, selv om de innehar felles sub-gruppekarakteristikk. Å klassifisere pasienter hvor man gir tilnærmet lik behandling innen

hver sub-gruppe vil derfor kunne medføre store begrensinger (Ibid). Det komplekse samspillet av ulike faktorer vanskeliggjør derfor bruk av tradisjonelle lineære årsaksforklaringer, hvor antagelsen om at samme årsak vil gi samme effekt (Kerry et al., 2012).

Selv om multidimensjonal stratifisert behandling viser lovende resultater og har mange tilhengere både blant klinikere og forskere, er forskningen og evidensgrunnlaget fortsatt sparsom. Samtidig påpekes det begrensinger ved dagens vitenskapelige metoder som grunnlag for å forstå kausalitet og evidens ved NSLBP. Dette taler for at mer forskning på området er nødvendig, samt å vurdere bruk av nye perspektiver og metoder innen forskning.

### **5.5. Begrensninger med denne systematiske oversikten**

I denne oversikten ble det kun søkt på engelsk språk noe som kan medføre en svakhet. I søket var det i midlertidig få studier på annet språk enn engelsk, og evidensen for eksklusjon av andre språk er motstridende (Furlan et al., 2015). Det ble ikke gjort noe forsøk på å identifisere upubliserte studier. Upubliserte studier er ofte små med ikke-signifikante resultater og kan føre til publikasjonsbias (Egger and Smith, 1998).

I utvelgelsen av studier anbefales det at to uavhengige forfattere foretar utvelgelse og vurdering av systematiske feil (Furlan et al., 2015). På grunn av praktiske årsaker og studiets omfang ble utvelgelse av studier og vurdering av inkluderte studier gjort av kun førsteforfatter. Selv om veileder ble brukt for å diskutere usikre tilfeller medfører dette en betydelig svakhet.

Tolkning av resultatene i denne systematiske oversikten er begrenset av få identifiserte studier, og nye studier vil potensielt ha en stor påvirkning på funnene. På grunn av få inkluderte studier og klassifiseringssystemer som er heterogene av natur ble det ikke foretatt noen form for meta-analyse. Selv om dette kan være en begrensning, understreker det også hvor lite data som er tilgjengelig med hensyn til effekt av multidimensjonal stratifisert behandling.

Alle forfatterne i denne studien er fysioterapeuter noe som kan representere en interessekonflikt. Selv om det er lite sannsynlig kan dette medføre en forutinntatthet i utvelgelse og vurdering av studier.

## **6. KONKLUSJON**

Resultatene av denne studien viser at multidimensjonal stratifisert behandling av NSLBP er mer effektivt enn standard konservativ behandling. Kun to RCT med to ulike klassifiseringssystemer ble inkludert. Begge studiene var av høy metodisk kvalitet. CFT viser i midlertidig betydelig større klinisk viktige effekter enn SB. På grunn av få studier, stor variasjon i effektstørrelser samt ulike studiekarakteristikker, må resultatene tolkes med varsomhet. Evidensgrunnlaget for multidimensjonal stratifisert behandling er derfor beskjedent og mer forskning kreves for å kunne gi spesifikke anbefalinger. Dette gjelder både valideringsstudier i andre settinger og impact-analyser.

## LITTERATURLISTE

- ABBOTT, A. 2016. Evidence base and future research directions in the management of low back pain. *World J Orthop*, 7, 156-61.
- APELDOORN, A. T., OSTELO, R. W., VAN HELVOIRT, H., FRITZ, J. M., KNOL, D. L., VAN TULDER, M. W. & DE VET, H. C. 2012. A randomized controlled trial on the effectiveness of a classification-based system for subacute and chronic low back pain. *Spine (Phila Pa 1976)*, 37, 1347-56.
- BAGNALL, D. L. 2010. Physiatry: What's the end game? *PM R*, 2, 3-5.
- BALAGUE, F., MANNION, A. F., PELLISE, F. & CEDRASCHI, C. 2012. Non-specific low back pain. *Lancet*, 379, 482-91.
- BENECIUK, J. M., ROBINSON, M. E. & GEORGE, S. Z. 2015. Subgrouping for patients with low back pain: a multidimensional approach incorporating cluster analysis and the STarT Back Screening Tool. *J Pain*, 16, 19-30.
- BERGSTROM, G., JENSEN, I. B., BODIN, L., LINTON, S. J. & NYGREN, A. L. 2001. The impact of psychologically different patient groups on outcome after a vocational rehabilitation program for long-term spinal pain patients. *Pain*, 93, 229-37.
- BOERSMA, K. & LINTON, S. J. 2005. Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *Clin J Pain*, 21, 38-43; discussion 69-72.
- BORKAN, J., VAN TULDER, 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 (Phila Pa 1976)*, 27, E128-32.
- 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 (03622436)*, 31, 623-631 9p.
- BROOKES, S. T., WHITELEY, E., EGGER, M., SMITH, G. D., MULHERAN, P. A. & PETERS, T. J. 2004. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol*, 57, 229-36.
- BUNZLI, S., MCEVOY, S., DANKAERTS, W., O'SULLIVAN, P. & O'SULLIVAN, K. 2016. Patient Perspectives on Participation in Cognitive Functional Therapy for Chronic Low Back Pain. *Phys Ther*, 96, 1397-407.
- BUTLER, C. C., EVANS, M., GREAVES, D. & SIMPSON, S. 2004. Medically unexplained symptoms: the biopsychosocial model found wanting. *J R Soc Med*, 97, 219-22.
- CAIRNS, M. C., FOSTER, N. E. & WRIGHT, C. 2006. Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. *Spine (Phila Pa 1976)*, 31, E670-81.
- CAMPBELL, P., FOSTER, N. E., THOMAS, E. & DUNN, K. M. 2013. Prognostic Indicators of Low Back Pain in Primary Care: Five-Year Prospective Study. *J Pain*, 14, 873-83.
- CHAN, A., FORD, J., HAHNE, A., SURKITT, L., SLATER, S., HINMAN, R. & TAYLOR, N. 2013. 12 month results of a randomised controlled trial comparing subgroup specific physiotherapy against advice for people with low back disorders. *Journal of Science and Medicine in Sport*, 16, e9.
- CHILDS, J. D., FRITZ, J. M., FLYNN, T. W., IRRGANG, J. J., JOHNSON, K. K., MAJKOWSKI, G. R. & DELITTO, A. 2004. A clinical prediction rule to identify patients with low back



- pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med*, 141, 920-8.
- CROOK, J., MILNER, R., SCHULTZ, I. Z. & STRINGER, B. 2002. Determinants of occupational disability following a low back injury: a critical review of the literature. *J Occup Rehabil*, 12, 277-95.
- DANKAERTS, W. & O'SULLIVAN, P. 2011. The validity of O'Sullivan's classification system (CS) for a sub-group of NS-CLBP with motor control impairment (MCI): overview of a series of studies and review of the literature. *Man Ther*, 16, 9-14.
- DANKAERTS, W., O'SULLIVAN, P., BURNETT, A., STRAKER, L., DAVEY, P. & GUPTA, R. 2009. Discriminating healthy controls and two clinical subgroups of nonspecific chronic low back pain patients using trunk muscle activation and lumbosacral kinematics of postures and movements: a statistical classification model. *Spine (Phila Pa 1976)*, 34, 1610-8.
- 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.
- DAY, S. J. & ALTMAN, D. G. 2000. Blinding in clinical trials and other studies. *BMJ*, 321, 504.
- DUNN, K. M. & CROFT, P. R. 2005. Classification of low back pain in primary care: using "bothersomeness" to identify the most severe cases. *Spine*, 30, 1887.
- EGGER, M. & SMITH, G. D. 1998. Bias in location and selection of studies. *Bmj*, 316, 61-6.
- ENDEAN, A., PALMER, K. T. & COGGON, D. 2011. Potential of magnetic resonance imaging findings to refine case definition for mechanical low back pain in epidemiological studies: a systematic review. *Spine (Phila Pa 1976)*, 36, 160-9.
- ERIKSEN, T. E., KERRY, R., MUMFORD, S., LIE, S. A. & ANJUM, R. L. 2013. At the borders of medical reasoning: aetiological and ontological challenges of medically unexplained symptoms. *Philos Ethics Humanit Med*, 8, 11.
- FAIRBANK, J., GWILYM, S. E., FRANCE, J. C., DAFFNER, S. D., DETTORI, J., HERMSMEYER, J. & ANDERSSON, G. 2011. The role of classification of chronic low back pain. *Spine (Phila Pa 1976)*, 36, S19-42.
- FERSUM, K. V., DANKAERTS, W., O'SULLIVAN, P. B., MAES, J., SKOUEN, J. S., BJORDAL, J. M. & KVALE, A. 2010. Integration of subclassification strategies in randomised controlled clinical trials evaluating manual therapy treatment and exercise therapy for non-specific chronic low back pain: a systematic review. *Br J Sports Med*, 44, 1054-62.
- FEYS, F., BEKKERING, G. E., SINGH, K. & DEVROEY, D. 2014. Do randomized clinical trials with inadequate blinding report enhanced placebo effects for intervention groups and nocebo effects for placebo groups? *Syst Rev*, 3, 14.
- FIELD, J. & NEWELL, D. 2012. Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. *Chiropr Man Therap*, 20, 17.
- FORD, J., STORY, I., O'SULLIVAN, P. & MCMEEKEN, J. 2007. Classification systems for low back pain: a review of the methodology for development and validation. *Physical Therapy Reviews*, 12, 33-42.
- FORD, J. J. & HAHNE, A. J. 2013. Pathoanatomy and classification of low back disorders. *Man Ther*, 18, 165-8.

- FOSTER, N. E., BISHOP, A., THOMAS, E., MAIN, C., HORNE, R., WEINMAN, J. & HAY, E. 2008. Illness perceptions of low back pain patients in primary care: What are they, do they change and are they associated with outcome? *Pain*, 136, 177-187.
- FOSTER, N. E. & DELITTO, A. 2011. Embedding psychosocial perspectives within clinical management of low back pain: integration of psychosocially informed management principles into physical therapist practice--challenges and opportunities. *Phys Ther*, 91, 790-803.
- FOSTER, N. E., HILL, J. C. & HAY, E. M. 2011. Subgrouping patients with low back pain in primary care: are we getting any better at it? *Man Ther*, 16, 3-8.
- FOSTER, N. E., HILL, J. C., O'SULLIVAN, P. & HANCOCK, M. 2013. Stratified models of care. *Best Pract Res Clin Rheumatol*, 27, 649-61.
- FOSTER, N. E., MULLIS, R., HILL, J. C., LEWIS, M., WHITEHURST, D. G. T., DOYLE, C., KONSTANTINOOU, K., MAIN, C., SOMERVILLE, S., SOWDEN, G., WATHALL, S., YOUNG, J. & HAY, E. M. 2014. Effect of Stratified Care for Low Back Pain in Family Practice (IMPACT Back): A Prospective Population-Based Sequential Comparison. *Ann Fam Med*, 12, 102-11.
- FRANKLIN, Z. C., SMITH, N. C. & FOWLER, N. E. 2016. A qualitative investigation of factors that matter to individuals in the pain management process. *Disabil Rehabil*, 38, 1934-42.
- FREBURGER, J. K., HOLMES, G. M., AGANS, R. P., JACKMAN, A. M., DARTER, J. D., WALLACE, A. S., CASTEL, L. D., KALSBECK, W. D. & CAREY, T. S. 2009. The rising prevalence of chronic low back pain. *Arch Intern Med*, 169, 251-8.
- FRITZ, J. M. 1998. *The efficacy of a homogenous treatment approach versus a classification approach to the treatment of work-related low back pain. A randomized clinical trial.* Ph.D., University of Pittsburgh.
- 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...including commentary by Gordon SL. *Spine (03622436)*, 28, 1363-1372 10p.
- FRITZ, J. M., PIVA, S. R. & CHILDS, J. D. 2005. Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. *Eur Spine J*, 14, 743-50.
- FURLAN, A. D., MALMIVAARA, A., CHOU, R., MAHER, C. G., DEYO, R. A., SCHOENE, M., BRONFORT, G., VAN TULDER, M. W., FROM THE EDITORIAL BOARD OF THE COCHRANE, B. & NECK, G. 2015. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. *Spine (Phila Pa 1976)*.
- FURLAN, A. D., VAN TULDER, M. W., CHERKIN, D. C., TSUKAYAMA, H., LAO, L., KOES, B. W. & BERMAN, B. M. 2005. Acupuncture and dry-needling for low back pain. *Cochrane Database Syst Rev*, CD001351.
- GEORGE, S. Z., ZEPPIERI, G., JR., CERE, A. L., CERE, M. R., BORUT, M. S., HODGES, M. J., REED, D. M., VALENCIA, C. & ROBINSON, M. E. 2008. A randomized trial of behavioral physical therapy interventions for acute and sub-acute low back pain (NCT00373867). *Pain*, 140, 145-57.
- GOPALAKRISHNAN, S. & GANESHKUMAR, P. 2013. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. *J Family Med Prim Care*, 2, 9-14.
- GREENHALGH, T., HOWICK, J. & MASKREY, N. 2014. Evidence based medicine: a movement in crisis? *BMJ*, 348.

- GROTLE, M., VOLLESTAD, N. K. & BROX, J. I. 2006. Clinical course and impact of fear-avoidance beliefs in low back pain: prospective cohort study of acute and chronic low back pain: II. *Spine (Phila Pa 1976)*, 31, 1038-46.
- HALDEMAN, S., KOPANSKY-GILES, D., HURWITZ, E. L., HOY, D., MARK ERWIN, W., DAGENAIS, S., KAWCHUK, G., STROMQVIST, B. & WALSH, N. 2012. Advancements in the management of spine disorders. *Best Pract Res Clin Rheumatol*, 26, 263-80.
- HALLIDY, M., PAPPAS, E., CLARE, H., HANCOCK, M. & FERREIRA, P. 2015. Short-term effects of mckenzie vs. motor control approach for patients with chronic lowback pain and a derangement classification. *Physiotherapy (United Kingdom)*, 101, eS506-eS507.
- HANCOCK, M., HERBERT, R. D. & MAHER, C. G. 2009. A guide to interpretation of studies investigating subgroups of responders to physical therapy interventions. *Phys Ther*, 89, 698-704.
- HASKINS, R., RIVETT, D. A. & OSMOTHERLY, P. G. 2012. Clinical prediction rules in the physiotherapy management of low back pain: a systematic review. *Man Ther*, 17, 9-21.
- HAY, E. M., DUNN, K. M., HILL, J. C., LEWIS, M., MASON, E. E., KONSTANTINOU, K., SOWDEN, G., SOMERVILLE, S., VOHORA, K., WHITEHURST, D. & MAIN, C. J. 2008. A randomised clinical trial of subgrouping and targeted treatment for low back pain compared with best current care. The STarT Back Trial Study Protocol. *BMC Musculoskelet Disord*, 9, 58.
- HAY, E. M., MULLIS, R., LEWIS, M., VOHORA, K., MAIN, C. J., WATSON, P., DZIEDZIC, K. S., SIM, J., MINNS LOWE, C. & CROFT, P. R. 2005. Comparison of physical treatments versus a brief pain-management programme for back pain in primary care: a randomised clinical trial in physiotherapy practice. *Lancet*, 365, 2024-30.
- HAYDEN, J. A., DUNN, K. M., VAN DER WINDT, D. A. & SHAW, W. S. 2010. What is the prognosis of back pain? *Best Pract Res Clin Rheumatol*, 24, 167-79.
- HAYDEN, J. A., VAN TULDER, M. W., MALMIVAARA, A. & KOES, B. W. 2005. Exercise therapy for treatment of non-specific low back pain. *Cochrane Database Syst Rev*, CD000335.
- HENRY, S. M., VAN DILLEN, L. R., OUELLETTE-MORTON, R. H., HITT, J. R., LOMOND, K. V., DESARNO, M. J. & BUNN, J. Y. 2014. Outcomes are not different for patient-matched versus nonmatched treatment in subjects with chronic recurrent low back pain: a randomized clinical trial. *Spine Journal: Official Journal of the North American Spine Society*, 14, 2799-810.
- HENSCHKE, N., KUIJPERS, T., RUBINSTEIN, S. M., VAN MIDDELKOOP, M., OSTELO, R., VERHAGEN, A., KOES, B. W. & VAN TULDER, M. W. 2012. Trends over time in the size and quality of randomised controlled trials of interventions for chronic low-back pain. *Eur Spine J*, 21, 375-81.
- HENSCHKE, N., OSTELO, R. W., VAN TULDER, M. W., VLAEYEN, J. W., MORLEY, S., ASSENDELFT, W. J. & MAIN, C. J. 2010. Behavioural treatment for chronic low-back pain. *Cochrane Database Syst Rev*, CD002014.
- HICKS, G. E., FRITZ, J. M., DELITTO, A. & MCGILL, S. M. 2005. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Arch Phys Med Rehabil*, 86, 1753-62.
- HILL, J. C., DUNN, K. M., LEWIS, M., MULLIS, R., MAIN, C. J., FOSTER, N. E. & HAY, E. M. 2008. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*, 59, 632-41.

- HILL, J. C., VOHORA, K., DUNN, K. M., MAIN, C. J. & HAY, E. M. 2010b. Comparing the STarT back screening tool's subgroup allocation of individual patients with that of independent clinical experts. *Clin J Pain*, 26, 783-7.
- HILL, J. C., WHITEHURST, D. G., LEWIS, M., BRYAN, S., DUNN, K. M., FOSTER, N. E., KONSTANTINOOU, K., MAIN, C. J., MASON, E., SOMERVILLE, S., SOWDEN, G., VOHORA, K. & HAY, E. M. 2011. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*, 378, 1560-71.
- HOY, D., BROOKS, P., BLYTH, F. & BUCHBINDER, R. 2010. The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol*, 24, 769-81.
- HOY, D., MARCH, L., BROOKS, P., BLYTH, F., WOOLF, A., BAIN, C., WILLIAMS, G., SMITH, E., VOS, T., BARENDREGT, J., MURRAY, C., BURSTEIN, R. & BUCHBINDER, R. 2014. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Annals of the Rheumatic Diseases*, 73, 968-974.
- JACOBS, W., VAN DER GAAG, N. A., TUSCHEL, A., DE KLEUVER, M., PEUL, W., VERBOUT, A. J. & ONER, F. C. 2012. Total disc replacement for chronic back pain in the presence of disc degeneration. *Cochrane Database Syst Rev*, 9, CD008326.
- JARVIK, J. G., HOLLINGWORTH, W., HEAGERTY, P. J., HAYNOR, D. R., BOYKO, E. J. & DEYO, R. A. 2005. Three-year incidence of low back pain in an initially asymptomatic cohort: clinical and imaging risk factors. *Spine (Phila Pa 1976)*, 30, 1541-8; discussion 1549.
- JORDAN, K., DUNN, K. M., LEWIS, M. & CROFT, P. 2006. A minimal clinically important difference was derived for the Roland-Morris Disability Questionnaire for low back pain. *J Clin Epidemiol*, 59, 45-52.
- KARAYANNIS, N. V., JULL, G. A. & HODGES, P. W. 2012. Physiotherapy movement based classification approaches to low back pain: comparison of subgroups through review and developer/expert survey. *BMC Musculoskelet Disord*, 13, 24.
- KENT, P., KEATING, J. L. & LEBOEUF-YDE, C. 2010. Research methods for subgrouping low back pain. *BMC Med Res Methodol*, 10, 62.
- KERRY, R., ERIKSEN, T. E., LIE, S. A., MUMFORD, S. D. & ANJUM, R. L. 2012. Causation and evidence-based practice: an ontological review. *J Eval Clin Pract*, 18, 1006-12.
- KIRKEHEI, I. & ORMSTAD, S. S. 2013. Litteratursøk. *Norsk epidemiologi*.
- KOES, B. W., VAN TULDER, M., LIN, C. W., MACEDO, L. G., MCAULEY, J. & MAHER, C. 2010. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J*, 19, 2075-94.
- KOES, B. W., VAN TULDER, M. W. & THOMAS, S. 2006. Diagnosis and treatment of low back pain. *BMJ*, 332, 1430-4.
- KONGSTED, A., JOHANNESSEN, E. & LEBOEUF-YDE, C. 2011. Feasibility of the STarT back screening tool in chiropractic clinics: a cross-sectional study of patients with low back pain. *Chiropr Man Therap*, 19, 10.
- LEDERMAN, E. 2011. The fall of the postural-structural-biomechanical model in manual and physical therapies: exemplified by lower back pain. *J Bodyw Mov Ther*, 15, 131-8.
- LINTON, S. J. 2001. Occupational psychological factors increase the risk for back pain: a systematic review. *J Occup Rehabil*, 11, 53-66.
- LINTON, S. J. 2005. *Understanding pain for better clinical practice: A psychological perspective*, Edinburgh, Elsevier.
- LÆRUM, E., BROX, J. I., STORHEIM, K., ESPELAND, A., HALDORSEN, E., MUNCH-ELLINGSEN, J., NIELSEN, L. L., ROSSVOLL, I., SKOUEN, J. S., STIG, L. C. & WERNER,

- E. L. 2007. Nasjonale kliniske retningslinjer. Korsryggsmerter - med og uten nerverotsaffeksjon. Formidlingsenheten for muskel- og skjelettlidelser (FORMI).
- LÆRUM, E. B., S.; IHLEBÆK, C.; JOHNSEN, K.; NATVIG, B.; AAS, E. 2013. Et muskel-skjelettregnskap. Forekomst og kostnader knyttet til skader, sykdommer og plager i muskel- og skjelettsystemet. Oslo: FORMI.
- MA, V. Y., CHAN, L. & CARRUTHERS, K. J. 2014. Incidence, prevalence, costs, and impact on disability of common conditions requiring rehabilitation in the United States: stroke, spinal cord injury, traumatic brain injury, multiple sclerosis, osteoarthritis, rheumatoid arthritis, limb loss, and back pain. *Arch Phys Med Rehabil*, 95, 986-995 e1.
- MAFI, J. N., MCCARTHY, E. P., DAVIS, R. B. & LANDON, B. E. 2013. Worsening trends in the management and treatment of back pain. *JAMA Intern Med*, 173, 1573-81.
- MAIN, C. J., SOWDEN, G., HILL, J. C., WATSON, P. J. & HAY, E. M. 2012. Integrating physical and psychological approaches to treatment in low back pain: the development and content of the STarT Back trial's 'high-risk' intervention (StarT Back; ISRCTN 37113406). *Physiotherapy*, 98, 110-117.
- MENEZES COSTA, L. C., MAHER, C. G., HANCOCK, M. J., MCAULEY, J. H., HERBERT, R. D. & COSTA, L. O. P. 2012. The prognosis of acute and persistent low-back pain: a meta-analysis. *Cmaj*, 184, E613-24.
- MILLER, E. R., SCHENK, R. J., KARNES, J. L. & ROUSSELLE, J. G. 2005. A comparison of the McKenzie approach to a specific spine stabilization program for chronic low back pain. *Journal of Manual & Manipulative Therapy (Journal of Manual & Manipulative Therapy)*, 13, 103-112 10p.
- MOHER, D., SHAMSEER, L., CLARKE, M., GHERSI, D., LIBERATI, A., PETTICREW, M., SHEKELLE, P., STEWART, L. A. & GROUP, P.-P. 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*, 4, 1.
- MORSO, L., ALBERT, H., KENT, P., MANNICHE, C. & HILL, J. 2011. Translation and discriminative validation of the STarT Back Screening Tool into Danish. *Eur Spine J*, 20, 2166-73.
- NASSER, M., VAN WEEL, C., VAN BINSBERGEN, J. J. & VAN DE LAAR, F. A. 2012. Generalizability of systematic reviews of the effectiveness of health care interventions to primary health care: concepts, methods and future research. *Fam Pract*, 29 Suppl 1, i94-i103.
- O'SULLIVAN, P. 2005. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. *Man Ther*, 10, 242-55.
- O'SULLIVAN, P. 2012. It's time for change with the management of non-specific chronic low back pain. *British Journal of Sports Medicine*, 46, 224-227.
- OSTELO, R. W., DEYO, R. A., STRATFORD, P., WADDELL, G., CROFT, P., VON KORFF, M., BOUTER, L. M. & DE VET, H. C. 2008. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*, 33, 90-4.
- PAGE, I., ABOUD, J., J. O. S., LAURENCELLE, L. & DESCARREAUX, M. 2015. Chronic low back pain clinical outcomes present higher associations with the STarT Back Screening Tool than with physiologic measures: a 12-month cohort study. *BMC Musculoskelet Disord*, 16, 201.
- PETERSEN, T., KRYGER, P., EKDAHL, C., OLSEN, S. & JACOBSEN, S. 2002. The effect of McKenzie therapy as compared with that of intensive strengthening training for

- the treatment of patients with subacute or chronic low back pain: A randomized controlled trial. *Spine (Phila Pa 1976)*, 27, 1702-9.
- PETERSEN, T., LARSEN, K., NORDSTEEN, J., OLSEN, S., FOURNIER, G. & JACOBSEN, S. 2011. The McKenzie Method Compared With Manipulation When Used Adjunctive to Information and Advice in Low Back Pain Patients Presenting With Centralization or Peripheralization: A Randomized Controlled Trial. *Spine (03622436)*, 36, 1999-2010 12p.
- PETERSEN T, L. M., THORSEN H, MANNICHE C, EKDAHL C, JACOBSEN S. 2003. Diagnostic classification of non-specific low back pain. A new system integrating path-anatomic and clinical categories *Physiother Theory & Practice*, 19, 213-37.
- PINCUS, T., KENT, P., BRONFORT, G., LOISEL, P., PRANSKY, G. & HARTVIGSEN, J. 2013. Twenty-five years with the biopsychosocial model of low back pain-is it time to celebrate? A report from the twelfth international forum for primary care research on low back pain. *Spine (Phila Pa 1976)*, 38, 2118-23.
- POLIT, D. F. & BECK, C. T. 2012. *Nursing research : generating and assessing evidence for nursing practice*, Philadelphia, Pa, Wolters Kluwer Health.
- ROBERTS, I., KER, K., EDWARDS, P., BEECHER, D., MANNO, D. & SYDENHAM, E. 2015. The knowledge system underpinning healthcare is not fit for purpose and must change. *Bmj*, 350, h2463.
- ROBINSON, M. E. & GEORGE, S. Z. 2012. Screening for problematic low back pain: STarT. *Pain*, 153, 2159-60.
- RUBINSTEIN, S. M., VAN MIDDELKOOP, M., ASSENDELFT, W. J., DE BOER, M. R. & VAN TULDER, M. W. 2011. Spinal manipulative therapy for chronic low-back pain. *Cochrane Database Syst Rev*, CD008112.
- SACKETT, D. L., ROSENBERG, W. M., GRAY, J. A., HAYNES, R. B. & RICHARDSON, W. S. 1996. Evidence based medicine: what it is and what it isn't. *Bmj*, 312, 71-2.
- SCHOLZ, J., MANNION, R. J., HORD, D. E., GRIFFIN, R. S., RAWAL, B., ZHENG, H., SCOFFINGS, D., PHILLIPS, A., GUO, J., LAING, R. J., ABDI, S., DECOSTERD, I. & WOOLF, C. J. 2009. A novel tool for the assessment of pain: validation in low back pain. *PLoS Med*, 6, e1000047.
- SHEERAN, L., COALES, P. & SPARKES, V. 2015. Clinical challenges of classification based targeted therapies for non-specific low back pain: What do physiotherapy practitioners and managers think? *Man Ther*, 20, 456-62.
- SHEERAN, L., VAN DEURSEN, R., CATERSON, B. & SPARKES, V. 2013. Classification-guided versus generalized postural intervention in subgroups of nonspecific chronic low back pain: a pragmatic randomized controlled study. *Spine (03622436)*, 38, 1613-1625 13p.
- SMART, K. M., BLAKE, C., STAINES, A. & DOODY, C. 2011. The Discriminative validity of "nociceptive," "peripheral neuropathic," and "central sensitization" as mechanisms-based classifications of musculoskeletal pain. *Clin J Pain*, 27, 655-63.
- SMEETS, R. J., VLAEYEN, J. W., HIDDING, A., KESTER, A. D., VAN DER HEIJDEN, G. J. & KNOTTNERUS, J. A. 2008. Chronic low back pain: physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain*, 134, 263-76.
- STANTON, T. R., HANCOCK, M. J., MAHER, C. G. & KOES, B. W. 2010. Critical appraisal of clinical prediction rules that aim to optimize treatment selection for musculoskeletal conditions. *Phys Ther*, 90, 843-54.

- VERBEEK, J., SENGERS, M. J., RIEMENS, L. & HAAFKENS, J. 2004. Patient expectations of treatment for back pain: a systematic review of qualitative and quantitative studies. *Spine (Phila Pa 1976)*, 29, 2309-18.
- VERKERK, K., LUIJSTERBURG, P. A., HEYMANS, M. W., RONCHETTI, I., POOL-GOUDZWAARD, A. L., MIEDEMA, H. S. & KOES, B. W. 2013. Prognosis and course of disability in patients with chronic nonspecific low back pain: a 5- and 12-month follow-up cohort study. *Phys Ther*, 93, 1603-14.
- VERKERK, K., LUIJSTERBURG, P. A., MIEDEMA, H. S., POOL-GOUDZWAARD, A. & KOES, B. W. 2012. Prognostic factors for recovery in chronic nonspecific low back pain: a systematic review. *Phys Ther*, 92, 1093-108.
- VIBE FERSUM, K., O'SULLIVAN, P., SKOEN, J. S., SMITH, A. & KVALE, A. 2013. Efficacy of classification-based cognitive functional therapy in patients with non-specific chronic low back pain: a randomized controlled trial. *Eur J Pain*, 17, 916-28.
- VIBE FERSUM, K., O'SULLIVAN, P. B., KVALE, A. & SKOEN, J. S. 2009. Inter-examiner reliability of a classification system for patients with non-specific low back pain. *Man Ther*, 14, 555-61.
- 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, 566-79.
- WADDELL, G. 1987. A new clinical model for the treatment of low-back pain. *Spine (Phila Pa 1976)*, 12, 632-44.
- WADDELL, G. 2005. Subgroups within "nonspecific" low back pain. *J Rheumatol*, 32, 395-6.
- WIDEMAN, T. H., HILL, J. C., MAIN, C. J., LEWIS, M., SULLIVAN, M. J. & HAY, E. M. 2012. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. *Pain*, 153, 2182-91.
- WILKENS, P., SCHEEL, I. B., GRUNDNES, O., HELLUM, C. & STORHEIM, K. 2013. Prognostic factors of prolonged disability in patients with chronic low back pain and lumbar degeneration in primary care: a cohort study. *Spine (Phila Pa 1976)*, 38, 65-74.

## **VEDLEGG**

Vedlegg 1: Søkehistorikk MEDLINE, EMBASE, PsycINFO, CINAHL, PEDro

Vedlegg 2: Karakteristikker av inkluderte studier og evaluering av risiko for systematiske feil, Hill et al 2011

Vedlegg 3: Karakteristikker av inkluderte studier og evaluering av risiko for systematiske feil, Vibe Fersum et al 2013



## Vedlegg 1

### **Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to 2016 February 09>**

9. feb.2016

---

- 1 Low Back Pain/ (16216)
- 2 low back pain.ti,ab,kw. (19340)
- 3 1 or 2 (25148)
- 4 (stratified or classification or subgroup\* or biopsychosocial or multidimensional).ti,ab,kw. (453955)
- 5 3 and 4 (1701)
- 6 randomized controlled trial.pt. or randomized.mp. (626541)
- 7 5 and 6 (333)

### **Ovid Embase <1974 to 2016 February 09>**

9.feb.2016

---

- 1 low back pain/ (41327)
- 2 low back pain.ti,ab,kw. (26784)
- 3 1 or 2 (46455)
- 4 (stratified or classification or subgroup\* or biopsychosocial or multidimensional).ti,ab,kw. (625688)
- 5 3 and 4 (2531)
- 6 randomized controlled trial.pt. or randomized.mp. (741354)
- 7 5 and 6 (428)

### **Ovid PsycINFO <1806 to 2016 February 09>**

9.feb.2016

---

- 1 back pain/ (3168)
- 2 low back pain.tw. (2797)
- 3 1 or 2 (3823)
- 4 (stratified or classification or subgroup\* or biopsychosocial or multidimensional).tw. (123664)
- 5 3 and 4 (418)
- 6 random\*.tw. (151425)
- 7 5 and 6 (47)

### **Cinahl (Ebsco) <1981 to 9.feb. 2016>**

- 
- |    |                                                                                  |         |
|----|----------------------------------------------------------------------------------|---------|
| S1 | (MH "Low Back Pain")                                                             | 10,409  |
| S2 | Low Back Pain                                                                    | 12,260  |
| S3 | S1 OR S2                                                                         | 12,260  |
| S4 | stratified OR classification OR subgroup* OR biopsychosocial OR multidimensional | 126,815 |
| S5 | S3 AND S4                                                                        | 1,078   |
| S6 | random*                                                                          | 163,060 |
| S7 | S5 AND S6                                                                        | 230     |

## **PEDro, søk opp til 9. Feb. 2016**

---

1. "low back pain" stratified avgrenset til Clinical trial: 16
2. "low back pain" classification avgrenset til Clinical trial: 17
3. "low back pain" subgroup\* avgrenset til Clinical trial: 56
4. "low back pain" multidimensional avgrenset til Clinical trial: 11

## Vedlegg 2

### KARAKTERISTIKKER AV INKLUDERTE STUDIER

Hill et al. 2011

| Metode       | Beskrivelse                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Deltakere    | <p>Land: England</p> <p>Antall deltakere: 851</p> <p>Rekruttering: 10 almenlegepraksiser i Keele General Practice Research Partnership, England</p> <p>Alder: 18 og oppover, gjennomsnitt 50 år</p> <p>Kjønn: begge, 59 % kvinner</p> <p>Inklusjonskriterier: korsryggsmerter uavhengig av varighet med eller uten bensmerter</p> <p>Eksklusjonskriterier: potensiell alvorlig patologi (eks: cauda equina syndrom, inflammatorisk artritt, malignitet), alvorlig sykdom eller komorbiditet, spinal kirurgi siste 6 mnd, gravide, og pasienter som mottok behandling for ryggen (unntatt i primærhelsetjenesten),</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Intervensjon | <p>Deltakerne ble randomisert til en av to intervensjonsgrupper:</p> <ol style="list-style-type: none"> <li>1. Start Back stratifiseringsmodell (SB): ved baseline ble deltakerne klassifisert med StarT Back Screening Tool i enten lav, medium eller høy risiko. Alle pasienter fikk initialt en konsultasjon med promotering og råd om aktivitet, inkludert tilbakekomst til arbeid, samt en brosjyre om øvelser og selvhjelpsgrupper. I tillegg fikk de se en utdanningsvideo på 15 min (Get Back Active) og gitt boken Back Book. Medium- og høy-risiko pasienter ble henvist til videre fysioterapibehandling. Medium-risiko pasienter fikk standardisert fysioterapi adressert mot fysiske symptomer og funksjon. Høy-risiko pasienter fikk fysioterapi med psykologisk tilnærming for å adressere psykologiske hindringer i tillegg til fysiske symptomer og funksjon. Behandling inntil 4 mnd.</li> <li>2. Kontrollgruppe: ved initial vurdering fikk pasientene råd om aktivitet og øvelser. Videre fysioterapibehandling ble bestemt av fysioterapeut ved baseline undersøkelse. Fysioterapeutene hadde generell trening innen fysikalske behandlingsmetoder samt noe trening innen psykologisk tilnærming, men ingen gjennomgikk spesiell trening relatert til dette studiet. Behandling inntil 4 mnd.</li> </ol> |
| Utfallsmål   | <p>Utfallsmål ble innsamlet ved baseline, og ved 4 og 12 mnd oppfølging</p> <ul style="list-style-type: none"> <li>• <b>Funksjon:</b> Roland and Morris Disability Questionnaire (RMDQ), skala fra 0-24 hvor høye score indikerer alvorlig funksjonsnedsettelse (primærutfallsmål)</li> <li>• <b>Smerte:</b> Numerisk rating skala fra 0 – 10 med minste og gjennomsnittlige smerte siste 2 uker, samt aktuelle smerte</li> <li>• <b>Livskvalitet:</b> EuroQol EQ-5D, Short Form 12 (SF-12)</li> <li>• <b>Andre utfallsmål:</b> <ul style="list-style-type: none"> <li>- Pain Catastrophizing scale</li> <li>- Tampa Scale of Kinesiophobia</li> <li>- Hospital Anxiety and Depression Scale</li> </ul> </li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|            | <ul style="list-style-type: none"> <li>- Reduksjon i risiko-gruppe med STarT Back Screening Tool</li> <li>- Oppfatning om samlet endring i ryggsmarter</li> <li>- Antall fysioterapi konsultasjoner</li> <li>- Helsetjenester – og kostnader over 12 mnd</li> <li>- Antall fraværsdager fra jobb på grunn av ryggsmarter</li> <li>- Pasienttilfredstillelse med intervensjonen</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Resultater | <p>Antall deltakere:</p> <ul style="list-style-type: none"> <li>• SB: 568</li> <li>• Kontrollgruppe: 283</li> <li>• Deltakere som manglet oppfølging ved 12mnd: 128 av 568 (23 %) i SB, 74 av 283 (26 %) i kontrollgruppe</li> </ul> <p>Effekt av behandling:</p> <ul style="list-style-type: none"> <li>• <b>Funksjon:</b> Stratifisert behandling viste signifikant bedring ved 4 mnd (<math>p &lt; 0,0001</math>) og 12 mnd (<math>p = 0,0095</math>)</li> <li>• <b>Smerte:</b> Stratifisert behandling viste signifikant bedring ved 4 mnd (<math>p = 0,0005</math>) men ikke 12 mnd (<math>p = 0,5002</math>)</li> <li>• <b>Livskvalitet:</b> Stratifisert behandling viste større gjennomsnittlig helseeffekter målt med EuroQol-5D. SF-12 viste signifikant bedring i favør SB ved 12 mnd oppfølging for den fysiske komponenten (<math>p &lt; 0,0001</math>) men ikke den mentale (<math>p = 0,4050</math>).</li> </ul> |

#### RISIKO AV SYSTEMATISKE FEIL

| Bias domene               | Kilde til bias                                                                                      | Vurdering | Beskrivelse                                                                                                                                                                                                      |
|---------------------------|-----------------------------------------------------------------------------------------------------|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Selection                 | 1. Was the method of randomization adequate? Adekvat randomiseringsmetode?                          | Ja        | Data-generert blokk randomisering                                                                                                                                                                                |
| Selection                 | 2. Was the treatment allocation concealed? Skjult tildeling av intervensjonsgruppe? Alle utfallsmål | Ja        | Fordeling av deltakere ble gjort av en uavhengig, blindet person                                                                                                                                                 |
| Performance and detection | 3. Was the patient blinded to the intervention? Blinding av pasient? Alle utfallsmål                | Nei       | Deltakerne fikk behandling som de var klar over, men kun fortalt at de fikk en av to behandlingsmodeller innen primærhelsetjenesten. Det er usikkert om de forstod hvilken av intervensjonsgruppene de tilhørte. |
| Performance and detection | 4. Was the care provider blinded to the intervention? Blinding av helse tilbyder? Alle utfallsmål   | Nei       | Administratorer av studien var fysioterapeuter som også utførte den aktive intervensjonen. Terapeuter som utførte kontrollbehandlingen var ikke kjent med detaljene i den stratifiserte modellen.                |

|                           |                                                                                                                                                        |    |                                                                                                                                                                                                  |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Performance and detection | 5. Was the outcome assessor blinded to the intervention? Blinding av assessor? Alle utfallsmål?                                                        | Ja | Pasientrapporterte utfallsmål: assessorer vurdert som blindet pga sammenlignbare intervensjonsgrupper. Bare pasientrapporterte data er benyttet i denne oversikten.                              |
| Attrition                 | 6. Was the drop-out rate described and acceptable? Akseptabel frafallsprosent? Alle utfallsmål                                                         | Ja | Frafall er beskrevet og balansert mellom intervensjonsgruppen og kontrollgruppen, med mindre enn 30 % frafall.                                                                                   |
| Attrition                 | 7. Were all randomized participants analyzed in the group to which they were allocated? Alle deltakere analysert - intention-to treat? Alle utfallsmål | Ja | Intention-to-treat analyse ble foretatt for begge gruppene                                                                                                                                       |
| Reporting                 | 8. Are reports of the study free of suggestion of selective outcome reporting? Fri for selektiv rapportering av utfallsmål?                            | Ja | Rapporterte utfallsmål er i overensstemmelse med protokollen (Hay et al. 2008)                                                                                                                   |
| Selection                 | 9. Were the groups similar at baseline regarding the most important prognostic indicators? Tilstrekkelig likhet av baseline karakteristikk?            | Ja | Demografiske faktorer, funksjonsnedsettelse og varighet av smerte var lik for begge gruppene ved baseline                                                                                        |
| Performance               | 10. Were co-interventions avoided or similar? Var ko-intervensjon unngått eller lik?                                                                   | Ja | Begge grupper mottok behandling som var sammenlignbare                                                                                                                                           |
| Performance               | 11. Was the compliance acceptable in all groups? Var compliance akseptabel for alle grupper?                                                           | Ja | Gjennomsnittlig antall behandlinger var likt for kontroll- og intervensjonsgruppen. Med unntak av første konsultasjon er det ikke beskrevet og sammenlignet antall og varighet av konsultasjoner |
| Performance               | 12. Was the timing of the outcome assessment similar in all groups? Likt tidspunkt for vurdering av utfallsmål for alle grupper?                       | Ja | Alle utfallsmål ble samlet før randomisering, og 4 mnd og 12 mnd etter, for begge gruppene                                                                                                       |
| Detection                 | 13. Are other sources of potential bias unlikely? Er andre kilder til bias usannsynlig?                                                                | Ja | Utfallsmål anses som valide, og forfatterne erklærer ingen interessekonflikter                                                                                                                   |

## Vedlegg 3

### KARAKTERISTIKKER AV INKLUDERTE STUDIER

Vibe Fersum et al. 2013

| Metode       | Beskrivelse                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Deltakere    | <p>Land: Norge</p> <p>Antall deltakere: 121</p> <p>Rekruttering: Private fysioterapiklinikker i Bergen og en poliklinisk avdeling for rygg ved Haukeland sykehus</p> <p>Alder: 18 – 65, gjennomsnitt: 42 år</p> <p>Kjønn: begge, 51 % kvinner</p> <p>Inklusjonskriterier: Alder mellom 18 og 65, NSLBP i mer enn 3 mnd som var assosiert med holdning, bevegelser eller aktiviteter, ODI score &gt; 14 % og smerte &gt; 2/10 målt ved pain intensity numeric rating scale (PINSR) siste 14 dager</p> <p>Eksklusjonskriterier: kontinuerlig sykmelding over 4 mnd (pga igangsetting av arbeidsrelaterte intervensjoner), akutt forverring av LBP ved undersøkelse for å unngå regresjon mot gjennomsnittet, spesifikk LBP, kirurgi av underekstremitet siste 3 mnd, kirurgi av lenderygg, graviditet, psykiatrisk diagnose, generalisert konstant uspesifikk smertetilstand, smerte uten tydelig mekanisk adferd, aktiv revmatologisk sykdom, progressiv nevrologisk sykdom, alvorlig hjerte- eller indremedisinsk tilstand, malign sykdom, akutte traumer, infeksjoner eller akutte vaskulære hendelser, mangel på compliance mer enn 50 %</p>                             |
| Intervensjon | <p>Deltakere ble randomisert til en av to intervensjonsgrupper:</p> <ol style="list-style-type: none"><li>1. Cognitive functional therapy (CFT) i 12 uker. Hver pasient mottok et individualisert program rettet mot å endre uhensiktsmessige atferd knyttet til kognisjoner, bevegelsesmønster og livsstilsatferd. Første konsultasjon varte 1 time, og oppfølgingstimer 30-45 min. De første 2-3 konsultasjonene ble foretatt en gang i uken, deretter en konsultasjon hver 2-3 uke i løpet av 12 uker. Alle instruksjoner ble gitt skriftlig, og hjemmeøvelser ble forventet gjort daglig samt føre dagbok for gjennomførte øvelser. Gjennomsnittlig antall behandlinger: 7,7 (range 4-16, SD 2,6)</li><li>2. Manuell terapi og øvelser (MT-EX) i 12 uker. Det bestod av leddmobilisering og manipulasjonsteknikker, samt at flesteparten av pasientene (82,5 %) fikk et generelt hjemmetreningsprogram. Treningsprogrammet inneholdt øvelser for motorisk kontroll med isolerte kontraksjoner av dype abdominalmuskler. En time ble brukt for første konsultasjon og oppfølgingstimer på 30 min. Gjennomsnittlig antall behandlinger: 8 (range 3-17, SD 2,9)</li></ol> |

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Utfallsmål | <p>Utfallsmål ble evaluert ved baseline, og ved 3- og 12 mnd oppfølging</p> <ul style="list-style-type: none"> <li>• <b>Funksjon:</b> Oswestry Disability Index (ODI)</li> <li>• <b>Smerte:</b> 11-punkts numerisk smerteskala (PNRS) målt ved hele tall fra 0-10, i løpet av de siste 2 ukene</li> <li>• <b>Andre utfallsmål:</b> <ul style="list-style-type: none"> <li>- Hopkins Symptoms Checklist for angst og depresjon.</li> <li>- Fear-Avoidance Beliefs Questionnaire (FABQ).</li> <li>- Totalt bevegelsesutsalg av lumbalcolumna målt med to inklinometer metoder</li> <li>- Pasienttilfredshet målt med spørreskjema ved 3 og 12 mnd oppfølging</li> <li>- Antall sykemeldingsdager ekstrahert fra Ørebro Screening Questionnaire</li> <li>- søk om helsehjelp mellom intervensjon og 12 mnd oppfølging ble vurdert av et enkelt spørreskjema</li> </ul> </li> </ul> |
| Resultater | <p>Antall deltakere:</p> <ul style="list-style-type: none"> <li>• CFT: 62</li> <li>• MT-EX: 59</li> <li>• Deltakere som manglet oppfølging ved 12mnd: 11 av 62 (17,7 %) i CFT, og 16 av 59 (27,1 %) i MT-EX</li> </ul> <p>Effekt av behandling</p> <ul style="list-style-type: none"> <li>• <b>Funksjon:</b> CFT viste signifikant bedring ved 3 mnd og 12 mnd sammenlignet med MT-EX, <math>p &lt; 0,001</math></li> <li>• <b>Smerte:</b> CFT viste signifikant bedring ved 3 mnd og 12 mnd sammenlignet med MT-EX, <math>p &lt; 0,001</math></li> <li>• <b>Andre utfallsmål:</b> CFT viste signifikant bedre resultater sammenlignet med MT-EX</li> </ul>                                                                                                                                                                                                                     |

## RISIKO AV SYSTEMATISKE FEIL

| Bias domene               | Kilde til bias                                                                                      | Vurdering | Beskrivelse                                                                                                                                                               |
|---------------------------|-----------------------------------------------------------------------------------------------------|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Selection                 | 1. Was the method of randomization adequate? Adekvat randomiseringsmetode?                          | Ja        | Randomisering ble utført i permuterte blokker på 16                                                                                                                       |
| Selection                 | 2. Was the treatment allocation concealed? Skjult tildeling av intervensjonsgruppe? Alle utfallsmål | Ja        | Forseglete, ugjennomsiktige brev                                                                                                                                          |
| Performance and detection | 3. Was the patient blinded to the intervention? Blinding av pasient? Alle utfallsmål                | Nei       | Deltakerne fikk behandling som de var klar over. De ble fortalt at de fikk en av to behandlingsmodeller hvor man for øyeblikket ikke visste hvilken som var mest effektiv |

|                           |                                                                                                                                                        |     |                                                                                                                                                                                                  |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Performance and detection | 4. Was the care provider blinded to the intervention? Blinding av helse tilbyder? Alle utfallsmål                                                      | Nei | Fysioterapeutene som ga behandling visste hvilken intervensjonsgruppe de tilhørte, MT-EX eller CFT                                                                                               |
| Performance and detection | 5. Was the outcome assessor blinded to the intervention? Blinding av assessor? Alle utfallsmål?                                                        | Ja  | Pasientrapporterte utfallsmål: assessorer vurdert som blindet pga sammenlignbare intervensjonsgrupper. Tester blindet for fordeling, intervensjon og 3 mnd oppfølging distribuerte spørreskjema. |
| Attrition                 | 6. Was the drop-out rate described and acceptable? Akseptabel frafallsprosent? Alle utfallsmål                                                         | Ja  | Frafall er beskrevet og balansert mellom intervensjonsgruppen og kontrollgruppen, med mindre enn 30 % frafall.                                                                                   |
| Attrition                 | 7. Were all randomized participants analyzed in the group to which they were allocated? Alle deltakere analysert - intention-to treat? Alle utfallsmål | Nei | Det ble ikke foretatt intention-to-treat analyse                                                                                                                                                 |
| Reporting                 | 8. Are reports of the study free of suggestion of selective outcome reporting? Fri for selektiv rapportering av utfallsmål?                            | Ja  | Utfallsmål beskrevet i metodekapitlet er i overensstemmelse med de som er rapportert i resultatene                                                                                               |
| Selection                 | 9. Were the groups similar at baseline regarding the most important prognostic indicators? Tilstrekkelig likhet av baseline karakteristikk?            | Ja  | Behandlingsgruppene var sammenlignbare med hensyn til baseline karakteristikk, med unntak av små men signifikante forskjeller i HSCL og FABQ work                                                |
| Performance               | 10. Were co-interventions avoided or similar? Var ko-intervensjon unngått eller lik?                                                                   | Ja  | Begge grupper mottok behandling som var sammenlignbare                                                                                                                                           |
| Performance               | 11. Was the compliance acceptable in all groups? Var compliance akseptabel for alle grupper?                                                           | Ja  | Gjennomsnittlig antall behandlinger var 7,7 (SD: 2,6) for CFT, og 8,0 (SD: 2,9) for MT-EX                                                                                                        |
| Performance               | 12. Was the timing of the outcome assessment similar in all groups? Likt tidspunkt for vurdering av utfallsmål for alle grupper?                       | Ja  | Alle utfallsmål ble samlet ved baseline, 3 mnd og 12 mnd etter, for begge gruppene                                                                                                               |
| Detection                 | 13. Are other sources of potential bias unlikely? Er andre kilder til bias usannsynlig?                                                                | Ja  | Utfallsmål anses som valide, og forfatterne erklærer ingen interessekonflikter                                                                                                                   |