

# Kliniske tester ved bekkenplager

*En studie av kliniske tester og sammenheng med egenrapportert smerte og funksjon*

**Magnhild Aasen**



Masteroppgave ved Det medisinske fakultet, Seksjon for helsefag

UNIVERSITETET I OSLO

April 2007



## Forord

Denne masteroppgaven hadde sitt utspring i en hyggelig invitasjon. En effektstudie av behandling ved bekkenplager etter fødsel var i emning, og jeg fikk muligheten til å delta med klinisk undersøkelse av pasientene. Arbeidet medførte behov for mer kunnskap om tester som jevnlig benyttes i klinisk praksis. Masterstudiet i helsefag ble en gylden mulighet til å bruke noe mer tid og tankevirksomhet på dette feltet. Prosjektlederens raushet var da stor ved å gi meg mulighet til å arbeide med datamaterialet for de gjennomførte kliniske undersøkelsene.

Så først varm takk til forskningsleder Britt Stuge ved kompetansesenter for klinisk forskning, Ullevål Sykehus. Hun inviterte til dette løpet og har fulgt meg hele veien, - positiv, energisk og stadig kunnskapssøkende. Samarbeidet i forskningsprosjektet var en fornøyelse, preget av god organisering og mye munterhet. Veiledning fra en engasjert fagperson i skjæringspunktet mellom klinikk og forskning har i neste omgang vært til stor hjelp og inspirasjon! Takk også til Gitle Kirkesola for betydelig bidrag i planleggingsfasen ved valg og beskrivelse av de kliniske testene.

Professor Nina Vøllestad ved Seksjon for helsefag, Universitetet i Oslo, stilte villig opp som hovedveileder. Uten hennes klarsyn, analytiske innspill og positive innstilling ville ikke denne oppgaven vært mulig å gjennomføre. Jeg har tilbakelagt en lang etappe uten god sikt i terrenget, men konstruktiv veiledning har fått mye til å klare i innspurten. Tusen takk for særdeles lærerike samtaler!

Trivsel i studiemiljøet er også sentralt for motivasjonen når det synes langt til mål. Jeg har satt pris på engasjerte medstudenter og dyktige ansatte ved seksjon for helsefag, og har hatt god nytte av stipendiat Hilde`s velvillige bidrag med litteratur. En spesiell takk til Turid for fellesskap, samarbeid og god støtte gjennom hele studiet. Jeg har også hatt stor glede av treffende og klartenkte bidrag fra Liv-Signe i ”gamle-kollokvien”. Sist, men ikke minst, takk til min nærmeste venn og kollega Lars for utfordrende spørsmål og nyttige diskusjoner underveis!

Molde, april 2007

Magnhild Aasen



# Sammendrag

## Kliniske tester ved bekkenplager og sammenheng med egenrapportert smerte og funksjon

**Bakgrunn:** Det finnes ingen objektive tester eller standardisert prosedyre for diagnostikk av bekkenplager. Forøvrig finnes lite kunnskap om hvilke kliniske tester som er positive hos kvinner med bekkenplager og hvordan testene samsvarer med pasientenes rapportering av smerte og funksjon.

**Hensikt:** Å beskrive forekomst av positive skår ved relevante kliniske tester i en populasjon av kvinner med bekkenplager etter fødsel og belyse hvilke sammenhenger som finnes mellom disse testene og egenrapportert smerte og funksjon.

**Materiale og metode:** 81 kvinner med bekkenplager ble undersøkt ved pretest ca 3 måneder etter fødsel og ved posttest etter en 20 ukers behandlingsperiode. 14 kliniske tester ble benyttet: Active straight leg raise (ASLR), posterior pelvic pain provocation (P4), rygg- og sideliggende kompresjon, distraksjon, Patrick`s, Menell`s og Trendelenburg`s test, to symfyseprovokasjonstester, palpasjon av m. iliopsoas, lange dorsale ligament (LDL) og sacrotuberale ligament, samt mål av diastase i m. rectus abdominis. Smerteintensitet ble målt ved Visuell Analog Skala (VAS) og funksjon ved Oswestry Disability Index (ODI). Korrelasjons- og regresjonsanalyser ble anvendt.

**Resultater:** Ved pretest viste 7 av 14 kliniske tester mer enn 75 % positive skår. De øvrige varierte fra 1 til 51 %. Etter behandling som medførte signifikant redusert smerte og bedret funksjon ( $p < 0.001$ ), viste 11 tester statistisk signifikant bedring i skår ( $p < 0.02$ ) sammenlignet med pretest. ASLR viste moderat korrelasjon med VAS ( $r = 0.35$ ) og ODI ( $r = 0.41$ ) ved pretest og bedret korrelasjon ( $r = 0.51$  og  $0.55$ ) ved posttest. De øvrige testene viste gjennomgående svak korrelasjon med VAS og ODI ved pretest ( $r < 0.35$ ) og moderat korrelasjon ved posttest ( $r < 0.50$ ). ASLR, LDL og Patrick`s test kunne samlet forklare ca 40 % av variansen i VAS og ODI ved posttest.

**Konklusjon:** 14 kliniske tester viste stor spredning i andel positive skår hos en gruppe kvinner med bekkenplager etter fødsel. For de fleste testene var andelen positive skår signifikant redusert fra før til etter behandling. Sammenheng mellom de kliniske testene og egenrapportering var gjennomgående svak til moderat. Kun ASLR, LDL og Patrick`s test ble funnet å ha klar sammenheng med egenrapportert smerteintensitet og funksjon.



## Abstract

Clinical tests for examining pelvic girdle pain (PGP) and correlation to self reported pain and disability

**Background:** There are so far no well-established objective tests or procedure for examining pelvic girdle pain (PGP). There is also limited knowledge on positive scores at relevant clinical PGP tests and how self reported pain and disability correlate to these tests.

**The aim:** To describe prevalence of positive scores of relevant clinical tests in a population of women suffering from PGP and to investigate the association between these tests and self reported pain and disability.

**Material and method:** 81 women suffering from PGP were studied approximately 3 months postpartum (pre-test) and after 20 weeks of treatment (post-test). The study included 14 clinical tests: Active straight leg raise (ASLR), posterior pelvic pain provocation (P4), compression in side-lying and in supine position, distraction, Patrick's, Menell's and Trendelenburgs test, two provocation tests for the pubic symphysis, palpation of the psoas muscle, the long dorsal ligament (LDL), the sacrotuberal ligament and diastasis in the rectus abdominis muscle. Visual Analogue Scale (VAS) was used to determine pain intensity, and disability was measured by Oswestry Disability Index (ODI). The collected data were analysed by correlation and regression statistics.

**Results:** 7 out of 14 clinical tests presented positive scores higher than 75 % at pre-test. The other tests varied between 1 and 51 % positive scores. Following treatment when pain intensity and disability were significantly reduced ( $p < 0.001$ ), 11 tests showed statistical significant improvement ( $p < 0.02$ ) compared with pre-test. ASLR was moderately correlated to VAS ( $r = 0.35$ ) and disability ( $r = 0.41$ ) at pre-test, and correlation was increased at post-test ( $r = 0.51$  and  $0.55$ ). The other tests were at pre-test weakly correlated to VAS and ODI ( $r < 0.35$ ), and moderately correlated at post-test ( $r < 0.50$ ). About 40 % of the variance in VAS and ODI at post-test could be explained by ASLR, LDL and Patrick's test.

**Conclusion:** In a population of women suffering from PGP, 14 clinical tests presented wide variation in positive test scores. Most tests showed significant reduction of positive scores after treatment compared with pre-test. Correlation between clinical tests and self report was weak to moderate. ASLR, LDL and Patrick's test were the only three tests significantly associated to self reported pain and disability.





---

# Innhold

<b>1. INNLEDNING .....</b>	<b>13</b>
1.1 BAKGRUNN FOR STUDIEN .....	13
1.2 FORMÅL .....	14
<b>2. TEORI.....</b>	<b>15</b>
2.1 BEKKENPLAGER .....	15
2.1.1 Terminologi.....	16
2.2 HVORDAN DIAGNOSTISERE BEKKENPLAGER? .....	18
2.2.1 Kliniske tester.....	19
2.2.2 Kombinasjon av kliniske tester .....	27
2.2.3 Smertelokalisasjon .....	28
2.3 MÅL FOR SMERTE OG FUNKSJON .....	30
2.3.1 Smerte .....	30
2.3.2 Funksjon.....	31
2.4 VIKTIGE EGENSKAPER FOR GODE MÅLEMETODER .....	33
2.4.1 Reliabilitet.....	34
2.4.2 Validitet.....	34
2.5 FORSKNINGSSPØRSMÅL.....	36
<b>3. MATERIALE OG METODE .....</b>	<b>37</b>
3.1 REKRUTTERING AV DELTAKERE .....	37
3.2 DESIGN .....	37
3.3 KRITERIER FOR INKLUSJON OG EKSKLUSJON .....	37

---

3.3.1	<i>Kliniske tester for eksklusjon</i> .....	38
3.4	BAKGRUNNSFAKTORER.....	39
3.5	EGENRAPPORTERTE MÅL.....	40
3.5.1	<i>Smerte</i> .....	40
3.5.2	<i>Funksjon</i> .....	41
3.6	KLINISKE TESTER.....	41
3.6.1	<i>Gjennomføring av testprosedyre</i> .....	41
3.6.2	<i>Beskrivelse av kliniske tester</i> .....	42
3.7	STATISTISKE ANALYSER.....	47
3.8	ETISKE HENSYN.....	49
<b>4.</b>	<b>RESULTATER</b> .....	<b>50</b>
4.1	BAKGRUNNSFAKTORER FOR INKLUDERTE KVINNER.....	50
4.2	SMERTE OG FUNKSJON.....	51
4.2.1	<i>Endring i smerte og funksjon fra før til etter behandling</i> .....	52
4.3	KLINISKE TESTER.....	54
4.3.1	<i>Active straight leg raise (ASLR) test</i> .....	54
4.3.2	<i>Bilaterale tester</i> .....	55
4.3.3	<i>Sideuavhengige tester</i> .....	56
4.4	SAMMENHENGER.....	57
4.4.1	<i>Sammenhenger mellom kliniske tester</i> .....	57
4.4.2	<i>Sammenhenger mellom kliniske tester og egenrapportert smerte og funksjon</i> .....	60
4.4.3	<i>Testenes selvstendige forklaringsverdi for smerte og funksjon</i> .....	66
4.4.4	<i>Kombinasjon av kliniske tester – assosiasjon med smerte og funksjon</i> .....	68

---

4.5	OPPSUMMERING AV RESULTATER.....	69
<b>5.</b>	<b>DISKUSJON AV METODISKE VALG .....</b>	<b>71</b>
5.1	DESIGN .....	71
5.2	UTVALG .....	71
5.3	MÅL FOR SMERTE OG FUNKSJON .....	72
5.4	EKSKLUSJONSKRITERIER.....	74
5.5	KLINISKE TESTER .....	76
5.6	GJENNOMFØRING AV TESTPROSEDYRE .....	77
5.7	STATISTISKE ANALYSER.....	78
<b>6.</b>	<b>DISKUSJON AV RESULTATER .....</b>	<b>80</b>
6.1	BESKRIVELSE AV BEKKENPLAGENE.....	80
6.1.1	<i>Smerte og funksjon.....</i>	<i>81</i>
6.1.2	<i>Kliniske tester.....</i>	<i>82</i>
6.1.3	<i>Endringer i testskår.....</i>	<i>84</i>
6.2	SAMMENHENGER .....	85
6.3	KLINISKE TESTERS BIDRAG TIL Å FORKLARE SMERTE OG FUNKSJON.....	87
6.4	VIDERE UTFORDRINGER VED DIAGNOSTIKK AV BEKKENPLAGER .....	89
<b>7.</b>	<b>KONKLUSJON.....</b>	<b>91</b>
	<b>KILDELISTE.....</b>	<b>92</b>



# 1. Innledning

Det var sein kveld, og noen navn ble skrevet ned på en serviett. Kreative klinikere hadde stukket hodene sammen for å finne medarbeidere til et forskningsprosjekt. I ettertid fikk jeg høre det var slik jeg ble en av flere aktuelle kandidater for å delta i planlegging og gjennomføring av en studie av behandlingseffekt ved bekkenplager etter fødsel. Invitasjonen kom, og jeg tok de første prøvende steg på veien for å bli bedre kjent med forskningens verden. Noen års erfaring i klinisk praksis hadde gitt en økende interesse for hvordan vi bedre kan utrede og diagnostisere pasientenes plager. Planlegging av testprosedyre for denne effektstudien avdekte behov for å innhente mer kunnskap om eksisterende dokumentasjon for de kliniske testene som ofte benyttes ved undersøkelse av bekkenplager.

## 1.1 Bakgrunn for studien

Min arbeidsoppgave i effektstudien til Stuge og medarbeidere (2004) var å gjennomføre en standardisert klinisk undersøkelse for å inkludere eller ekskludere kvinner som meldte seg som aktuelle deltakere (Stuge, Laerum, Kirkesola, & Vollestad, 2004). De inkluderte kvinnene skulle også retestes med samme prosedyre etter en behandlingsperiode. Etter samarbeid med prosjektleder og medforfatter om utarbeidelse og utførelse av testprosedyrene, gikk jeg da inn i prosjektet som ”blindet” undersøker. Jeg hadde ikke kjennskap til hvilke behandlingsgrupper pasientene havnet i, og heller ikke inngående kjennskap til behandlingsalternativene i den planlagte effektstudien. Skår på de kliniske testene i dette arbeidet utgjorde et stort datamateriale. I ettertid så man behovet for å gjøre analyser på det aktuelle testmaterialet.

Ved oppstart av effektstudien forelå det ingen studier som kunne gi et godt klinisk bilde av kvinner med bekkenplager etter fødsel. Det var derfor vesentlig å kunne beskrive de inkluderte pasientene ut fra hvordan de skåret på kliniske tester. Det var også interessant å finne ut om det kliniske bildet som viste seg ved denne testprosedyren ville ha noe samsvar med kvinnenens egen rapportering av smerte og funksjon. Mitt ståsted var i klinisk arbeid med undersøkelse og behandling. Motivasjonen for å ta fatt på oppgaven som ble min studie, lå derfor i å søke kunnskap som kunne bidra til forbedret klinisk praksis.

## 1.2 Formål

Hvordan skal bekkenplager diagnostiseres? Dette spørsmålet har blitt stilt av mange innen utdanning, forskning og klinisk praksis. Svarene kan være ulike, og det foreligger ingen klar felles oppfatning av hvordan dette arbeidet skal gjøres. Spørsmålet ble også overordnet for denne studien.

Studien ble planlagt for å kartlegge og beskrive det kliniske bildet av en gruppe kvinner med bekkenplager etter fødsel ved hjelp av et utvalg kliniske tester. Samtidig hadde disse kvinnene rapportert egen smerteintensitet og funksjon ved bruk av standardiserte mål. Hensikten var ut fra dette samlede materialet å gjøre en deskriptiv framstilling av pasientgruppen og undersøke eventuelle sammenhenger mellom skår på kliniske tester og kvinnenenes egen rapportering av plager.

Kvinnene ble testet på to ulike tidspunkter, før og etter en behandlingsperiode. Før behandling var dette en selektert pasientpopulasjon med bekkenplager. Etter behandling hadde den samme populasjonen endret status slik at en andel var friske og en andel hadde varierende grad av plager. Studien hadde også som hensikt å beskrive endringer i skår på kliniske tester og hvilken sammenheng dette kunne ha med rapportering av smerteintensitet og funksjon.

Formålet med studien ble ut fra dette å bidra til en utvidet beskrivelse av bekkenplager og belyse hvilke sammenhenger som kunne finnes mellom kliniske tester, smerteintensitet og funksjon i det aktuelle materialet.

## 2. Teori

### 2.1 Bekkenplager

Bekkenplager forekommer hos begge kjønn, men er mest vanlig hos kvinner. Slike plager kan opptre uavhengig av graviditet, men er oftest assosiert med svangerskap og fødsel (Ostgaard, Andersson, & Karlsson, 1991; Ostgaard, Zetherstrom, Roos-Hansson, & Svanberg, 1994b; Endresen, 1995; Sturesson, Uden, & Uden, 1997). Svangerskap og barseltid har til alle tider vært en livsfase forbundet med glede og forventning. De fleste aksepterer også at denne perioden kan være forbundet med noe rygg- og bekkenplager. Mange kvinner har likevel erfart at svangerskap eller fødsel har medført plager over lengre tid og med større intensitet enn de var forberedt på. I møte med helsevesenet beskriver disse kvinnene ofte ønske eller klare forventninger om at plagene skal avta etter fødsel. Smerter og begrensninger i funksjon erfares som særlig vanskelig når dette reduserer mulighetene til omsorg for et nyfødt barn.

Vitenskapelige arbeider har dokumentert varierende forekomst av bekkenplager hos kvinner både i graviditet og etter fødsel. En studie fra 1996 viste at hver tredje gravide kvinne erfarte bakre bakkensmerter, og hver niende ryggsmertter (Ostgaard, Roos-Hansson, & Zetherstrom, 1996). Bakre bakkensmerter var mer intense i svangerskapet, og ryggsmertter mer intense og mer vanlig etter fødsel. Det er også vist ulik forekomst av plager i tiden etter fødsel. En oppfølgingsstudie av gravide viste at 71 % hadde rygg- og bekkenplager i svangerskapet. Etter 6 år hadde 16 % av disse kvinnene fortsatt plager (Ostgaard, Zetherstrom, & Roos-Hansson, 1997). En kohortestudie fra 2006 viste at 90 % av kvinner med bekkenplager i svangerskapet var friske etter fødsel, mens 10 % hadde vedvarende plager ett og et halvt år seinere (Rost, Jacqueline, Kaiser, Verhagen, & Koes, 2006). To og Wong (2003) viste høyere forekomst to år etter fødsel, i dette materialet hadde 21 % av kvinnene vedvarende ryggsmertter (To & Wong, 2003). Resultatene samsvarte med en tre års oppfølging av gravide, der 20 % av kvinner med ryggsmertter i svangerskapet fortsatt hadde smertter tre år seinere (Noren, Ostgaard, Johansson, & Ostgaard, 2002). En review viste at gjennomsnittlig 25 % av kvinner inkludert i de undersøkte forskningsarbeidene var plaget av smertter i lumbal eller bekken etter fødsel (Wu et al., 2004). Variasjonen i resultatene var likevel stor mellom de ulike studiene (fra 0.3 % til 67 %, median 22 %). Wu og medarbeidere (2004) rapporterte at gjennomsnittlig 7 % av kvinnene hadde alvorlige plager etter svangerskap.

Albert og medarbeidere (2001) gjennomførte en kohortestudie for å beskrive prognose for kvinner med ulike subgrupper av svangerskapsrelaterte bekkenleddsmerter (Albert, Godskesen, & Westergaard, 2001). Flertallet av de inkluderte kvinnene var smertefri innen en måned etter fødsel. Etter 2 år hadde 8.5 % fortsatt bekkenplager, men varighet av plagene varierte sterkt mellom subgruppene. Forfatterne konkluderte med at tre av fire subgrupper hadde svært god prognose etter fødsel. Ingen med kun symfysesmerter hadde plager 6 måneder etter fødsel, derimot hadde 21 % av kvinner med smerter fra alle tre bekkenledd fortsatt plager etter 2 år.

Felles for flere av studiene synes å være at et stort flertall av kvinner med bekkenplager i svangerskapet er kvitt sine plager i løpet av en kortere periode etter fødsel. Langtids prevalens for lumbal- og bekkenplager (2- 6 år etter fødsel) varierer fra rundt 10 % til vel 20 %. Variasjonen ser ut til å kunne avhenge av ulike faktorer. Definisjonene av rygg- og bekkenplager er ofte ulike, og de diagnostiske prosedyrene er gjennomført på mange forskjellige måter. Det er også stor variasjon i terminologi som er benyttet i studier innenfor dette området.

### **2.1.1 Terminologi**

Bekkenplager i forbindelse med svangerskap har vært beskrevet helt siden Hippokrates' tid, og er omtalt på ulike måter opp gjennom historien (Bastiaanssen, de Bie, Bastiaenen, Essed, & van den Brandt, 2005). I norske medisinske miljøer har flere termer vært foreslått og benyttet (Kogstad & Biornstad, 1990). "Bekkenløsning" ble i 1988 beskrevet av Kogstad som "stillingsforandringer og/ eller funksjonsforandringer i ileosacralledd og symfyse ved smertefulle provokasjonstester". Denne bruken av begrepet står i kontrast til forståelsen av bekkenløsning som en normal fysiologisk prosess på grunn av øket produksjon av relaxin fra rundt tiende svangerskapsuke og fram mot fødsel (Gray, 1989). I 1990 presenterte Den norske Lægeforening to andre termer (Endresen, 1995).

"Symptomgivende bekkenløsning" betegnet rygg- og bekkensmerter med gangforstyrrelser og positive provokasjonstester på bekkenets leddforbindelser de første 6 månedene etter fødsel.

"Bekkenleddssyndrom" betegnet de samme plagene dersom de vedvarte ut over 6 måneder. Disse to termene ble videre benyttet av Rikstrykdeverket som diagnoser knyttet til rettigheter om honorartakst ved fysikalsk behandling.

I internasjonal litteratur finnes også en rekke termer brukt om smerter i lumbal- og bekkenregionen. De fleste forfatterne har så langt ikke gjort noen klar distinksjon mellom lumbale smerter og smerter i bekken (Wu et al., 2004). "Back pain" er benyttet for å beskrive smerter både fra lumbal og bekken og kombinasjon av disse (Kristiansson & Svardsudd, 1996a; Ostgaard et al., 1997; Noren et al.,



---

2002). I noen studier er kliniske tester benyttet for differensiering av plagene (Kristiansson et al., 1996a; Noren et al., 2002), i andre studier er det kun benyttet spørreskjema for kartlegging av smerter (To et al., 2003) eller spørreskjema i kombinasjon med VAS smertemåling (Ostgaard et al., 1997; Mousavi, Parnianpour, & Vleeming, 2007).

Östgaard og medarbeidere (1994b) påpekte behovet for å skille mellom ”low back pain” (LBP) og ”posterior pelvic pain” (PPP) ved lumbal- og bekkenplager relatert til svangerskap. Bakre bekkensmerter (PPP) ble i denne studien beskrevet som karakteristiske smerter distalt og lateralt for L5, med eller uten radierende smerter til dorsalside av lår og kne. De viste til at begreper som ”sacroiliac joint dysfunction”, ”pelvic girdle relaxation” og ”pelvic insufficiency” var vanskelig å identifisere. To år seinere ble det også vist til at få studier hadde differensiert mellom rygg smerter og bakre bekkensmerter i graviditet, og ingen studier hadde presentert oppfølgingsdata med hensyn til smertetyper (Ostgaard et al., 1996).

Etter hvert ble det presentert flere lignende termer for bekkenplager med definisjoner mer og mindre knyttet til graviditet. ”Posterior pelvic pain after pregnancy” (PPPP) ble definert som bakre bekkensmerter etter fødsel med debut i svangerskap eller innen tre uker etter fødsel, uten indikasjon for noen spesifikk sykdom som kunne forklare plagene (Mens, Vleeming, Snijders, Koes, & Stam, 2002a). Mens og medarbeidere mente det kunne diskuteres om PPPP var et spesifikt syndrom eller kun uspesifikke lumbal/bekkenmerter med debut i svangerskap eller ved fødsel. De valgte i sin definisjon å utelukke smerter fra symfyseområdet. Bastiaanssen og medarbeidere (2005) valgte annerledes og brukte begrepet ”pregnancy-related pelvic girdle pain” (PPGP) (Bastiaanssen et al., 2005). Deres definisjon ble begrenset til smertedebut i svangerskap i nedre del av rygg, symfyse, lyske og/ eller radierende smerter til legg. En review av Wu og medarbeidere (2004) viste til evidens for å gjøre diagnostisk differensiering mellom lumbale plager og bekkenplager i forbindelse med svangerskap. Termene ”pregnancy-related pelvic girdle pain” (PPP) og ”pregnancy-related low back pain” (PLBP) ble anbefalt for videre bruk i klinikk og forskning.

”Pelvic girdle pain” (PGP) ble i 1999 definert som bekkensmerter influert av posisjon og bevegelse, lokalisert posterørt så vel som anteriørt (Mens, Vleeming, Snijders, Stam, & Ginai, 1999). En videre subgruppering av PGP ble gjort av Albert og medarbeidere (Albert, Godskesen, & Westergaard, 2000). Bekkenplager ble delt i fem subgrupper ut fra smertelokalisering og kliniske funn. Disse subgruppene var smerter fra alle tre bekkenledd, smerter kun fra symfyse, ensidig iliosacral syndrom, bilateralt iliosacral syndrom, og en gruppe for smerter fra ett eller flere ledd med inkonsistente objektive funn. PGP har etter hvert blitt en hyppig brukt betegnelse for smerter

lokalisert mellom lumbosacral overgang og glutealfold, inkludert smerter i symfyseområde og somatisk refererte smerter distalt for glutealfold (Laslett, 2005).

Termen er bearbeidet og endret, og følgende definisjon for PGP ble i 2004 foreslått av the World Group 4, European Cost Commission (Vleeming, Albert, Ostgaard, Stuge, & Sturesson, 2004):

*”Pelvic girdle pain (PGP) generally arises in relation to pregnancy, trauma or reactive arthritis. Pain is experienced between the posterior iliac crest and the gluteal fold, particularly in the vicinity of the sacroiliac joints (SIJ). The pain may radiate in the posterior thigh and can also occur in conjunction with / or separately in the symphysis. The endurance capacity for standing, walking and sitting is diminished. The diagnosis of PGP can be reached after exclusion of lumbar causes. The pain or functional disturbances in relation to PGP must be reproducible by specific clinical tests.”*

Denne forståelsen av PGP er så langt det nærmeste fagmiljøene har kommet en internasjonal konsensus om definisjon av bekkenplager. Avgrensningen fra European Cost Commission hadde til hensikt å ekskludere patologiske tilstander innen gynekologi og urologi samt plager med årsak i lumbalcolumna. På denne måten skulle termen PGP relateres kun til bekkenplager med årsak i muskel-skjelettsystemet. Samtidig ga definisjonen ingen beskrivelse av subgrupper innenfor området bekkenplager.

Uklarheter både i terminologi og underliggende diagnostiske kriterier er gjennomgående i vitenskapelige arbeider på dette feltet. Heiberg Endresen valgte i 1995 følgende tittel for en artikkel om temaet: ”Bekkenløsning- en tilstand med mange navn og uklare kriterier” (Endresen, 1995). Utsagnet har mer enn ti år seinere ennå ikke mistet sin aktualitet. Uklarheter angående diagnostiske kriterier og terminologi er fortsatt en stor utfordring både innenfor Rikstrygdeverkets forskrifter, undervisning og forskning, og ikke minst i klinisk praksis.

I fortsettelsen av denne oppgaven brukes ”bekkenplager” som norsk betegnelse for PGP.

## 2.2 Hvordan diagnostisere bekkenplager?

I klinisk arbeid er det tradisjon for å benytte anamnese og kartlegging av smerteutbredelse i tillegg til ulike kliniske tester ved utredning av bekkenplager. Den samlede informasjonen benyttes av den enkelte undersøker i en klinisk resonneringsprosess for å kunne stille en diagnose. Det finnes fortsatt ingen objektive målemetoder som kan påvise, diagnostisere eller differensiere bekkenplager. Radiologiske utredninger og blodprøver kan være hjelpemidler i differensialdiagnostikk, men gir alene ingen klare svar på årsak til eller lokalisering av plagene. Det kan også være noe

---

sammenfallende smerteutbredelse for plager som er relatert til gynekologi og plager relatert til muskel-skjelettsystemet (Gray, 1989). Med dette grunnlaget er både pasienter, klinikere og forskere avhengige av den kliniske undersøkelsen for å best mulig utrede denne type plager. Kliniske tester benyttes i stor grad for å avdekke funksjonsforandringer og lokalisere smertegivende vevsstrukturer (Kirkesola & Solberg, 1998).

### **2.2.1 Kliniske tester**

I klinisk praksis er det benyttet en mengde forskjellige tester for å diagnostisere og differensiere plager i lumbal- og bekkenområdet. Disse kliniske testene har etter hvert blitt delt i to kategorier. Palpasjonstester er benyttet for å vurdere posisjon og grad av bevegelse, og smerteprovokasjonstester er benyttet for å reprodusere symptomer (Van der Wurff, Hagmeijer, & Meyne, 2000a; Albert et al., 2000; Laslett, Aprill, McDonald, & Young, 2005). Palpasjonstester er generelt vist å være lite reliable, en del provokasjonstester er dokumentert som mer pålitelige (Van der Wurff et al., 2000a; Robinson et al., 2007). Et problem ved vurdering av kliniske tester er at mange av de aktuelle studiene er vurdert til å være av dårlig metodisk kvalitet (Van der Wurff et al., 2000a; Van der Wurff, Meyne, & Hagmeijer, 2000b; Zelle, Gruen, Brown, & George, 2005).

Ved planlegging av effektstudien til Stuge og medarbeidere (2004) ble en gjennomgang av aktuelle vitenskapelige studier i kombinasjon med klinisk erfaring og skjønn lagt til grunn for valg av kliniske tester. De valgene som da ble gjort, ble også avgjørende for hvilket materiale som presenteres i denne kohortestudien. I fortsettelsen av dette kapittelet vil det bli redegjort for de ulike kliniske testene som ble benyttet.

#### ***Posterior pelvic pain provocation (P4) test***

Posterior pelvic pain provocation (P4) test ble i 1994 introdusert som en smerteprovokasjonstest for å differensiere mellom smerter i nedre lumbalcolumna og bakre bekken hos gravide kvinner (Ostgaard, Zetherstrom, & Roos-Hansson, 1994a). Testen var også beskrevet i 1991 uten at den var navngitt som P4 (Roos-Hansson & Zetherstrom, 1991). I følge disse forfatterne var testen benyttet siden 1983 for å differensiere mellom lumbale smerter og bakre bekkensmerter hos gravide. Östgaards studie viste til sterk korrelasjon mellom positiv P4-test og en historie med bakre bekkensmerter. Det ble understreket at testen ikke var spesifikk med tanke på anatomiske strukturer og ikke utelukket andre typer ryggsmertor. Det ble derfor anbefalt å alltid bruke P4 i kombinasjon med en grundig klinisk undersøkelse av ryggen. Vleeming og medarbeidere (2002) hevdet at positiv P4-test delvis kunne forklares ut fra foroverrotasjon av ilium i testprosedyren (Vleeming, de Vries, Mens, & van

Wingerden, 2002). Dorsal kompresjon gjennom femurs lengdeakse vil gi en aksial kraft gjennom ilium som medfører foroverrotasjon og smerteprovoserende stress av long dorsal ligament (LDL) på samme side. Sensitivitet for P4 er dokumentert å være 0.80, spesifisitet 0.81 (Ostgaard et al., 1994a). Det er også vist intertester-enighet for P4 på 0.84 for høyre og 0.87 for venstre side (Robinson et al., 2007).

”The femoral compression test” ble benyttet i en studie av Kristiansson og Svärdsudd som en av flere tester for å se på forholdet mellom egenrapportert smerte og status for rygg og bekken vurdert ved kliniske tester (Kristiansson et al., 1996a). Denne testen ble beskrevet tilsvarende som P4 med unntak av fiksering av motsatt sides ilium, og er av andre forfattere oppfattet å være samme test (Albert et al., 2000). Sensitivitet for testen var 0.69, spesifisitet 0.90. I denne studien ble undersøkerens kompresjonskraft ved utførelse av testen angitt å være 50-150 Newton.

P4 ble også evaluert som del av en klassifiseringsprosedyre ved svangerskapsrelaterte bekkenleddssmerter (Albert et al., 2000). Denne studien viste at P4 hadde sensitivitet på 0.90 for smerter fra alle tre bekkenledd, 0.84 for ensidige iliosacralledds-smerter (IS-smerter), og 0.93 for bilaterale IS-smerter hos gravide. Spesifisitet ble angitt til 0.98. Albert og medarbeidere presenterte altså høyere sensitivitet for P4 enn Kristiansson og medarbeidere, og de mente dette sannsynligvis kunne forklares ut fra at deres studie hadde strammere kriterier for smertelokalisasjon. Testen var i Alberts studie definert positiv kun ved smerte i eller i direkte tilknytning til leddene.

”Thigh thrust” er etter beskrivelsen også tilsvarende test som P4, men er utført med lett adduksjon av pasientens hofte (Laslett, Young, Aprill, & McDonald, 2003). Dreyfuss og medarbeidere undersøkte denne testen som en del av et testbatteri og sammenlignet med anesthesiinjeksjoner (Dreyfuss, Michaelsen, Pauza, McLarty, & Bogduk, 1996). Testen viste seg å ikke være egnet til å diagnostisere smerte spesielt fra IS-leddene. Laslett undersøkte også sensitivitet for thigh thrust, men konkluderte med helt andre resultater enn Dreyfuss og medarbeidere (Laslett et al., 2005). Thigh thrust var den mest sensitive av fem provokasjonstester. Denne testen ble anbefalt benyttet først ved diagnostikk av patologi i IS-ledd, sammen med distraksjonstest (separasjon).

”POSH” (posterior shear) er en fjerde benevnelse for tilsvarende utførelse ved provokasjon av bakre strukturer i bekkenet. I en studie av Broadhurst og Bond ble sensitivitet for denne testen beregnet til 0.80 og spesifisitet til 1.0 (Broadhurst & Bond, 1998).

P4 er også undersøkt som del av en testkombinasjon med ulike andre tester. Dette belyses nærmere i kapittel 2.2.2.

### *Active straight leg raise (ASLR) test*

ASLR ble i 1999 presentert som en funksjonstest (Mens et al., 1999). Ved aktivt strakt beinløft i ryggliggende stilling skulle funksjonen ved overføring av krefter fra lumbosacralcolumna til underekstremitetene vurderes. Testen ble skåret på en skala fra 0 til 4 på hver side. Dette ble gjort etter en samlet vurdering av pasientens rapportering av svakhet, smerte eller sideforskjell, undersøkers observasjon av hastighet og bevegelseskvalitet på strakt beinløft, samt vurdering av pasientens kroppsspråk og verbale uttrykk. Beinløftet var beskrevet å skulle være 5 cm opp fra benken.

I en studie fra 2001 var utførelse og skåringsprosedyre for ASLR blitt noe endret (Mens, Vleeming, Snijders, Koes, & Stam, 2001). Beinløftet skulle være 20 cm opp fra benken. Undersøkers deltakelse i skåringsprosedyren var utelatt. Skala for skåring var endret til 0-10 (0-5 på hvert bein) etter pasientens vurdering av hvor tungt det var å løfte beinet, 0 tilsvarte negativ test. Sensitivitet og spesifisitet for ASLR ble sammenlignet med P4, og forfatterne konkluderte med at reliabilitet, sensitivitet og spesifisitet for ASLR var høy. Test-retest reliabilitet var 0.87. Best balanse mellom sensitivitet (0.87) og spesifisitet (0.94) ble funnet når skår 0 var definert som negativ og 1-10 som positiv. Sensitivitet for ASLR (0.87) ble da funnet å være høyere enn for P4 (0.69). Forfatterne mente samtidig at en høyere cut-off muligens var å foretrekke i en del kliniske situasjoner. Ved cut-off mellom 2 og 3 ble sensitiviteten for ASLR endret til 0.66 og spesifisiteten til 1.0. Korrelasjon mellom ASLR og P4 i denne studien var 0.27, og forfatterne mente den lave korrelasjonen skyldtes at testene så ut til å måle ulike aspekter ved bekkenplager.

En validitetsstudie fra 2002 konkluderte med at ASLR også var valid for å måle alvorlighetsgrad av bakre bekkenplager etter fødsel (Mens et al., 2002a). Alvorlighetsgrad ble her definert ut fra sykehistorie, tid før smerteøkning ved stående, gange, sykling, sittende og liggende, og ved et utvalg provokasjonstester. Resultatene fra denne studien viste korrelasjon på 0.70 mellom ASLR og Quebeck Back Pain Disability Scale (QBPD).

En annen studie fra 2002 som vurderte respons på utfallsmål ved behandling av bekkenplager etter fødsel, konkluderte med at ASLR var nyttig for å gi informasjon om bedring eller forverring av plagene (Mens et al., 2002b). Forfatterne mente testen kunne brukes som standard i framtidig forskning.

Motoriske kontrollstrategier ved ASLR ble undersøkt hos en gruppe kvinner og menn som hadde klinisk diagnostiserte iliosacralledssmerter og sammenlignet med en kontrollgruppe (O'Sullivan et al., 2002). Denne pasientgruppen hadde smertedebut etter traume, plagene var ikke relatert til

svangerskap. Det ble observert endret kinematikk av diafragma og bekkenbunn under utførelse av ASLR, og avvik i respirasjonsmønster var assosiert med disse forandringene.

Skår ved ASLR er også sammenlignet med målinger av bevegelighet (Doppler vibrasjonsmål) i IS-leddene. Ved fiksering med belte over bekkenet fanget ASLR opp de samme endringene i bevegelighet som Doppler-målingene. Forfatterne mente ut fra dette at ASLR kunne erstatte disse målingene for vurdering av endret bevegelighet i IS-leddene (Mens, Damen, Snijders, & Stam, 2006).

### *Long dorsal ligament (LDL) test*

Patofysiologi ved bekkensmerter ville i følge Vleeming og medarbeidere (2002) kunne belyses ytterligere ved å legge til en enkel smertetest av LDL. Dette ligamentet har utspring på posteriere spina iliaca superior (PSIS) og nærliggende ilium med forløp til laterale kant av tredje og fjerde sakrale segment, og av og til også til femte segment. LDL har forbindelse med aponeurosen til erector spinae, bakre lag av thoracolumbale fascie, overfladiske lag av gluteus maximus og en del av sacrotuberale ligament. Nutasjon av sacrum reduserer tensjon i LDL, konternutasjon øker tensjon (Vleeming et al., 1996). Disse ligamentstrukturene er også benevnt som ligamentum iliosacrale longitudinale posterior (Gray, 1989).

En studie fra 1998 indikerte at LDL ofte var palpasjonsømt ved bakre bekkenplager etter fødsel (Vleeming, Mens, de Vries, van Wingerden, & Pool, 1998). Forfatterne stilte spørsmål om dette kunne være et resultat av belastning av LDL på grunn av framoverrotert ilium og dermed konternutert sacrum. Testen ble kalt smerteprovokasjonstest og utført ved palpasjon rett caudalt for PSIS. Den ble skåret negativ når pasienten ikke indikerte smerter og positiv når smerter ble klart indikert. Skåringsprosedyren var endret i en seinere studie fra 2002, her ble smerte ved palpasjon gradert på en skala fra 0 (ingen smerte) til 3 (utholdelig smerte) på hver side. Sensitivitet for LDL ble beregnet for kvinner med bekkenplager etter fødsel ut fra kriteriene positiv ASLR eller P4 på minst en side. Resultatet viste sensitivitet på 0.86. Ved strengere kriterier for ALSR (minimum 3 poeng) og P4 (minst 2 på smerteskala 0-3) var sensitivitet for LDL 0.98. Friske var ikke inkludert i studien, og spesifisitet for testen ble derfor ikke beregnet. Test av LDL på gravide er vist å ha 97 % intertester-enighet og kappa-verdi 0.34 (Albert et al., 2000).

Kristiansson og medarbeidere (1996a) viste at provokasjonstester hadde større diskriminerende evne i nedre lumbalcolumna og bekken enn i øvre deler av columna. Smerteprovokasjon over PSIS var den kliniske testen som viste høyest prediktiv verdi i bekkenregionen. Det er mulig at palpasjon av

dette området samsvarte helt eller delvis med test av LDL, men det ble ikke gjort anatomiske distinksjoner mellom PSIS og LDL i studien (Vleeming et al., 2002). Slik distinksjon ble heller ikke gjort av Dreyfuss og medarbeidere (1996) som gjennomførte en studie for å identifisere nyttige kliniske tester til diagnostikk av patologi i IS-ledd. Anestesiinjeksjoner ble vurdert som den mest pålitelige metoden for å påvise slik patologi. Ømhet over sulcus sacralis rett medialt for PSIS var den mest sensitive faktoren for å predikere resultatet av denne type injeksjoner, men testen hadde lav spesifisitet.

### *Provokasjonstest av symfysis pubis*

Provokasjonstest av symfysis pubis er beskrevet som palpasjon av hele forsiden av symfyse og har til hensikt å avdekke plager fra denne leddforbindelsen (Albert et al., 2000). Albert og medarbeidere registrerte testen som positiv dersom palpasjon utløste smerte som vedvarte mer enn 5 sekund. Dette arbeidet viste at palpasjon av symfysis pubis og Trendelenburg's test var de mest sensitive testene ved smerter i symfyse hos gravide. Sensitivitet for provokasjon av symfyse var 0.60 ved symfysemerter og 0.81 ved smerter fra alle tre bekkenledd. Spesifisitet var 0.99 ved begge disse tilstandene. Også i en tidligere studie ble direkte trykk mot symfyse benyttet på gravide med ryggplager for å vurdere om det var smerter tilstede (Kristiansson et al., 1996a). Sensitivitet for testen ble da beregnet til 0.87, spesifisitet 0.85. Ingen kombinasjoner av andre tester viste seg å være bedre.

### *Modifisert Trendelenburgs test*

Trendelenburgs test har til hensikt å teste kraft i m. gluteus medius og bedømme hoftens stabilitet ved stående på ett bein (Hoppenfeld, 1976; Kirkesola et al., 1998). I studien til Albert og medarbeidere (2000) ble Trendelenburg's test brukt differensialdiagnostisk for å utelukke kraftsvikt i glutealmuskulatur. Dersom testen provoserte smerter i bekkenleddene, ble den brukt som klassifisering til subgrupper ut fra smerter i symfyse, ett IS-ledd, begge IS-ledd eller alle tre bekkenledd. Sensitivitet for denne testen var 0.60 for smerter fra alle tre bekkenledd og 0.62 for symfysemerter, spesifisitet var 0.99. Utførelsen av Trendelenburgs test er også blitt benyttet for registrering av smerter i symfyse og er hevdet å ha diskriminerende evne hos gravide (Wormslev et al., 1994).

### *Kompresjon*

Kompresjon har i følge Laslett og medarbeidere (1994) til hensikt å gi strekk på dorsale iliosacrale ligamenter eller komprimere ventrale del av iliosacralleddet. De konkluderte med at kompresjon var

en reliabel test for diagnostisering av IS-ledd (Laslett & Williams, 1994). En seinere studie viste at kompresjonstest burde benyttes i utredning av symptomatiske IS-ledd dersom ikke thigh thrust eller distraksjon var positive (Laslett et al., 2005). Testen var i følge dette arbeidet et tredjevalg ved utredning av symptomer fra IS-leddene, og både sensitivitet og spesifisitet var 0.69. I begge disse studiene ble testen utført med pasienten i sideliggende stilling. Dersom kompresjon var positiv sammen med enten distraksjon eller thigh thrust, var ytterligere testing av IS-leddene ikke nødvendig.

Robinson og medarbeidere (2007) viste intertester-enighet for kompresjon i sideliggende stilling på 0.82 ved test på høyre side og 0.88 ved test på venstre side. Potter og Rohststein (1985) fant også intertesterreliabilitet på 0.76 for denne testen (Potter & Rothstein, 1985), men studien ble kritisert for begrenset generaliserbarhet på grunn av lavt antall inkluderte og manglende justering for tilfeldig enighet (Freburger & Riddle, 2001). Kristiansson og medarbeidere (1996a) fant lav sensitivitet (0.07) og høy spesifisitet (0.99) for kompresjon i ryggliggende stilling.

Ved sammenligning mellom tester og diagnostiske injeksjoner for IS-ledd, viste Maigne og medarbeidere (1996) at kompresjonstesten ikke var noen nyttig prediktor for IS-smerter (Maigne, Aivaliklis, & Pfefer, 1996). Mc. Combe og medarbeidere (1989) viste til lav reliabilitet for testen utført i ryggliggende stilling, dette framkom ved sammenligning av testresultat mellom ortopedisk kirurg og fysioterapeut og uten standardisert tesprodyre (McCombe, Fairbank, Cockersole, & Pynsent, 1989). De anbefalte testen brukt med forsiktighet. Albert og medarbeidere (2000) viste at kompresjon og separasjon hadde lavere sensitivitet enn andre smerteprovokasjonstester, og gikk et steg videre ved å anbefale å forkaste disse testene til fordel for mer reliable tester som Menell's, P4 og Patrick's test.

### *Distraksjon*

Distraksjon har til hensikt å provosere ventrale iliosacrale ligamenter og kapsel samt komprimere dorsale strukturer i IS-leddene (Laslett et al., 1994; Levin, Nilsson-Wikmar, & Stenstrom, 2005). Distraksjon var den mest spesifikke testen (0.81) og hadde høyest positiv prediktiv verdi (0.60) av fem undersøkte tester i Lasletts studie fra 2005, sensitivitet var 0.60. Den ble anbefalt brukt sammen med thigh thrust. Dersom disse to testene var positive, hevdet forfatterne at det var unødvendig med flere tester for utredning av smerter i IS-leddene. I tidligere studier var høyeste rapporterte sensitivitet for distraksjon 21 % (Levin & Stenstrom, 2003). Det er også vist høy intertesterreliabilitet (0.94) for distraksjon (Potter et al., 1985). Kristiansson og medarbeidere (1996a) fant testen nyttig for diskriminerende diagnostikk i sacralområdet. For distraksjon alene var det lav sensitivitet (0.15) og



høy spesifisitet (0.98). Robinson og medarbeidere (2007) fant intertester-enighet på 0.82 for denne testen. Det er også funnet indikasjoner for at standardisering av kraft og varighet av distraksjonstesten kan være av betydning for reliabilitet og sensitivitet (Levin, Nilsson-Wikmar, Stenstrom, & Lundeberg, 1998; Levin, Nilsson-Wikmar, Harms-Ringdahl, & Stenstrom, 2001; Levin et al., 2003). Her ble det anbefalt 20 sekunders varighet på testen dersom den ikke måtte avbrytes på grunn av smerter. Kraft ved utførelse av testen er beskrevet til 250-300 Newton (Kokmeyer, Van der, Aufdemkampe, & Fickenscher, 2002). Distraksjon er vist å være mer pålitelig når den er benyttet som del av en testkombinasjon enn som enkeltstående test (Kristiansson et al., 1996a; Robinson et al., 2007).

Som for kompresjon, viste annen litteratur også til lav reliabilitet (McCombe et al., 1989), sensitivitet (Albert et al., 2000) og manglende prediktiv verdi (Maigne et al., 1996) for distraksjons-testen. Samtidig ble det i en review-studie vist til dokumentasjon som støtter bruk av både kompresjons- og distraksjonstest for å identifisere dysfunksjon i IS-ledd (Freburger et al., 2001).

### *Patrick`s test*

Patrick`s test har til hensikt å stresse IS-leddet ved fleksjon, abduksjon og utadrotasjon av hofteleddet (Dreyfuss et al., 1996). Testen reflekterer tilstanden for IS-leddet dersom hoftebevegeligheten er normal (van Deursen, Patijn, Ockhuysen, & Vortman, 1990). Det er også beskrevet at Patrick`s test gir foroverrotasjon av ilium. Dermed reproduseres ikke bare smerter fra IS-leddet. Dysfunksjon av iliolumbale ligamenter og lumbosacralt bevegelsessegment vil også kunne gi smerter (Broadhurst et al., 1998).

Testen er beskrevet med ulike utførelser og tolkning. Noen har definert positiv test kun ved smerteprovokasjon (Slipman, Sterenfeld, Chou, Herzog, & Vresilovic, 1998), noen kun ved nedsatt bevegelighet (van Deursen et al., 1990). Andre har definert positiv test både ved smerter i dorsale region rundt IS-leddet på testsiden og ved nedsatt bevegelighet med øket motstand i ytterstilling av bevegelsen (Strender, Sjoblom, Sundell, Ludwig, & Taube, 1997).

Patrick`s test viste akseptabel intertester enighet (0.82), men ble vurdert som ikke valid ved diagnostikk av iliosacralleddssmerter i en studie av Dreyfuss og medarbeidere (Dreyfuss et al., 1996; Van der Wurff et al., 2000a). Strender og medarbeidere (1997) fant at enighet ved testen var 0.92, men at prevalensen for patologiske tegn var mindre enn 0.1. En studie av Van Deursen og medarbeidere viste både lav prediktiv verdi og lavere enighet mellom undersøkernes enn ved ren tilfeldighet. Slipman og medarbeidere (1998) konkluderte med at testen verken isolert eller i

kombinasjon med andre tester kunne verifisere en diagnose for IS-ledd. Maigne og medarbeidere (1996) hevdet at den ikke var nyttig som prediktor for IS-smerter.

I motsetning til de nevnte studiene, hevdet Broadhurst og Bond (1998) at Patrick`s test sammen med to andre provokasjonstester og smertetegning var nyttig i diagnostikk av IS-ledd. De fant høy sensitivitet (0.77) for Patrick`s test, og spesifisitet var 1.0. Arbeidet ble seinere kritisert for mangelfull design (Van der Wurff et al., 2000b). Studien til Albert og medarbeidere (2000) viste samsvar med resultatene til Broadhurst og Bond angående spesifisitet (0.99), samt sensitivitet varierende mellom 0.70 (for bekkenleddsyndrom) og ca 0.40 (for ensidig og bilateralt IS-syndrom). Robinson og medarbeidere (2007) fant også inter-tester enighet for Patrick`s test på 0.80 på høyre side og 0.74 på venstre side. Reliabiliteten var bedret når testen ble benyttet i kombinasjon med (tre eller fem) andre smerteprovokasjonstester.

### *Menell`s test*

Menell`s test har til hensikt å gi kompresjon og traksjon i hoftleddet. Kraften påføres gjennom strakt bein og gir dermed en sagittal bevegelse i bekkenet (Albert et al., 2000). Testen er evaluert i få studier. Albert og medarbeidere (2000) viste til spesifisitet på 1.0 og sensitivitet fra 0.70 for bekkenleddsyndrom til 0.65 for bilateralt IS-syndrom og 0.54 for unilateralt IS-syndrom. Testen ble anbefalt brukt for diagnostikk sammen med ”thigh thrust” og Patrick`s test. En studie av Rantanen og Airaksinen fra 1985 viste sensitivitet på 0.50 for Menell`s test, men studien ble i en seinere review vurdert som ikke valid (Van der Wurff et al., 2000b).

### *Palpasjon av m. iliopsoas*

Denne palpasjonstesten har til hensikt å avdekke smerte eller ømhet i m. iliopsoas, og er tidligere benyttet av Wormslev og medarbeidere (1994). I deres studie ble palpasjon utført medialt for spina iliaca anterior superior (SIAS), og testen var definert positiv dersom det var ømhet ved palpasjon av muskelbuken. Testen viste diskriminerende evne (kappa 0.60 på høyre side og 0.50 på venstre side) mellom friske og gravide med bekkensmerter. Forfatterne antydte at et diagnostisk kriterium for bekkenløsning kunne være to positive tester av følgende: palpasjon av m. iliopsoas, palpasjon av symfyisen, stående symfysetest, Patrick`s test og palpasjon av ligamentum sacrotuberale. Et tverrfaglig ekspertpanel vurderte også ømhet i m. iliopsoas som en av 20 aktuelle tester for utredning av IS-smerter, men her kom testen ikke blant de 12 som ble videre undersøkt (Dreyfuss et al., 1996).

### *Mål av diastase i m. rectus abdominis*

Mål av diastase i m. rectus abdominis ble også benyttet av Wormslev og medarbeidere (1994). Testen ble definert positiv dersom diastase hos en pasient i ryggliggende stilling kunne palperes med en bredde over 0.5 cm. Denne registreringen viste ingen diskriminerende evne mellom friske og gravide med bekkensmerter.

### *Palpasjon av lig. sacrotuberale*

Ømhet ved palpasjon av lig. sacrotuberale er vist å ha diskriminerende evne mellom friske og gravide med bekkensmerter (Wormslev et al., 1994). Intertester enighet viste kappaverdi på 0.41 på venstre side og 0.54 på høyre side. Testen var en av fem som kunne være aktuell som diagnostisk kriterium ved bekkenløsning (se over, palpasjon av m. iliopsoas).

## **2.2.2 Kombinasjon av kliniske tester**

I klinisk praksis har det tradisjonelt sett vært benyttet en rekke kliniske tester for å utrede, diagnostisere og vurdere prognose for pasienter. Ulike testkombinasjoner har vært benyttet avhengig av utdanningstradisjon, fagmiljø og tilstanden til den enkelte pasient. Klinikere har også på ulike måter blitt opplært til resonneringsprosesser rundt tolkning av aktuelle testresultater.

Laslett og medarbeidere (2003) sammenlignet bruk av provokasjonstester alene og bruk av disse testene i kombinasjon med en klinisk resonneringsprosess ved undersøkelse av IS-ledd (Laslett et al., 2003). Arbeidet viste at den diagnostiske presisjonen var større ved en klinisk undersøkelse sammen med en resonneringsprosess enn ved provokasjonstester alene. Van Deursen og medarbeidere (1990) fant tidligere at den diagnostiske verdien av seks kliniske tester var overraskende lav, og de antydte også at verdien av testene ville vært større i kombinasjon med annen informasjon i den diagnostiske prosedyren. En review fra 2001 konkluderte med at kombinasjon av smerteprovokasjonstester og pasientens beskrivelse av smertelokalisering syntes å være nyttig for å avdekke dysfunksjon i IS-ledd (Freburger et al., 2001).

Studier har fokusert på ulike kombinasjoner av kliniske tester for bekkenplager. Mange av disse ble vurdert i en review fra 2005, der det ble anbefalt å kombinere tester for å reliabelt diagnostisere IS-ledd som smertekilde (Zelle et al., 2005).

Kristiansson og Svärdsudd (1996a) hevdet også at beste måten å diskriminere mellom kvinner med og uten lumbal- og bekkensmerter, var å benytte en kombinasjon av kliniske tester.

Testkombinasjonen som i følge deres arbeid hadde størst diskriminerende evne for lumbosacrale smerter, var kompresjon av femur (P4), palpasjon av sacrospinale ligamenter, PSIS og lumbal, samt test av lumbale bevegelser. De mest nyttige testene for sacralområdet var kompresjon av femur (P4), palpasjon av sacrospinale ligamenter, PSIS og lumbal, test av lumbale bevegelser og distraksjon. Disse testkombinasjonene hadde større diskriminerende evne enn enkelttester både ved lumbale og sacrale smerter. Reliabilitet for en kombinasjon av provokasjonstester for smerter fra IS-leddene ble også undersøkt av Robinson og medarbeidere (2007). Cluster på både tre og fem smerteprovokasjonstester ble vist å ha god inter-tester enighet, henholdsvis 0.89 - 0.91 for testene P4, Patrick's test og unilateral innadrotasjon av hofte, og 0.76 – 0.90 for testene distraksjon, P4, Patrick's test, og uni- og bilateral innadrotasjon av hofte. Resultatet av disse testkombinasjonene var bedre enn for den enkelte test alene. Samtidig ble behovet for ytterligere validering av testene understreket.

Laslett og medarbeidere (2003) undersøkte en kombinasjon av fem smerteprovokasjonstester. Disse testene var thigh thrust, kompresjon, distraksjon, Gaenslen's test og sacral thrust. Sensitivitet for denne testkombinasjonen var 0.91, spesifisitet var 0.87 ved identifisering av symptomatiske IS-ledd.

En seinere studie av Laslett og medarbeidere (2005) viste at thigh thrust og distraksjon burde brukes først ved undersøkelse av IS-ledd. Dersom disse var positive, var det unødvendig med flere tester. Dersom bare en av disse var positiv, ble det anbefalt å legge til kompresjon, og eventuelt sacral thrust dersom kompresjon var negativ. Thigh thrust var en av 12 tester som også ble undersøkt av Dreyfuss og medarbeidere (1996). De konkluderte med at kliniske tester og anamnese ikke var pålitelig i diagnostikk av IS-ledd, og hevdet at blokade var eneste måten å stille en slik diagnose på.

De aktuelle studiene gir altså ulike anbefalinger for kombinasjon av kliniske tester ved utredning av bekkenplager. Felles konklusjon for flere av disse er likevel at et utvalg kliniske tester bør sammenholdes med klinisk resonnering og smertelokalisasjon.

### **2.2.3 Smertelokalisasjon**

Bekkenplager kan ut fra klinisk erfaring synes å ha varierende smerteutbredelse. Med utgangspunkt i anatomiske sammenhenger kan smerter lokalisert til dorsale områder av bekkenet ha utspring i funksjonsforandringer i et større område. I tillegg til lokal smerteårsak, kan smerter relateres til dermatom, myotom, sklerotom og viscerotom (Gray, 1989; Netter, 1991; Kuchera & Kuchera, 1993). Hudområdet mellom L5 og glutealfold har dermatom innervert fra L2 til S4. Myotom for samme område relateres til segmentene fra L1 til S2 (Gray, 1989). Iliopsoas har innervasjon fra L1 og L2, obturatorius eksternus fra L4, glutealmuskulatur og obturatorius internus fra L5 og S1, piriformis har

---

innervasjon fra S1 og S2. Sympaticus-innervasjon til strukturer i bekkenet kommer fra segmentene Th11 til L2, og parasympatisk innervasjon fra S2 til S4 (Forsberg, 1988; Gray, 1989; Netter, 1991; Kuchera et al., 1993). Dette medfører at funksjonsforandringer i columna helt fra Th 11 til S4 samt skade eller annen patologi i urogenetalia og ligament- og bindevevsstrukturer i det lille bekken kan være medvirkende årsak til smertelokalisasjon i området mellom L5 og glutealfold.

Iliosacralleddene har vært særlig mye diskutert som årsak til smerter i lumbosacralområdet og bekken. Innervasjon til IS-leddene er beskrevet på ulike måter fra nivåene L2 til S4 i columna. Dette kan også medføre stor variasjon i smerteutbredelse (Fortin, Dwyer, West, & Pier, 1994; Slipman et al., 2000; Laslett et al., 2005). Fortin og medarbeidere (1994b) gjennomførte en kartlegging av smerteutbredelse fra IS-ledd som viste et 3 x 10 cm stort referanseområde for smerte lokalisert rett inferiørt for PSIS. Utenfor dette angitte området var smerteutbredelsen variabel. Dette aktuelle ”smertekartet” ble videre brukt i en studie av pasienter med bekkenplager, og forfatterne konkluderte med at dysfunksjon i iliosacralledd kunne screenes ved hjelp av smertetegning (Fortin, April, Ponthieux, & Pier, 1994). Dreyfuss og medarbeidere (1996) gjennomførte også en studie som viste varierende utbredelse av refererte smerter fra IS-ledd. Felles for pasientene som fikk påvist respons fra IS-leddene, var at de ikke hadde smerter proksimalt for L5-nivå. Andre arbeider har vist resultater som ikke samsvarer med dette. Slipman og medarbeidere (2000) viste at over 70 % av pasienter med diagnostiserte plager fra IS-ledd hadde smerteutbredelse til nedre lumbalcolumna.

Proksimal avgrensning av smertelokalisasjon til distalt for L5 eller øvre cristakant brukes likevel som kriterium for identifisering av bekkenplager i et større antall studier (Ostgaard et al., 1994a; Mens et al., 2002a; Vleeming et al., 2002; O'Sullivan et al., 2002; Hungerford, Gilleard, & Lee, 2004). Noen forfattere begrenser distal smerteutbredelse til glutealfold (Mens et al., 2002a; Vleeming et al., 2002). Andre inkluderer varierende grad av smerteutbredelse til underekstremitetene (Ostgaard et al., 1994a; Dreyfuss et al., 1996; Laslett et al., 2005). Bekkenplager er i følge Laslett (2005) hyppig beskrevet som smerter lokalisert mellom lumbosacral overgang og glutealfold. Plagene inkluderer smerter i området for sacrum, glutealmuskulatur og symfyse, samt refererte smerter distalt for glutealfold. Dette samsvarer med konklusjonene i en review fra 2005 som viste til at smerte fra IS-leddene kan radiere til ulike anatomiske områder inkludert sete, lyske og hele underekstremiteten. Her hevdes det at dersom smerten er lokalisert cranialt for L5-nivå, er det lite sannsynlig at IS-leddet er smerteårsak (Zelle et al., 2005).

## 2.3 Mål for smerte og funksjon

Opplevelse av helseplager er ofte relatert til graden av smerte og begrensning av funksjon i daglig aktivitet. Det var derfor av interesse å finne hensiktsmessige målemetoder for smerte og funksjon for den aktuelle gruppen kvinner med bekkenplager etter fødsel.

### 2.3.1 Smerte

Definisjonen på smerte er i følge The International Association for the Study of Pain (IASP) ”en ubehagelig sensorisk og følelsesmessig opplevelse forbundet med vevsødeleggelse eller mulig vevsødeleggelse, eller beskrevet som slik ødeleggelse” (Knardahl, 2002). Opplevelsen av smerte kan forsås som hjernens tolkning av tilgjengelig informasjon om kroppens tilstand (Knardahl, 2002; Brodal, 2005). Denne informasjonen kan være ufullstendig og konfliktfylt, og vil kunne være avhengig av faktorer som tidligere opplevelser, forventninger, holdninger, tro og emosjoner. Sammenhengen smerten opptrer i, vil også kunne være av betydning (Williamson & Hoggart, 2005). Smerteperspeksjon omfatter både en kognitiv, en sensorisk og en affektiv dimensjon (Gifford & Butler, 1997; Melzack, 2001). Smerte beskrives derfor også som en multidimensjonal erfaring (Von Korff, Jensen, & Karoly, 2000; Melzack, 2001). Den enkeltes erfaringer produseres av et vidt distribuert nevralt nettverk i hjernen (nevromatrix). Det er mønstrene for signaler ut fra dette systemet som produserer både den individuelle smerteerfaringen og den enkeltes adferdsmessige respons (Melzack, 2001).

Ulike dimensjoner av smerte er brukt som effektmål i klinisk forskning. Kvantifisering av en smerteopplevelse som ikke kan verifiseres av andre, er likevel forbundet med vansker. Det finnes ingen enkeltstående målemetoder som kan framstille alle de ulike dimensjonene ved smerte. Von Korff og medarbeidere (2000) hevdet at det kan være nyttig å betrakte alvorlighetsgrad av smerte som et globalt konstrukt. De viste til økende dokumentasjon for sammenheng mellom smerteintensitet og smertens påvirkning på aktiviteter, og at disse faktorene er tilstrekkelig korrelert til å forme en enkel underliggende dimensjon for alvorlighetsgrad av smerte. En valid og reliabel smertere registrering for klinisk forskning kan i følge disse forfatterne derfor oppnås ved å måle smerteintensitet sammen med smertens påvirkning på aktiviteter. Smerteintensitet er så langt målt på mange ulike skalaer, både visuelle, verbale og numeriske (Bolton & Wilkinson, 1998).

### *Visuell analog skala (VAS) som mål for smerteintensitet*

VAS er hyppig brukt som målemetode for smerteintensitet i klinisk forskning. Skalaen beskrives som en endimensjonal smertemålingsmetode (Kropmans, Dijkstra, Stegenga, Stewart, & de Bont, 1999). Studier har vist god test-retest reliabilitet ved akutte smerter (Revill, Robinson, Rosen, & Hogg, 1976; Williamson et al., 2005) og moderat til god test-retest reliabilitet ved lumbale smerter (Roach, Brown, Dunigan, Kusek, & Walas, 1997). VAS har gode evaluerende egenskaper og er vist å være et valid mål for smerteintensitet (Larsen et al., 1999; Von Korff et al., 2000). Det er derimot vist lav sensitivitet og spesifisitet for VAS ved måling av en pasientgruppe med kroniske lumbale smerter (Grotle, Brox, & Vollestad, 2004).

VAS måles på en 100 mm lang skala. Variabilitet ved bruk av VAS som smertemål er angitt å være rundt 20 % (Rosier, Iadarola, & Coghill, 2002). Ved å sammenligne med andre mål for smerteintensitet, er det funnet at VAS-skår på 30 mm tilsvarer moderate smerter, 54 mm eller mer tilsvarer sterke smerter (Collins, Moore, & McQuay, 1997). Klinisk signifikant smertereduksjon målt ved VAS er antydnet å være rundt 30-33 % (Williamson et al., 2005). Minste målbare forskjell ved VAS angis å være 22 mm for minimal og maksimal smerte og 28 mm for aktuell smerte. Forfatterne mener at disse resultatene er generaliserbare til heterogene grupper av smertepasienter (Kropmans et al., 1999; Finch, Brooks, Stratford, & Mayo, 2002).

### **2.3.2 Funksjon**

Funksjon er et begrep som forstås og benyttes på ulike måter relatert til helse og sykdom. Verdens helseorganisasjon har utarbeidet en internasjonal klassifisering av funksjon, funksjonshemming og helse (ICF) (WHO, 2001). Denne klassifiseringen reflekterer en distinksjon mellom tre helseperspektiver; kroppslig, personlig og sosialt.

I forbindelse med klinisk forskning er det utviklet en rekke spørreskjemaer med tanke på å måle ulike aspekter av funksjon. Funksjonsmål kan enten være generiske (uavhengige av sykdom) eller tilstands-spesifikke mål. Det har så langt vært lite fokus på hvilke aspekter ved funksjon som måles. Fokus på utførelse, det vil si hva pasienten faktisk gjør, er sannsynligvis den mest valide tilnærmingen i spørreskjema (Grotle et al., 2004).

Det finnes ingen validerte spørreskjema som er spesifikke for å måle funksjon ved bekkenplager. I tidlige studier er det benyttet ulike funksjonsmål. Mens og medarbeidere (2002a) benyttet QBPDS, en måleskala for uspesifikke korsryggssmerter. De hevdet at denne skalaen i en pilotstudie viste seg å være egnet for pasienter med bekkensmerter. Nilsson-Wikmar og medarbeidere (2005) benyttet et

annet ryggspesifikt spørreskjema, Disability Rating Index (DRI), i undersøkelse av gravide (Nilsson-Wikmar, Holm, Oijerstedt, & Harms-Ringdahl, 2005). Måling av funksjon er også løst ved å inkludere enkelte funksjonsvariabler i et mer omfattende spørreskjema (Rost, Jacqueline, Kaiser, Verhagen, & Koes, 2004). Wormslev og medarbeidere (1994) valgte å bruke en elleve punkts bokskala som mål for funksjon og disability (funksjonshemming) hos gravide. Laslett og medarbeidere (2003) benyttet Roland-Morris Disability Questionnaire (RMDQ) sammen med Dallas Pain Questionnaire for å måle disability ved inklusjon av pasienter med lumbal- og bekkensmerter. Oswestry Disability Index (ODI) er også benyttet for å måle funksjonsendringer ved behandling av iliosacralleddssmerter (O'Sullivan & Beales, 2006; Gutke, Ostgaard, & Oberg, 2006). Disse ulike studiene viser at det er gjort ulike skjønsmessige vurderinger av hvilke funksjonsmål som kunne være best egnet for pasienter med bekkenplager. I fortsettelsen presenteres ODI på bakgrunn av at dette spørreskjemaet har vist gode måleegenskaper og derfor ble benyttet i effektstudien til Stuge og medarbeidere (2004).

### *Oswestry disability index (ODI) som mål for funksjon*

ODI er utviklet for å kartlegge funksjonsnivå ved smerter i lumbalcolumna. Første utgave av dette spørreskjemaet ble initiert av John O'Brien i 1976 og publisert i 1980 (Fairbank, Couper, Davies, & O'Brien, 1980; Fairbank & Pynsent, 2000). Her ble begrepet disability definert som en pasients utførelse eller prestasjon sammenlignet med friske. De hevdet at mål for disability var en viktig komponent i evaluering av alle muskel- og skjelettplager (Baker, Pynsent, & Fairbank, 1990). Forfatterne viste til at kombinasjonen av lukkede spørsmål og selvadministrering ga et reliabelt format, og at skjemaet var en valid indikator på funksjonsnivå.

Spørreskjemaet ble tatt i bruk innen klinisk forskning i ulike miljøer, og nye versjoner ble utviklet både av forfatterne og andre. I en artikkel publisert i 2000 anbefalte Fairbank og medarbeidere versjon 2.0 av ODI, og hevdet at dette skjemaet var et valid mål for tilstands-spesifikk disability (Fairbank et al., 2000). Måleskala var fra 0 til 100 poeng. Under 20 poeng tilsvarte minimal disability, 21–40 poeng moderat, og over 40 poeng alvorlig disability.

En revidert versjon av ODI ble publisert av en gruppe kiropraktorer i 1989 (Hudson-Cook, Tomes-Nicholson, & Breen, 1989). De hevdet at den reviderte utgaven produserte høyere gjennomsnitt og større spredning i skår på disability, og mente dette øket sensitiviteten for spørreskjemaet når det ble benyttet av pasienter som var mindre plaget. I denne versjonen ble seksjonen om seksualliv utelatt og erstattet av spørsmål om endring av smertenivå. Fairbank kritiserte ODI Chiropractic Version for at denne utgaven blandet impairment med disability, og hevdet at skjemaet av den grunn ikke var



akseptabelt (Fairbank, 2000). Beurskens og medarbeidere (1996) påpekte også at evnen til å avdekke klinisk viktige endringer ved Hudson-Cook's versjon kunne være forskjellig fra originalversjonen siden den produserte høyere gjennomsnittsskår for disability, og at ODI Chiropractic Version derfor måtte etterprøves (Beurskens, de Vet, & Koke, 1996).

Både ODI originalversjon, versjon 2.0 og ODI Chiropractic Version ble etter hvert retestet i mange andre studier og viste gode måleegenskaper (Deyo et al., 1998; Grotle, Brox, & Vollestad, 2005). De modifiserte versjonene av ODI var like med hensyn til innhold. De inneholdt hovedsakelig aktivitetsbegrensinger, med noen få tema som reflekterte impairments og sosial funksjon. Grotle og medarbeidere (2005) konkluderte i motsetning til Fairbank og Beurskens med at ODI Chiropractic Version var godt validert og kunne anbefales brukt uten ytterligere valideringsstudier. De viste også til at dette skjemaet inneholdt en blanding av alle funksjonsområder vurdert ut fra kategoriene i ICF (kroppsfunksjoner, kroppsstrukturer, aktivitet, deltakelse) (Grotle et al., 2004; Grotle et al., 2005).

Evnen til å avdekke klinisk viktige endringer ved bruk av ODI har blitt sammenlignet med RMDQ og VAS. Sensitivitet for endring var lavere for ODI enn for de andre måleinstrumentene, samtidig var ODI mest spesifikk for endring (Beurskens et al., 1996).

Det er valgt ulike minimumsgrenser for klinisk signifikante endringer for ODI, varierende fra 4 til 15 poeng (Fairbank et al., 2000). Grotle og medarbeidere (2003) fant at 11 poeng var estimat for målefeil og minste målbare endring (Grotle, Brox, & Vollestad, 2003). Dette var nesten identisk med resultater i studien til Holm og medarbeidere (Holm, Friis, Storheim, & Brox, 2003). I studien til Beurskens og medarbeidere (1996) var den beste cut-off-grensen for klinisk signifikant endring 4-6 poeng.

## 2.4 Viktige egenskaper for gode målemetoder

Det har innen mange områder av forskning og klinisk arbeid vært en utfordring å finne gode og pålitelige målemetoder for de parametrene en ønsker å undersøke. Kliniske tester har hatt en sentral plass i praksis, men nøyaktighet og pålitelighet av testresultatene har vært omdiskutert. Resultatet av en test som har til hensikt å provosere smerter, vil kunne påvirkes av variasjoner både hos undersøker, pasient og i omgivelsene.

I klinisk forskning er fenomener som smerte og fysisk funksjonsnivå relevante som mål, men heller ikke disse ligger godt til rette for objektive målinger siden de ikke er direkte observerbare. Det er en

metodisk utfordring at det ikke finnes ”gullstandard” for måling av kliniske endringer (Grotle et al., 2004). Målemetoder vurderes ut fra noen sentrale begreper som er vesentlige for måleegenskapene.

### **2.4.1 Reliabilitet**

Reliabilitet er et uttrykk for pålitelighet og reproduserbarhet (Potter et al., 1985; Andresen, 2000). En reliabel test skal være nøyaktig og kunne gi samme resultat både når den utføres av samme person gjentatte ganger under samme forhold (intratester-reliabilitet), og når den utføres av ulike personer under like forhold (intertester-reliabilitet) (Hicks, 1999). Testen skal dermed gi samme resultat hver gang så lenge det ikke er oppstått en reell endring (Andresen, 2000). Reproduserbarhet vil kunne avhenge både av målemetode og av personen som utfører målingen. Med dette utgangspunktet er det viktig å kartlegge hvilke faktorer ved den enkelte målemetode som skaper variasjon (Benestad & Laake, 2004). Personen som utfører målinger, gjør i mange tilfelle subjektive vurderinger som en del av denne registreringen. I slike tilfelle må vedkommende betraktes som et instrument. Bestemte forventninger hos denne personen kan påvirke resultatet. ”Blinding” av undersøker er derfor brukt for å redusere usikkerheten ved målingen. Dette innebærer at undersøker ikke er kjent med hvilken kategori eller gruppe pasienten tilhører, og dermed er utelukket fra systematisk påvirkning av testresultatet (Altman, 1999). Standardisering av det øvrige utstyret som benyttes er også vesentlig for pålitelighet av målingen. På grunn av noe usikkerhet forbundet med både undersøker, pasient og målemetode vil det foreligge en grad av variasjon ved alle målinger. Det er derfor vesentlig at variasjon vurderes sammen med måleresultatet.

En test må være reliabel dersom resultatene av målingen skal være meningsfulle (Potter et al., 1985). Det er samtidig viktig å ikke avvise kliniske tester med lav reliabilitet i første omgang dersom dette skyldes manglende standardisering av teknikk eller evaluering av testresultatene (Strender et al., 1997). Ulike studier har antydnet at betydelig bedre standardisering er nødvendig for å oppnå høyere reliabilitet for kliniske tester (Strender et al., 1997; Levin et al., 2005). Andre har hevdet det er mindre sannsynlig at metodologisk oppgradering av testene vil føre til mer reliable resultater (Van der Wurff et al., 2000a).

### **2.4.2 Validitet**

Validitet er et uttrykk for i hvilken grad en målemetode faktisk måler det den har til intensjon å måle (Hicks, 1999). Metoder som velges i forskningsarbeid må relateres til ulike typer validitet.

Begrepsvaliditet (konstrukt validitet) er knyttet til gyldigheten av selve begrepet som studeres. Det

måler i hvor stor grad målevariabelen gjenspeiler eller reflekterer den variabelen man virkelig ønsker å studere. For å bedre begrepsvaliditet, må begrepet avklares og gis en operasjonell definisjon som viser hvordan variabelen skal måles (Domholdt, 1993). Innholdsvaliditet viser i hvilken grad et mål er representativt for det aktuelle begrepet eller området det er ønskelig å måle (Domholdt, 1993; Pallant, 2004). Både begrepsvaliditet og innholdsvaliditet benyttes i stor grad ved evaluering av spørreskjemaer.

Kliniske tester vurderes både i forhold til hvor gyldige de er på det tidspunktet de benyttes, og med hvilken sikkerhet de kan forutsi et utfall fram i tid. Samtidig validitet er et uttrykk for i hvilken grad et mål er relatert til en annen målestANDARD. Ved vurdering av samtidig validitet er det vesentlig at målingene er nær i tid for i størst mulig grad å måle samme fenomen (Domholdt, 1993). For å vurdere validitet av en klinisk test er det nødvendig å sammenligne resultatene med de som er oppnådd ved en referansetest. En svært godt egnet referansetest benevnes ofte som gullstandard. Men i klinisk forskning finnes få tester som oppnår gullstandard, og de som finnes kan i en del tilfeller både være tidkrevende og kostbare, av og til også med en testprosedyre som innebærer risiko for pasienten (Van der Wurff et al., 2000b). Det er også ofte av interesse både for klinikere og forskere å vite i hvilken grad en test kan forutsi et framtidig resultat. Dette benevnes som testens prediktive validitet (Hicks, 1999).

Validiteten av en testprosedyre er definert som evnen til å bestemme korrekt om en patologisk tilstand er tilstede eller ikke (Benestad et al., 2004). Validitet er på dette grunnlaget direkte knyttet til sensitivitet og spesifisitet.

### ***Sensitivitet og spesifisitet***

Sensitivitet representerer den andelen i en populasjon som har patologisk tilstand og samtidig positivt resultat av den diagnostiske testen. En test som kan identifisere korrekt alle undersøkte med den aktuelle patologiske tilstanden, har sensitivitet på 1.0. Spesifisitet representerer andelen friske i populasjonen som har negativt resultat av den diagnostiske testen. En test som kan identifisere korrekt alle undersøkte som ikke har den aktuelle patologiske tilstanden, har spesifisitet på 1.0 (Potter et al., 1985; Altman, 1999; Cibulka & Koldehoff, 1999; Van der Wurff et al., 2000b; Benestad et al., 2004). Sensitivitet og spesifisitet er direkte relatert til hverandre. Tester som har både høy sensitivitet og spesifisitet er de mest pålitelige. Noen hevder at tester med svært høy sensitivitet og svært lav spesifisitet, eller omvendt, er av liten verdi (Benestad et al., 2004). Andre mener en test med svært høy sensitivitet kan være klinisk nyttig selv om den har lav spesifisitet siden en negativ test da har svært liten sannsynlighet for å være falsk negativ (Laslett, 2005).

I dette kapitlet er det presentert en rekke kliniske tester som benyttes ved utredning av bekkenplager. Ut fra tilgjengelig dokumentasjon er testenes reliabilitet og validitet forsøkt framstilt. Det er også presentert verktøy som kan benyttes for subjektivt å angi smerteintensitet og funksjonsnivå relatert til plagene. I studien som presenteres i denne oppgaven er egenrapportert smerteintensitet målt ved hjelp av VAS og funksjon målt ved ODI. Skår på kliniske tester holdes opp mot disse egenrapporterte målene for å få et innblikk i samtidig validitet av de aktuelle testene.

## 2.5 Forskningsspørsmål

Teori som er presentert i dette kapitlet har vist noe av både mangfold, muligheter og mangler når det gjelder målemetoder og diagnostiske prosedyrer for pasienter med bekkenplager. Klinisk praksis er som nevnt varierende, og forskere har gjort ulike metodevalg og benyttet ulike innfallsvinkler for å kartlegge fagfeltet. Beskrivelse av det kliniske bildet av denne pasientgruppen har vist seg å være mangelfull.

På bakgrunn av denne gjennomgangen var det av interesse å gjøre en beskrivelse av den aktuelle gruppen kvinner med bekkenplager ut fra forekomst av positive skår på kliniske tester. For å få et innblikk i samtidig validitet for testene, var det også ønskelig å undersøke sammenheng mellom skår på testene og kvinnenes egen rapportering av smerteintensitet og funksjon.

Problemstillingene som ønskes belyst er:

1. Hvordan fordeles skår ved aktuelle kliniske tester og egenrapportert smerteintensitet og funksjon hos kvinner med bekkenplager etter fødsel og etter påfølgende behandling?
2. Hvordan er sammenhengen mellom skår på aktuelle kliniske tester og egenrapportert smerteintensitet og funksjon ved bekkenplager etter fødsel?
3. Endres skår på kliniske tester hos en gruppe kvinner som gjennomsnittlig har bedret funksjon og redusert smerteintensitet etter behandling av bekkenplager etter fødsel?
4. Endres sammenheng mellom kliniske tester og egenrapportert smerte og funksjon i løpet av denne behandlingsperioden?
5. Hvilke av de aktuelle kliniske testene har selvstendig forklaringsverdi for egenrapportert smerteintensitet og funksjon?

### 3. Materiale og metode

Datamateriale for denne studien ble innhentet i forbindelse med en effektstudie gjennomført av Stuge og medarbeidere (2004). Mål for effektstudien var å evaluere om spesifikke stabiliserende øvelser ved behandling av bekkenplager (PGP) etter svangerskap reduserte kvinnes smerter, bedret funksjon og helse relatert livskvalitet etter behandlingsperioden og ett år etter fødsel mer enn fysioterapi uten spesifikke stabiliserende øvelser. Datainnsamlingen ble utført i tidsrommet 1999-2001. Et større antall kliniske tester ble benyttet for inklusjon, eksklusjon og evaluering. Det ble i ettertid vurdert som interessant å studere dette materialet nærmere. En ny studie av de aktuelle kliniske testene ble derfor designet.

#### 3.1 Rekruttering av deltakere

Populasjonen for datainnsamling til effektstudien var kvinner i Oslo og Akershus med bekkenplager etter fødsel. Skriftlig informasjon ble sendt ut til relevant helsepersonell i primærhelsetjenesten i de to aktuelle fylkene med forespørsel om å rekruttere pasienter. En stor andel kvinner ble henvist, noen tok også kontakt på eget initiativ. Til sammen 95 kvinner ble funnet aktuelle for undersøkelse etter et gjennomgått telefonintervju. De som ble inkludert i effektstudien, ble i neste omgang også deltakere i studien som presenteres i denne masteroppgaven.

#### 3.2 Design

Denne studien ble designet som en prospektiv kohortestudie (Altman, 1999). Samtlige inkluderte kvinner ble analysert som én gruppe. I forbindelse med effektstudien til Stuge og medarbeidere (2004) ble disse kvinnene undersøkt ved hjelp av kliniske tester etter en fastlagt prosedyre ved inklusjon (pretest), etter 20 uker ved avsluttet behandling (posttest), og ett år etter inklusjon. I denne kohortestudien er data fra pretest og posttest benyttet.

#### 3.3 Kriterier for inklusjon og eksklusjon

Følgende kriterier ble valgt for inklusjon:

- Nordiske kvinner som forstår og snakker norsk
- Smerter fra bekkenregionen som har debutert i siste svangerskap, i forbindelse med fødsel, eller innen tre uker etter fødsel
- Siste fødsel 6-16 uker før inklusjon
- Smertelokalisasjon svarende til bekkenregion distalt for L5-nivå, angitt ved smertetegning som valideres ved å la kvinnen peke på angitte smerteområder på egen kropp
- Kliniske tester for inklusjon (nærmere beskrivelse i kap. 3.6.2):
  - Posterior pelvic pain provocation (P4) test
  - Provokasjonstest av symfysis pubis eller symfysemerter ved modifisert Trendelenburgs test
  - Active straight leg raise (ASLR) test, 3 poeng på en side eller minimum 4 poeng bilateralt
  - Palpasjonstest av long dorsal ligament (LDL)

P4 eller ALSR må være positiv på en eller begge sider, samt minst en av de andre testene.

Følgende kriterier ble valgt for eksklusjon:

- Ryggplager med indikasjon på radikulopati. Tester for eksklusjon er positiv SLR, Slumps test, Crams test eller test av nervus femoralis
- Revmatologisk sykdom, malign eller annen alvorlig sykdom
- Kontinuerlige bekkenplager fra tidligere svangerskap eller debut av bekkenplager *før* starten av siste graviditet

Ved tvilstilfeller for tolkning av de aktuelle kriteriene ble ekstern undersøker tilkalt for vurdering.

### **3.3.1 Kliniske tester for eksklusjon**

Både i klinikk og forskning har det vært en utfordring å differensiere mellom patologi i lumbalcolumna og bekken. Nervestrekktester har vært benyttet for å vurdere om nevralt vev kunne være involvert ved muskel-skjelettplager (Kerr, Vujnovich, & Bradnam, 2002). Kliniske tegn som indikerer kompresjon av nerverøtter har vært hyppig brukte kriterier for differensiering eller eksklusjon i studier relatert til plager i lumbal og bekken (Kristiansson et al., 1996a; Albert et al., 2000; Mens et al., 2002a; Laslett et al., 2005). Nervestrekktester benyttes for å reproducere eller

forsterke pasientens smerter. Dura, nerverøtter og perifere nerver kan under patologiske forhold gi smerter ved strekk. Utstrålende smerter fra nerverøtter og dorsale rotganglier er beskrevet som intense, skjærende, lette å lokalisere og med utbredelse som en streng eller smal stripe til ekstremiteten (Kirkesola et al., 1998; Bogduk, 2005). Det finnes en rekke ulike nervestrekktester. De fire følgende ble benyttet som eksklusjonskriterier i effektstudien: Straight Leg Raising (SLR) test, Crams test, Slump test og strekktest av nervus femoralis.

SRL er en nervestrekktest utviklet for å identifisere nerverotsmerter. Positiv test har vært betraktet som et av de mest vesentlige kliniske tegn for å diagnostisere prolaps eller protrusjoner av discus intervertebralis (CRAM, 1953). Det har vært motstridende resultater angående sensitivitet og spesifisitet for SLR. Det er rapportert at Kosteljanetz i 1988 fant lav sensitivitet (0.33) og høy spesifisitet (0.88) for testen (Stankovic, Johnell, Maly, & Willner, 1999). En review-studie rapporterte sensitivitet på 0.91 og spesifisitet på 0.26. Her ble det hevdet at lav spesifisitet begrenset testens diagnostiske presisjon (Deville, van der Windt, Dzaferagic, Bezemer, & Bouter, 2000).

Crams test ble utviklet som en videreføring av SLR. En eksperimentell studie viste at tensjonen i nervus ischiadicus økte ved kompresjon av nervus tibialis i knehasen. Testen skulle bidra til å eliminere falske positive testsvar ved SLR. Falske positive svar kunne ha sin årsak i refererte smerter som ikke var assosiert med trykk på nerverøtter (CRAM, 1953).

Slump test ble først presentert av Maitland, og er en reliabel test for nevralt strukturer i lumbal og underekstremiteten (Kerr et al., 2002). I motsetning til SLR og Crams test utføres denne testen i vektbærende stilling der nervestrukturere provoseres ved ekstensjon av kne og fleksjon av cervicalcolumna (Kirkesola et al., 1998).

Strekktest av nervus femoralis forventes å gi smerteutstråling til ventralside av lår ved kompresjon av nerverot fra segmentene L2, L3 eller L4 (Estridge, Rouhe, & Johnson, 1982). Validitet for denne nervestrekktesten er ukjent (Kreitz, Cote, & Yong-Hing, 1996).

### 3.4 Bakgrunnsfaktorer

Opplysninger om kvinnenens bakgrunn, familieforhold og helse ble innhentet ved spørreskjema utarbeidet for effektstudien. Alder, høyde, vekt, sivilstand, utdanningsnivå, antall barn og antall uker siden siste fødsel ble registrert. Aktuell sykehistorie ble kartlagt ved varighet av plagene, variasjon i

smertes, sykemelding i og etter svangerskap, tidligere bekkenløsning eller skader i bekkenet, tidligere ryggplager, samt grad av fysisk aktivitet før det aktuelle svangerskapet.

## 3.5 Egenrapporterte mål

Smerteintensitet og funksjon ble valgt som relevante effektmål i studien til Stuge og medarbeidere (2004). I tidligere studier er det gjort ulike metodiske valg for å måle disse variablene. Ingen aktuelle målemetoder er spesielt tilpasset for bruk ved bekkenplager. Visuell analog skala (VAS) er hyppig benyttet i sammenlignbare studier og ble valgt som smertemål. Det ble også gjort en skjønnsmessig vurdering av egnede validerte funksjonsmål for den aktuelle pasientgruppen, og Oswestry Disability Index (ODI) Chiropractic Version ble valgt som mål for funksjon (Hudson-Cook et al., 1989). Alle pasienter fikk utlevert et omfattende spørreskjema for utfylling i forkant av klinisk undersøkelse både ved pretest og posttest. Inkludert i dette spørreskjemaet var VAS for rapportering av smerte og ODI i norsk oversettelse for rapportering av funksjon. Kvinnene ble bedt om å svare på alle spørsmål. Dersom aktuelle svaralternativer ikke passet helt, ble de i forkant oppfordret til å velge det mest nærliggende alternativet. De ble hjulpet til en komfortabel hvilestilling og var deretter alene under utfylling av spørreskjemaet. Etter den kliniske undersøkelsen og vurdering av inklusjonskriteriene, ble det undersøkt om alle spørsmålsfelt var besvart. Dersom svar manglet, ble kvinnene bedt om å fylle ut disse.

### 3.5.1 Smerte

VAS som mål for egenrapportert smerte ble presentert som en 100 mm lang horisontal linje i spørreskjemaet. Lengde på skalaen ble kontrollmålt i kopierte skjemaer. Venstre endepunkt på linjen var markert med teksten ”ingen smerte”, og høyre endepunkt med ”uutholdelig smerte”. Data ble registrert i antall mm fra 0 til 100, målt fra venstre endepunkt til avmerket kryss på linjen. Kvinnene ble bedt om å vurdere sitt ”nåværende smertenivå” på det verste om morgenen etter at de hadde stått opp, og på det verste om kvelden før de gikk til sengs. I de følgende analysene ble smerte om kvelden valgt ut fra klinisk erfaring og dokumentasjon for at denne pasientgruppen ofte rapporterer økende plager utover dagen (Mens, Vleeming, Stoeckart, Stam, & Snijders, 1996).



### 3.5.2 Funksjon

Spørreskjemaet ODI Chiropractic Version består av 10 deler. Hver del består av 6 utsagn som beskriver ulik grad av funksjonsdyktighet. Høyest mulig skår for en del er 5. Dersom det er krysset av for første utsagn, er skår lik 0, dersom det er krysset av for siste utsagn er skår lik 5. Dersom det er krysset av for flere utsagn i samme seksjon, velges det utsagnet som har høyest skår. Endelig skår på ODI Chiropractic Version regnes ut ved å angi samlet skår i prosent av høyeste mulige sumskår (50) (Fairbank et al., 1980; Hudson-Cook et al., 1989).

## 3.6 Kliniske tester

I tillegg til kvinnes egen rapportering av smerte og funksjon, ble kliniske tester benyttet for kartlegging av plager. Før oppstart av effektstudien ble aktuell vitenskapelig dokumentasjon for relevante kliniske tester gjennomgått. I tillegg ble et internasjonalt ekspertpanel konsultert med tanke på utvalg og utførelse av testene. Vurdering av denne informasjonen i kombinasjon med klinisk erfaring og skjønn resulterte i en kombinasjon av tester, hovedsakelig smerteprovokasjonstester. Standardisering og utførelse av de aktuelle testene ble beskrevet, og de ble videre delt i kategoriene inklusjonstester, eksklusjonstester og evaluerende tester. Det ble utført 14 kliniske tester for inklusjon og evaluering, samt fire nervestrekktester for eksklusjon. Bilaterale tester ble skåret som 0 = negativ, 1 = unilateral positiv, 2 = bilateral positiv test. Sideuavhengige tester ble skåret som 0 = negativ og 1 = positiv test. Unntaket var ASLR som ble skåret på en skala fra 0 til 10 poeng der 0 tilsvarte negativ test.

### 3.6.1 Gjennomføring av testprosedyre

Testskjemaet ble redigert med tanke på hensiktsmessig rekkefølge av testene. Kvinnene skulle først testes i stående, videre i ryggliggende, sideliggende, framliggende og sittende. Minimalt antall stillingsendringer ble valgt for å redusere den totale belastningen under testprosedyren. Antall tester ble også vurdert og begrenset ut fra klinisk erfaring angående smerteprovokasjon ved undersøkelse av denne pasientgruppen. Det var vesentlig at testprosedyren ikke skulle oppleves så belastende at kvinnene ville vegre seg for retesting. Ved innkalling til pretest fikk hver enkelt kvinne velge om hun ville testes på formiddag eller ettermiddag, og posttest ble i neste omgang gjennomført på samme tidspunkt på dagen. Dette resulterte i en ganske jevn fordeling av tester tidlig og seint på dagen.

Flere treningsesjoner ble gjennomgått sammen med prosjektleder for å avklare den praktiske utførelsen av de ulike testene. Kompresjonskraft ble testet ut med gjentatte repetisjoner med MicroFet håndholdt dynamometer for P4 og kompresjonstester (se testbeskrivelse i kap. 3.6.2). Trykkgrad for palpasjonstester ble også innøvd ved gjentatte repetisjoner og tilbakemeldinger for best mulig standardisering av testene. Forfatter av denne oppgaven utførte samtlige kliniske tester og registrerte skår for disse på undersøkelsesskjema både ved pretest og posttest. Det ble underveis gjort noen mangelfulle skriftlige registreringer som er synliggjort ved redusert antall (missing) for de aktuelle testene i presentasjon av resultatene.

Samme lokale, undersøkelsesbenk og tilleggsutstyr (pøller, puter) ble brukt ved alle målingene. Benk og øvrig utstyr hadde fast plassering i lokalet. Ved P4-testen var benken senket til minimumsnivå for å gi undersøkeren tilnærmet lik utgangsstilling ved alle målinger. Muntlig instruksjon for den enkelte test var standardisert, og det ble lagt vekt på å unngå ledende formuleringer av spørsmål. Dersom kvinnene hadde oppklarende spørsmål underveis i testprosedyren, ble disse besvart. I det følgende kapittelet er de kliniske testene presentert i den rekkefølgen de ble utført under testprosedyren.

### **3.6.2 Beskrivelse av kliniske tester**

#### *Trendelenburgs test*

Pasientens utgangsstilling er stående på begge bein. Undersøker sitter bak pasienten.

Utførelse: Pasienten står på ett bein og flekterer til 90 grader i hofte og kne på motsatt side. Pasienten skal ikke lene seg over standbeinet. Testen utføres bilateralt (Kirkesola et al., 1998; Albert et al., 2000).

Testen er positiv dersom pasienten ikke kan holde bekkenet horisontalt.

#### *Registrering av symfysesmerter ved modifisert Trendelenburgs test*

Utgangsstilling og utførelse er som beskrevet over.

Test for symfysesmerter i stående er positiv dersom smerter provoseres i symfyser ved stående på ett bein (skiller ikke om dette er på høyre, venstre eller begge sider) (Albert et al., 2000).

### ***Provokasjonstest av symfysis pubis***

Pasientens utgangsstilling er ryggliggende med en pølle under knærne. Undersøker står ved siden av benken.

Utførelse: Undersøker palperer craniale del av symfysis pubis, gir deretter et lett til fast trykk i vertikal retning og holder inntil 5 sekunder. Testen gjentas litt mer caudalt i symfysis pubis dersom første trykk resulterer i negativ test.

Testen er positiv hvis provosert smerte vedvarer mer enn 5 sekunder etter avsluttet trykk (Albert et al., 2000).

### ***Posterior pelvic pain provocation (P4) test***

Pasientens utgangsstilling er ryggliggende med 90 grader fleksjon (uten adduksjon) i hofte og tilnærmet maksimalt flektert kne på testsiden. Undersøker står ved benken ved siden av pasientens flekterte hofte.

Utførelse: Undersøker plasserer en hånd på flektert kne, motsatt hånd stabiliserer pasientens spina iliaca anterior superior. Det gis et skånsomt aksialt/ vertikalt trykk gjennom femurs longitudinelle akse. Dersom dette ikke gir smerteprovokasjon, økes kraften gradvis og trykket holdes inntil 5 sekunder. Testen utføres bilateralt.

Testen er positiv hvis den utløser eller øker smerte i IS-regionen på testsiden (Ostgaard et al., 1994a).

### ***Active straight leg raise (ASLR) test***

Pasientens utgangsstilling er ryggliggende med strake bein. Undersøker står ved enden av benken.

Utførelse: Pasienten bes om å løfte ett bein 5-10 cm over benken med ekstendert kne. Følgende spørsmål stilles: "Kan du vurdere hvor vanskelig eller tungt det er å løfte beinet?" Pasienten bes om å skåre muntlig på testen ut fra følgende poengskala: 0 = ingen problemer, 1 = svært lite vanskelig (evt. tungt), 2 = litt vanskelig, 3 = vanskelig, 4 = svært vanskelig, 5 = ikke mulig. Testen gjentas på motsatt side (Mens et al., 2001; Mens et al., 2002a).

Testen er positiv som inklusjonstest ved minimum 3 poeng på en side eller minimum 4 poeng i sum bilateralt.

### *Kompresjon ryggliggende*

Pasientens utgangsstilling er ryggliggende med pølle under knærne. Undersøker står ved siden av benken på pasientens høyre side.

Utførelse: Undersøkers hender plasseres lateralt for SIAS bilateralt. Ilium komprimeres i medial retning bilateralt og holdes noen sekunder, deretter gis et forsterket trykk (Kristiansson et al., 1996a; Kirkesola et al., 1998).

Testen er positiv hvis trykket provoserer smerter eller eventuelt øker smerter i IS-område eller symfyseregion.

### *Distraksjon*

Pasientens utgangsstilling er ryggliggende med pølle under knærne. Undersøker står ved siden av benken på pasientens høyre side.

Utførelse: Undersøkers hender plasseres i kryssgrep på medialsiden av SIAS bilateralt. Ilii skyves i posterolateral retning bilateralt og trykket holdes noen sekunder (Kirkesola et al., 1998; Laslett et al., 2005).

Testen er positiv hvis skyvet provoserer smerter eller eventuelt øker smerter i IS-område eller symfyseregion.

### *Patrick`s test*

Pasientens utgangsstilling er ryggliggende med testbeinets hæl plassert på motsatt sides kne. Undersøker står ved siden av benken på pasientens testside.

Utførelse: Undersøker plasserer en hånd på medialsiden av testbeinets kne, den andre hånden på pasientens SIAS på motsatt side. Pasienten bes om å senke kneet ned mot benken. Dersom kneet ikke kan senkes helt ned, gis et skånsomt vertikalt trykk mot kneets medialside. Testen utføres bilateralt.

Testen er positiv hvis a) bevegelsesutslaget er betydelig redusert, eller b) smerte provoseres i IS-område på testsiden ved ytterstilling av normalt bevegelsesutslag (Streuder et al., 1997; Broadhurst et al., 1998; Kirkesola et al., 1998; Albert et al., 2000).

### ***Menell`s test***

Pasientens utgangsstilling er ryggliggende med strakt bein, 30 grader abduksjon og 10 grader fleksjon i hoften. Undersøker står ved enden av benken.

Utførelse: Undersøker tar grep med begge hender rundt pasientens ankel. Det utføres først kompresjon i cranial retning med undersøkers bekken mot pasientens fot, deretter traksjon i distal retning. Testen utføres bilateralt (Albert et al., 2000).

Testen er positiv hvis smerte provoseres eller eventuelt øker i IS-område eller symfyseregion under kompresjon og/ eller traksjon.

### ***Palpasjon av m. iliopsoas***

Pasientens utgangsstilling er ryggliggende med flekterte hofter og knær. Undersøker står ved siden av benken.

Utførelse: M. psoas palperes med ”flatt” fingergrep. Pasienten bes flekere i hoften mot motsand fra undersøker for å bekrefte at riktig struktur palperes. Følgende spørsmål stilles til pasienten: ”Er dette trykket tydelig ømt eller vondt? Hvis du er usikker, kan du sammenligne med motsatt side”. Testen utføres bilateralt.

Testen er positiv hvis pasienten angir tydelig ømhet eller smerte ved palpasjonstrykket (Wormslev et al., 1994).

### ***Mål av diastase i m. rectus abdominis***

Pasientens utgangsstilling er ryggliggende med flekterte hofter og knær og armene langs siden. Undersøker står ved benken på pasientens høyre side.

Utførelse: Palpasjon utføres like caudalt for umbilicus mens pasienten ruller opp i fleksjon av columna til scapulae er fri fra underlaget. Testen gjentas med palpasjon 4-5 cm cranialt og caudalt for umbilicus for kontrollmåling. Diastase registreres i antall fingerbredder.

Testen er positiv hvis diastase like caudalt for umbilicus måles til to fingerbredder eller mer (Boissonnault & Blaschak, 1988; Wormslev et al., 1994; Gilleard & Brown, 1996).

### ***Kompresjon sideliggende***

Pasientens utgangsstilling er liggende på venstre side med lett flekterte hofter og knær og pute mellom beina. Undersøker står ved siden av benken bak pasienten.

Utførelse: Undersøkers hender plasseres over hverandre lateralt for spina iliaca anterior superior (SIAS) på pasientens høyre side. Ilium komprimeres i medial/vertikal retning og holdes noen sekunder, deretter gis et forsterket trykk (Kirkesola et al., 1998; Laslett et al., 2005).

Testen er positiv hvis kompresjonen provoserer smerter eller eventuelt øker smerter i IS-område eller symfyseregion.

### ***Palpasjon av long dorsal ligament (LDL)***

Pasientens utgangsstilling er framliggende med strake bein og pølle under anklene. Undersøker står ved benken på testsiden.

Utførelse: Palpasjon rett caudalt for spina iliaca posterior superior (SIPS). Følgende spørsmål stilles til pasienten: ”Gjør dette trykket vondt? Hvis du er usikker, kan du vurdere sammenlignet med motsatt side”. Testen avsluttes før innøvd palpasjonstrykk er oppnådd dersom smerte rapporteres, trykket holdes ikke over tid. Testen utføres bilateralt (Vleeming et al., 2002).

Testen er positiv hvis pasienten angir smerte ved palpasjon.

### ***Palpasjon av lig. sacrotuberale***

Pasientens utgangsstilling er framliggende med strake bein og pølle under anklene. Undersøker står ved benken på testsiden.

Utførelse: Palpasjon lateralt for coccygeus i cranial og lateral retning. Følgende spørsmål stilles til pasienten: ”Gjør dette trykket vondt? Hvis du er usikker, kan du vurdere sammenlignet med motsatt side”. Testen avsluttes før innøvd palpasjonstrykk er oppnådd dersom smerte rapporteres, trykket holdes ikke over tid. Testen utføres bilateralt.

Testen er positiv hvis pasienten angir smerte ved palpasjon (Wormslev et al., 1994).

## 3.7 Statistiske analyser

VAS som mål for smerte og ODI som mål for funksjon ble begge skåret på skala 0-100. ASLR ble skåret på skala 0-10. Disse tre variablene ble analysert som kontinuerlige data. Alle de øvrige kliniske testene representerte kategoriske data. Av disse ble sideuavhengige tester skåret som 0 = negativ og 1 = positiv test. Bilaterale tester ble skåret som 0 = bilateral negativ, 1 = ensidig positiv og 2 = bilateral positiv test. I tillegg ble demografiske data analysert, disse representerte både kontinuerlige og kategoriske variabler.

Statistiske analyser ble utført ved hjelp av SPSS, versjon 12.0.1. Usikkerhetsmål (spredning) for kontinuerlige data ble angitt med 95 % konfidensintervall (CI) og standarddeviasjon (SD). Spredning for bakgrunnsvariabler ble angitt med standarddeviasjon, minimums- og maksimumsverdier. Signifikansnivå for sammenhenger og endringer ble satt til 5 % med tosidig testing.

For en deskriptiv framstilling av skår på de kontinuerlige variablene VAS, ODI og ASLR ble det benyttet histogram samt gjennomsnittsverdier, standarddeviasjon og konfidensintervall. For de 13 kliniske testene med kategoriske data ble søylediagram, antall og prosent benyttet. ASLR var i utgangspunktet definert som en kontinuerlig variabel, men testen ble også dikotomisert i den deskriptive framstillingen. Dette ble gjort ut fra anbefalt cut-off mellom 0 som negativ og 1-10 som positiv skår i original beskrivelse av testen (Mens et al., 2001).

Statistisk signifikansnivå for endring av ASLR ble undersøkt ved parett-test. Endring i skår på alle testene med kategoriske data ble framstilt ved prosentangivelser. Alle disse registreringene var parede data for kohorten registrert ved pretest og posttest. Sideuavhengige tester ble analysert med 2 x 2-tabeller, og statistisk signifikansnivå for endring ble angitt ved Mc. Nemar's test. Bilaterale tester ble analysert med r x c-tabeller, og for disse testene ble Mc. Nemar Bowker's test benyttet for framstilling av statistisk signifikansnivå.

De 14 kliniske testene ble analysert med tanke på mulige assosiasjoner/ sammenhenger mellom testene. For ASLR ble Pearson korrelasjonsanalyse benyttet. For de 13 testene som var kategoriske variabler, ble det benyttet tabellanalyser med kjikvadrat-test. For 2 x 2-tabeller ble Yates' Correction oppgitt i stedet for kjikvadratet (Altman, 1999). I tilfeller der kravet om forventet antall i cellene ikke var oppfylt, ble Fischer's eksakte test benyttet for å angi p-verdi. Kjikkvadrattest ble også brukt for å analysere r x c-tabeller. Et flertall av disse tabellanalysene oppfylte ikke forutsetningen om forventet antall større enn 5 for 80 % av cellene i tabellen. Også her ble Fischer's eksakte test benyttet.

Mulige sammenhenger mellom de kliniske testene og VAS og/ eller ODI ble undersøkt ved korrelasjonsanalyser. Ved pretest viste VAS, ODI og ASLR tilnærmet normalfordelte data, ved posttest var det noe avvik fra normalfordeling for alle tre variablene. Korrelasjonsanalysene ble gjort med både parametriske (Pearson) og ikke-parametriske (Spearman`s) tester med tilnærmet likt resultat. Videre analyser ble ut fra dette basert på Pearson korrelasjonskoeffisient for dermed å unngå å gjøre logg-transformering av datamaterialet. Boksplott ble i tillegg benyttet for å framstille eventuelle forskjeller i skår på VAS og ODI mellom gruppene med positiv og negativ testskår.

Trinnvis multippel regresjon ble gjennomført for å analysere i hvilken grad skår på kliniske tester kunne forklare variansen i VAS og ODI, og for å vurdere hvilke tester som eventuelt kunne ha en selvstendig forklaringsverdi for smerte og funksjon. Det ble benyttet 10 uavhengige variabler i denne regresjonsmodellen. Antallet tester ble redusert fra 14 til 10 på følgende måte: Testene kompresjon ryggliggende og kompresjon sideliggende ble slått sammen til ”kompresjon”. Kompresjon ble da registrert som positiv dersom minst en av utgangsstillingene for testen hadde gitt positiv skår.

Trendelenburgs test, registrering av diastasemål og symfysemerter i stående (modifisert Trendelenburgs test) ble også fjernet fra regresjonsanalysen på grunnlag av innledende analyseresultater (se kap. 4.4.3). 10 uavhengige variabler var noe i overkant av anbefalinger i litteraturen relatert til antall inkluderte pasienter (Altman, 1999), men ble vurdert som akseptabelt. Dataene ble sjekket for multikollinearitet og outliers ved vurdering av scatterplot og normalplot for standardiserte residualer, samt ved å sjekke at verdier for Mahalanobis distances ikke oversteg kritiske verdier. Korrelasjon mellom avhengige variabler og hver enkelt av de uavhengige variablene ble kontrollert med tanke på at korrelasjonen helst skulle være over 0.30. Det ble også kontrollert at korrelasjonen mellom de uavhengige variablene ikke var over 0.70 (Pallant, 2004). De samme analysene for trinnvis multippel regresjon ble gjentatt med ASLR som dikotom variabel for å sammenligne resultatene med tanke på selvstendig forklaringsverdi for egenrapportert smerte og funksjon.

På bakgrunn av at datamaterialet inneholdt skåringer på et stort antall kliniske tester, ble det til slutt i analysene valgt å presentere en samlet framstilling av testskår både ved pretest og posttest. Antall positive skår ble summert for hver enkelt kvinne og framstilt ved søylediagram, og sumskår ble deretter plottet sammen med skår på VAS og ODI ved både pretest og posttest. Dette ble gjort for å synliggjøre variasjon og eventuelle trender. Korrelasjonsanalyse ble benyttet for å undersøke sammenheng mellom antall positive tester og egenrapportert smerte eller funksjon.



### 3.8 Etiske hensyn

Før oppstart av effektstudien ble det søkt og innvilget godkjenning fra Regional Etisk Komite (REK). Godkjenning fra Norsk Samfunnsvitenskapelig Datatjeneste (SND) var ikke påkrevd for dette forskningsarbeidet. Samtlige pasienter som ble rekruttert til effektstudien mottok muntlig og skriftlig informasjon og undertegnet erklæring om frivillig samtykke før eventuell inklusjon. De ble også informert om mulighet for å trekke seg på ethvert tidspunkt underveis i prosjektet. Alle data ble anonymisert etter inklusjonsnummer, og videre analyser ble gjort ut fra denne nummereringen.

Kliniske tester ble brukt for inklusjon/ eksklusjon ved pretest og for evaluering ved måletidspunktene etter behandling. Valg av aktuelle tester ble gjort på bakgrunn av rådende praksis i fagfeltet og i samsvar med relevant vitenskapelig dokumentasjon på det aktuelle tidspunktet. Varighet for testprosedyren var omtrent lik en vanlig klinisk undersøkelse, og kvinnene ble på bakgrunn av dette ikke utsatt for ekstra fysiske belastninger verken ved inklusjon eller posttest.

## 4. Resultater

### 4.1 Bakgrunnsfaktorer for inkluderte kvinner

95 kvinner med bekkenplager etter fødsel ble funnet aktuelle for undersøkelse for eventuell inklusjon i effektstudien. Undersøkelsen førte til at 14 kvinner ble ekskludert og 81 ble inkludert. Dette utgjorde materialet både til effektstudien og i neste omgang til denne kohortestudien. De inkluderte viste seg å være en gruppe med relativt høy utdanning, gjennomsnitt i antall studieår lå over høyskolenivå (tab.1). De hadde fra ett til fire barn, 48 % (39 kvinner) var førstegangsfødende. Nesten alle levde sammen med ektefelle eller samboer. 76 % hadde vært helt eller delvis sykemeldt i perioder av svangerskapet. 13 % var sykemeldt i fødselspermisjonen, samtlige av disse var 100 % sykemeldt. Kvinnene hadde i gjennomsnitt hatt plager rundt 8 måneder ved inklusjonstidspunktet, men variasjonen var stor (fra 1 til 36 mnd).

Smertelokalisasjon ble angitt ved validert smertetegning, og 25 % av kvinnene markerte smerter i området rundt alle tre bekkenledd. Kun symfysemerter ble markert av 12 % i dette materialet, og 63 % viste til dorsale bekkensmerter eller annen smertelokalisering. Ved spørsmål om smertevariasjon rapporterte 11 % av kvinnene å ikke ha smertefri perioder i løpet av dagen. 57 % hadde av og til smertefri perioder, og 32 % hadde ofte smertefri perioder i løpet av dagen.

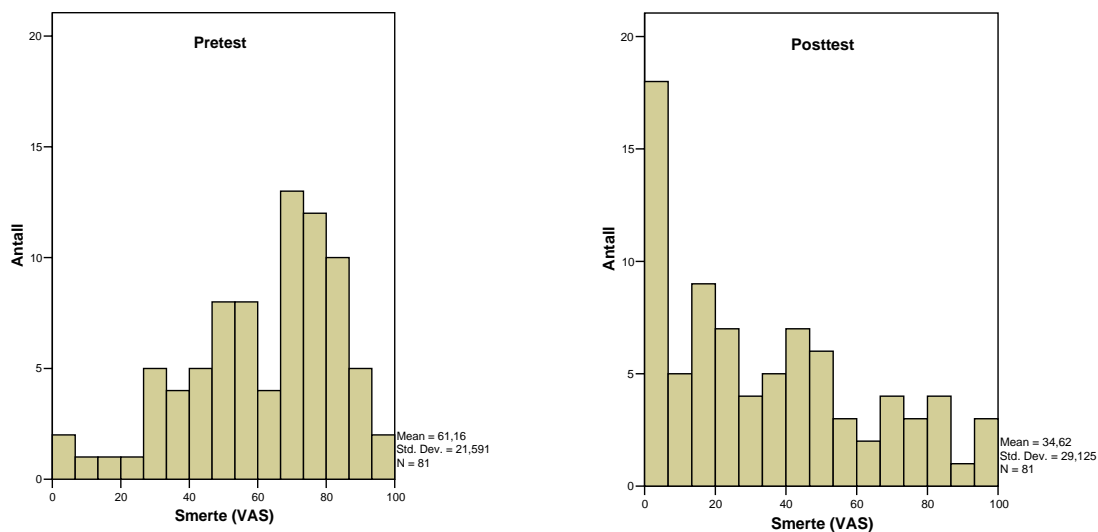
**Tabell 1: Deskriptive data (bakgrunnsvariabler) for 81 inkluderte kvinner ved pretest**

	N	Gjennomsnitt (SD)	Range (min-max)
Alder (år)	81	32.3 (3.9)	22-41
Utdanning (antall år)	81	16.1 (2.5)	9-23
Vekt (kg)	80	68 (10.5)	53-105
Høyde (cm)	81	168 (5.7)	152-180
Antall barn	81	1.7 (0.8)	1-4
Gift/ samboende (n, %)	81	79 (97.5 %)	
Antall uker siden fødsel	81	9.8 (3.3)	4-18
Varighet av plager (mnd)	81	7.6 (4.1)	1-36
Sykemeldt i svangerskapet (n, %)	79	64 (76 %)	
Sykemeldt i fødselspermisjon (n, %)	78	10 (13 %)	

Av de 42 kvinnene som hadde født flere barn, hadde 32 hatt bekkenløsning i tidligere svangerskap. Kun 2 kvinner rapporterte å ha skadet bekkenet tidligere, for eksempel ved fall eller ulykke. Nesten halvparten (48 %) hadde hatt vondt i korsryggen før siste svangerskap. 72 % av kvinnene hadde drevet regelmessig fysisk trening i løpet av siste året før det aktuelle svangerskapet. Trening var her definert til aktivitet mer enn 30 minutter pr. økt som resulterte i andpustenhet og svette (Siconolfi, Lasater, Snow, & Carleton, 1985). De hyppigst brukte treningsaktivitetene var turgang og aerobic/trimparti. Ti kvinner hadde drevet løpetrening. Kun én var med i konkurranseidrett.

## 4.2 Smerte og funksjon

Histogrammene for intensitet av kveldssmerte målt ved VAS viste at dataene fordelte seg over hele skalaen både ved pretest og posttest (fig. 1). Ved pretest var gjennomsnittet for skår på VAS 61 mm (SD 22, CI 56, 66). Spredningen var fra ”ingen smerte” til ”uutholdelig smerte”, men pasientgruppen samlet skåret relativt høyt for smerte om kvelden. Kun 5 % skåret lavere enn 20 mm på VAS ved pretest, og samtidig var det 40 % av kvinnene som skåret høyere enn 70 mm.



**Fig. 1: Fordeling av rapportert smerteintensitet om kvelden målt ved VAS (0-100 mm) ved pretest og posttest**

Ved posttest var dette bildet endret til en gjennomsnittlig skår på 35 mm (SD 29, CI 28, 41). På dette tidspunktet skåret 40 % av kvinnene lavere enn 20 mm på VAS, 17 % skåret høyere enn 70 mm. Dataene hadde ikke optimal normalfordeling ved pretest, og ved posttest viste histogrammet en tydelig forskyvning til venstre på VAS-skala. Denne forskyvningen viste altså at mange kvinner rapporterte å være smertefri eller ha minimale smerter ved posttest.

Funksjon målt ved ODI viste en tilnærmet normalfordeling ved pretest (fig. 2). Gjennomsnittlig skår på ODI var 40 poeng (SD 14, CI 37, 44). Spredningen av data var stor her også, men ingen skåret helt på minimum eller maksimum av skalaen. 10 % av kvinnene hadde skår lavere enn 20 poeng, og kunne ut fra dette kategoriseres til å ha minimale funksjonsproblemer (Fairbank et al., 2000). 52 % av kvinnene skåret høyere enn 40 poeng og rapporterte dermed betydelige funksjonsproblemer.

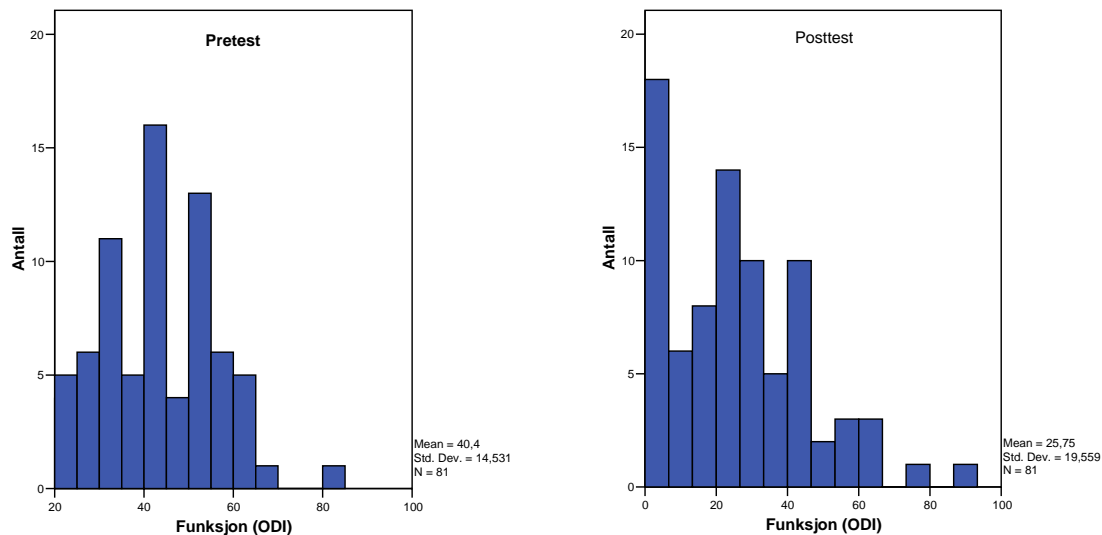


Fig. 2: Fordeling av rapportert funksjon ved skår på Oswestry Disability Index (0-100 poeng) ved pretest og posttest

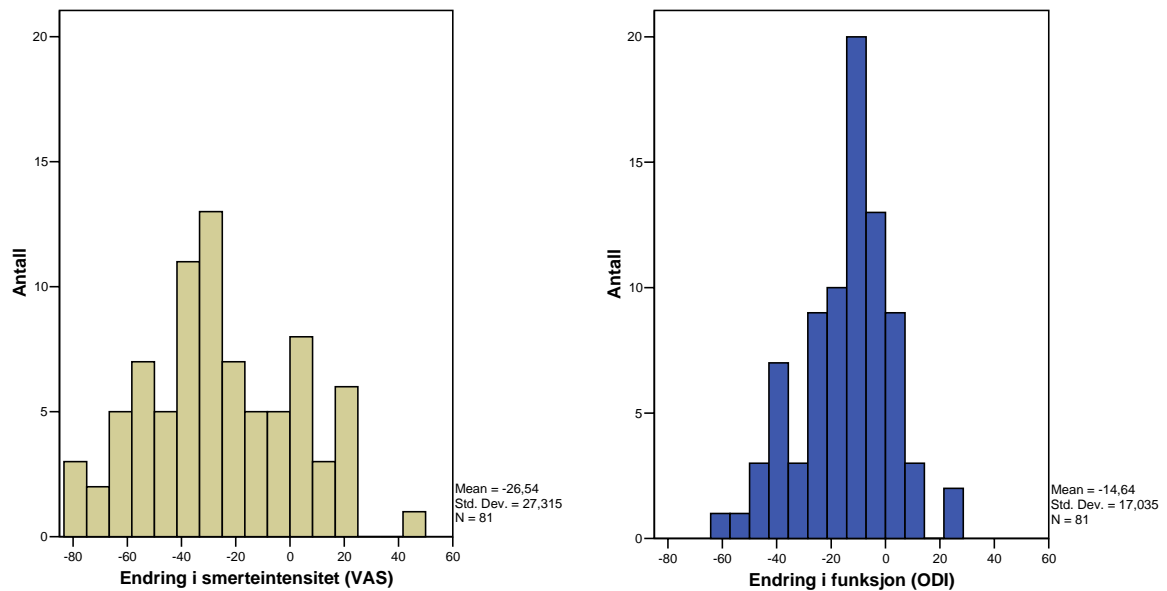
Ved posttest var det en tydelig reduksjon i skår på ODI. Gjennomsnitt var 26 poeng (SD 20, CI 21, 30), og dataene var ikke lenger normalfordelt. 40 % av kvinnene skåret lavere enn 20 poeng, 18 % skåret høyere enn 40. Dette viste at det var tre ganger så mange kvinner som hadde minimale funksjonsproblemer ved posttest. Samtidig var antallet kvinner med betydelige funksjonsproblemer redusert med ca to tredjedeler.

#### 4.2.1 Endring i smerte og funksjon fra før til etter behandling

Fordelingen av skår på både smerte og funksjon var betydelig endret fra før til etter behandling. Det kunne derfor være av interesse å se nærmere på hvordan de individuelle endringene fordelte seg i materialet.

Endringer fra pretest til posttest viste tilnærmet normalfordelte data både for VAS og ODI (fig. 3). Gjennomsnittlig reduksjon i intensiteten av kveldssmerte fra pretest til posttest målt ved VAS var 26 mm (SD 27, CI -33, -20). Denne reduksjonen var statistisk signifikant ( $p < 0.001$ ). 78 % av kvinnene

rapporterte smertereduksjon fra pretest til posttest. 18.5 % rapporterte smerteøkning, 2.5 % hadde uendret smertenivå.



**Fig. 3: Endring i rapportering av smerteintensitet om kvelden målt ved VAS (0-100 mm) og funksjon målt ved ODI (0-100 poeng) fra pretest til posttest.**

Gjennomsnittlig reduksjon i skår på ODI fra pretest til posttest var 15 poeng (SD 17, CI -18, -11). Denne reduksjonen var også statistisk signifikant ( $p < 0.001$ ). 83 % av kvinnene rapporterte bedring målt ved ODI, 6 % hadde uendret funksjon, og 11 % viste forverring fra pretest til posttest.

## 4.3 Kliniske tester

### 4.3.1 Active straight leg raise (ASLR) test

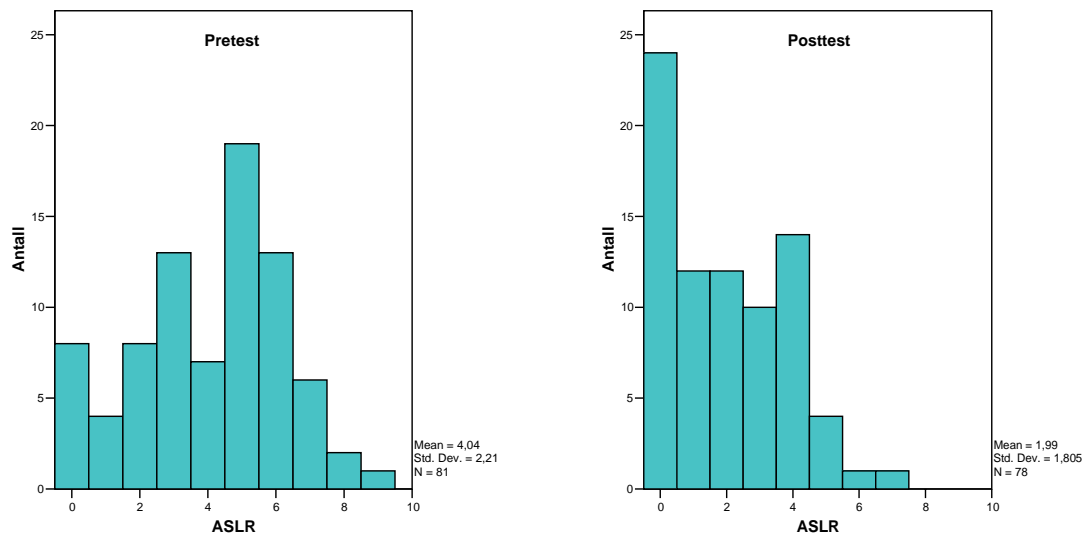


Fig. 4: Fordeling av skår på ASLR (0- 10 poeng) ved pretest og posttest

ASLR-skår varierte ved pretest fra 0 til 9, med en gjennomsnittlig skår på 4.0 (SD 2.2, CI 3.5, 4.5) (fig. 4) og median på 5. Åtte kvinner skåret 0 ved pretest. Ved posttest var fordelingen forskjøvet mot venstre. 24 kvinner skåret da 0 på testen, og kun 6 kvinner skåret høyere enn 4. Gjennomsnittlig skår var endret til 2.0 (SD 1.8, CI 1.6, 2.4), median var også 2 ved posttest.

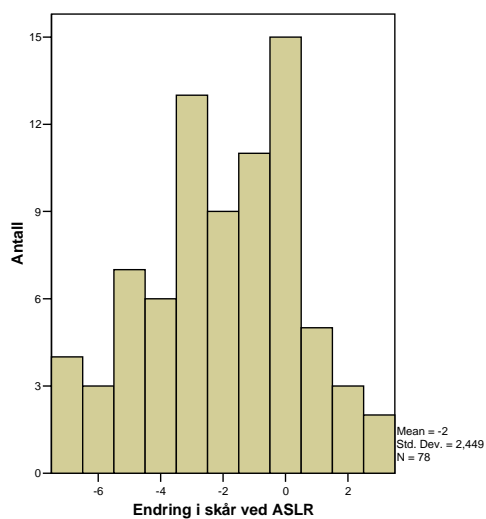


Fig. 5: ASLR, endring i skår fra pretest til posttest

Endring i skår ved ASLR fra pretest til posttest viste en tilnærmet normalfordeling av dataene (fig. 5). Gjennomsnittlig endring var -2 (SD 3, CI -3, -2). Dette var en statistisk signifikant reduksjon ( $p < 0.001$ ). 68 % av kvinnene skåret lavere på ASLR ved posttest, men med ulik grad av bedring vurdert ut fra denne testen. 19 % skåret uendret, 13 % viste høyere skår.

### 4.3.2 Bilaterale tester

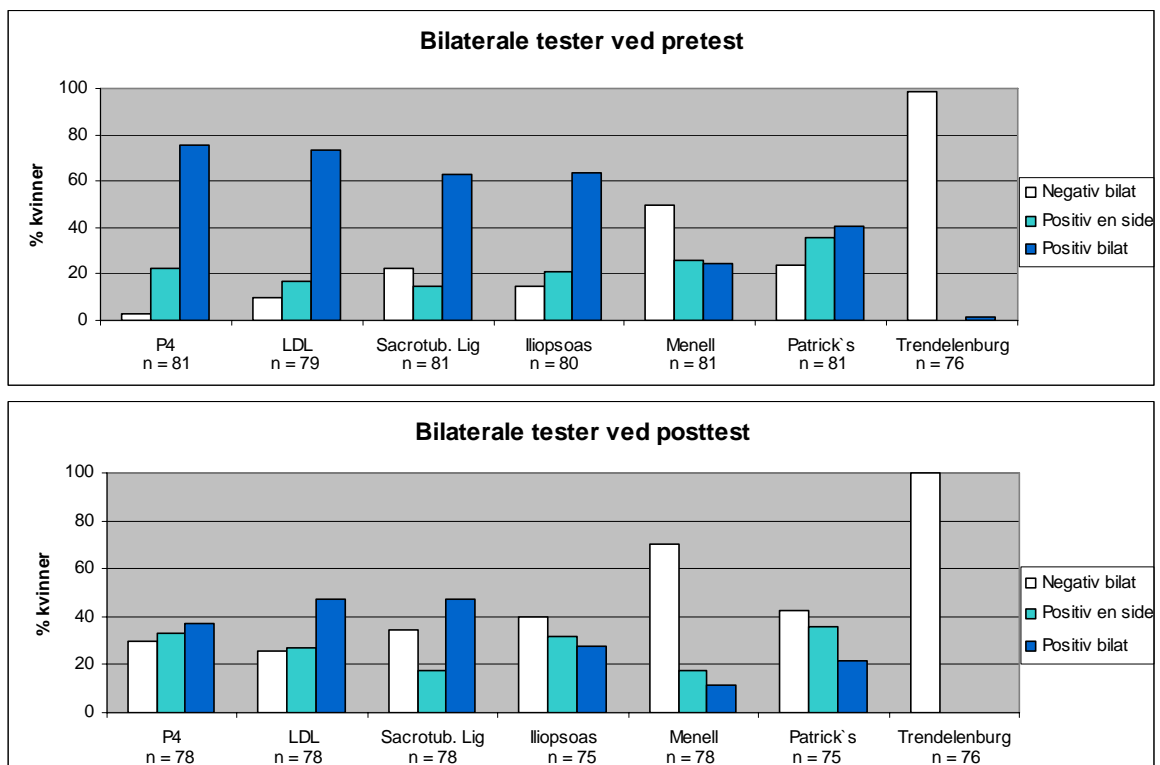


Fig. 6: Fordeling av skår for bilaterale tester ved pretest og posttest

Trendelenburgs test skilte seg klart ut i den grafiske framstillingen av bilaterale tester ved pretest ved at samtlige pasienter med unntak av én skåret negativt (fig. 6). Denne testen fanget derfor ikke opp noe aspekt ved de aktuelle plagene i dette materialet, og ble på det grunnlaget ikke tatt med i videre analyser. Menell's test skilte seg også ut ved å ha en klar overvekt av bilateral negativ skår ved pretest, halvparten av pasientene rapporterte ingen smerteprovokasjon ved denne testen. Patrick's test viste en svak overvekt av bilateralt positive skår. For de øvrige bilaterale testene var det stor overvekt av positiv skår på begge sider. P4 viste størst andel bilateralt positive (75 %) og lavest andel bilateralt negative (3 %) skår. Palpasjon av LDL hadde noen færre bilateralt positive (73 %), men hadde for øvrig et lignende skåringsmønster som P4. Tilsvarende mønster viste seg også for palpasjon av sacrotuberale ligament og iliopsoas med ca 63 % bilateralt positive skår.

### Endring i skår på bilaterale tester fra før til etter behandling

Skår på bilaterale tester viste betydelig endring fra pretest til posttest (fig. 6). Unntaket var Trendelenburgs test, denne var uendret negativ fra pretest. Generelt viste de kliniske testene endring i retning av større andel negative skår. De to ligamenttestene palpasjon av LDL og palpasjon av sacrotuberale ligament hadde fortsatt klar overvekt av bilateralt positive skår. P4 viste et lignende skåringsmønster som LDL og palpasjon av sacrotuberale ligament. P4 var også den testen som hadde størst reduksjon i andel bilateralt positive skår, fra 75 % til 37 %. Palpasjon av m. iliopsoas viste nesten like stor endring som P4 fra pretest til posttest. Menell's test skilte seg ut ved pretest med den største andelen bilateralt negative skår (49 %), denne andelen var øket til 70 % ved posttest. Palpasjon av sacrotuberale ligament endret seg mindre fra pretest til posttest enn palpasjon av LDL. Endring fra pretest til posttest var statistisk signifikant ( $p < 0.02$ ) for alle bilaterale tester unntatt palpasjon av sacrotuberale ligament og Trendelenburgs test.

#### 4.3.3 Sideuavhengige tester

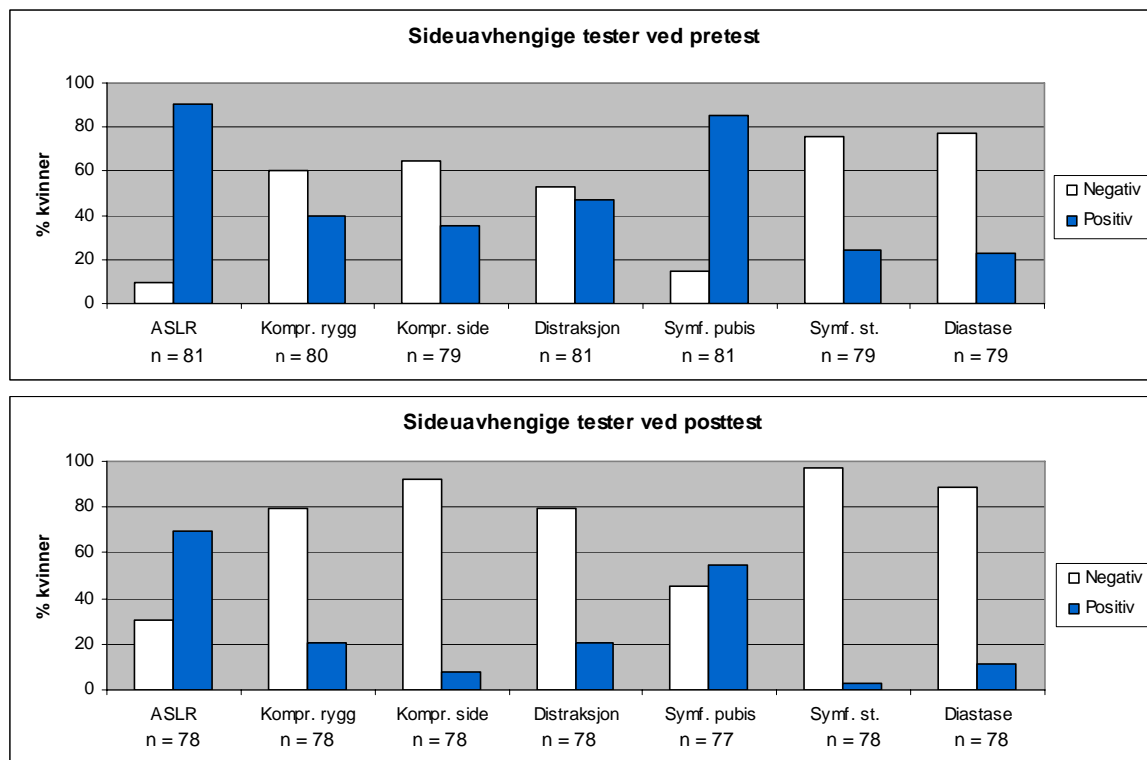


Fig. 7: Fordeling av skår for sideuavhengige tester ved pretest og posttest

ASLR ble først analysert som kontinuerlig variabel (kap. 4.3.1). Testen ble også analysert etter en dikotomisering der 0 representerte negativ skår og 1 til 10 positiv skår (fig. 7). Ved pretest var ASLR



den testen som viste høyest andel positive skår (90 %), deretter kom provokasjonstest av symfysis pubis med 85 % positive skår. Distraksjon var den eneste testen med tilnærmet lik andel positive og negative skår. Et fellestrekk for de øvrige sideuavhengige testene var en noe større andel negative enn positive skår. Kompresjonstest ble utført både i ryggliggende og sideliggende utgangsstilling. Disse to utgangsstillingene for testen skåret relativt likt, men med noen flere positive i ryggliggende enn i sideliggende. Registrering av diastase viste størst andel negative skår (77 %). Symfysesmerter i stående viste også en stor andel negative skår (76 %). Denne testen ga dermed et resultat som var svær forskjellig fra provokasjonstest av symfysis pubis. Samtlige som skåret positivt på symfysesmerter i stående, skåret samtidig positivt på test av symfysis pubis. Test av symfysesmerter i stående ble på dette grunnlaget tatt ut av videre analyser.

### *Endring i skår på sideuavhengige tester fra pretest til posttest*

Samtlige sideuavhengige tester viste reduksjon i andel positive skår fra pretest til posttest (fig. 7). Provokasjonstest av symfysis pubis viste størst reduksjon, men det var fortsatt en liten overvekt av positive skår (55 %) ved posttest. Mål av diastase i m. rectus abdominis var minst endret. ASLR viste fortsatt en klar overvekt av positive skår (69 %) ved posttest. Endringen var statistisk signifikant ( $p < 0.005$ ) for alle sideuavhengige tester unntatt mål av diastase.

## 4.4 Sammenhenger

### 4.4.1 Sammenhenger mellom kliniske tester

Korrelasjonsanalyser var av interesse for å framstille eventuelle sammenhenger mellom de ulike kliniske testene. Disse innledende resultatene ble benyttet som kartlegging før videre regresjonsanalyser av testenes forklaringsverdi for smerte og funksjon. Ved inklusjon til effektstudien ble ASLR, P4, LDL og provokasjonstest av symfysis pubis brukt som inklusjonstester (kap. 3.3). Dette medførte et lavt antall negative skår på disse testene, og dermed lite spredning i fordelingen av data ved pretest. Liten spredning i datamaterialet medfører vansker med å avdekke eventuelle sammenhenger. Ved posttest var gruppen endret slik at en andel av kvinnene også var godt fungerende (se resultater for ODI, kap. 4.2.1), og mange skåret negativt på inklusjonstestene. Med dette grunnlaget ble kun posttest valgt for analyse av sammenhenger mellom de kliniske testene.

Generelt viste korrelasjonsanalysene mange statistisk signifikante sammenhenger mellom de ulike testene (tab. 2). De ulike utgangsstillingene for kompresjonstest (ryggliggende og sideliggende) viste statistisk signifikant sammenheng. P4 viste statistisk signifikant sammenheng med LDL, distraksjon, Patrick's test og Menell's test. Resultatene for Menell's test viste også statistisk signifikant sammenheng med tester både av iliopsoas, symfyse og bakre strukturer.

Det var manglende sammenheng mellom flere av testene som etter beskrivelsen skulle gi informasjon om de samme anatomiske strukturene (tab. 2). For eksempel viste kompresjonstestene ingen sammenheng med P4. Disse ulike testene skulle i følge aktuell litteratur være provokasjonstester for bakre strukturer rundt iliosacralleddene (Ostgaard et al., 1994a; Laslett et al., 1994).

Kompresjonstestene viste heller ikke sammenheng med direkte provokasjonstest av LDL. Samtidig framkom det statistisk signifikante sammenhenger mellom tester som etter beskrivelsen skulle gi informasjon om ulike anatomiske strukturer. For eksempel viste provokasjonstest av symfysis pubis sammenheng med LDL, Menell's test og palpasjon av iliopsoas.

ASLR viste ingen statistisk signifikant korrelasjon med kompresjon, distraksjon eller Patrick's test. Samtidig var det korrelasjon mellom ASLR og andre tester av strukturer rundt iliosacralleddene (P4, LDL, sacrotuberale ligament og Menell's test) samt korrelasjon med palpasjon av iliopsoas. Mål av diastase viste svak negativ korrelasjon med ASLR og manglende sammenheng med alle de andre testene. Dette resultatet skilte seg klart ut fra de øvrige i korrelasjonsanalysen (tab. 2).

Et relativt stort antall ( $n = 78$ ) var inkludert i analyser av sammenheng mellom testene. Med dette antallet vil også svake sammenhenger kunne være statistisk signifikante. Mange tester i denne analysen viser statistisk høysignifikant sammenheng samtidig som Pearson korrelasjonskoeffisient og kjikvadratet viser svake til moderate sammenhenger (tab. 2). Analysene viser med disse resultatene på tross av høyt statistisk signifikansnivå noe usikkerhet rundt klinisk relevans av de aktuelle sammenhengene.

Tabell 2: Sammenheng mellom kliniske tester ved posttest. r = Pearson korreleksjonskoeffisient,  $\chi^2$  = kjikvadrat. p < 0.05 er vist med uthevet skrift

		P4	LDL	Sacrotub. ligament	Kompresj. ryggligg.	Kompresj. sideligg.	Distraksjon	Menell's test	Patrick's test	Iliopsoas	Symfysis pubis	Diastase n = 78
ASLR n = 78	r p	<b>0.46</b> <b>&lt; 0.001</b>	<b>0.43</b> <b>&lt; 0.001</b>	<b>0.23</b> <b>0.04</b>	0.22 0.06	0.16 0.15	0.16 0.15	<b>0.38</b> <b>0.001</b>	0.16 0.18 n=75	<b>0.47</b> <b>&lt; 0.001</b> n = 75	0.23 0.05 n = 77	<b>- 0.24</b> <b>0.03</b>
P4 n = 78	$\chi^2$ p		<b>12.5</b> <b>0.014</b>	2.9 0.584	5.4 0.072	5.1 0.058	<b>15.7</b> <b>&lt;0.001</b>	<b>11.6</b> <b>0.014</b>	<b>14.3</b> <b>0.006</b>	4.4 0.359	4.3 0.117	0.6 0.822
LDL n = 78	$\chi^2$ p			<b>16.2</b> <b>0.002</b>	1.6 0.449	3.1 0.126	4.3 0.135	<b>10.3</b> <b>0.025</b>	4.2 0.385	<b>10.8</b> <b>0.029</b>	<b>6.6</b> <b>0.037</b>	0.4 0.905
Sacrotuberale ligament n = 78	$\chi^2$ p				2.0 4.0	1.1 0.745	2.4 0.305	<b>14.6</b> <b>0.003</b>	2.3 0.709	2.0 0.749	5.5 0.064	1.9 0.462
Kompresjon ryggliggende n = 78	$\chi^2$ p					<b>15.7</b> <b>0.001</b>	0.25 0.729	4.3 0.109	4.3 0.113	<b>6.4</b> <b>0.036</b>	1.0 0.317	0.55 0.676
Kompresjon sideliggende n = 78	$\chi^2$ p						3.5 0.097	<b>6.6</b> <b>0.029</b>	3.3 0.170	1.8 0.427	2.2 0.212	0.85 p = 1.0
Distraksjon n = 78	$\chi^2$ p							<b>13.2</b> <b>0.001</b>	5.8 0.053	4.5 0.099	2.5 0.118	0.55 0.676
Menell n = 78	$\chi^2$ p								4.8 0.310	<b>11.5</b> <b>0.014</b>	<b>10.5</b> <b>0.005</b>	3.1 0.172
Patrick n = 75	$\chi^2$ p									<b>9.5</b> <b>0.046</b> n = 72	3.8 0.151 n = 74	1.1 0.593 n = 75
Iliopsoas n = 75	$\chi^2$ p										<b>12.3</b> <b>0.002</b> n = 74	1.8 0.414 n = 75
Symfysis pubis n = 77	$\chi^2$ p											1.85 0.286 n = 77

#### **4.4.2 Sammenhenger mellom kliniske tester og egenrapportert smerte og funksjon**

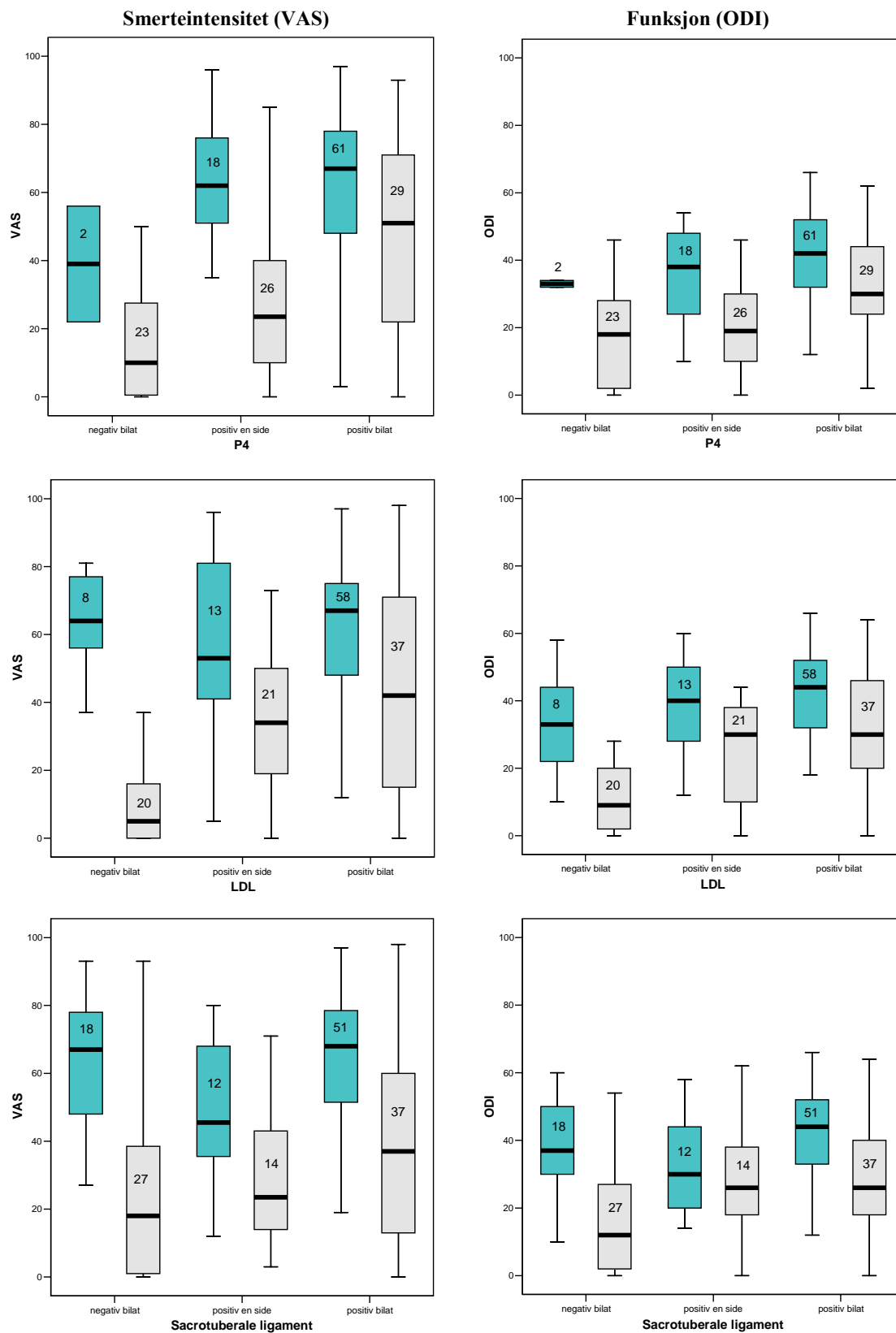
Kvinnene i dette materialet ble bedt om å beskrive smerteintensitet og funksjonsnivå ved hjelp av standardiserte måleinstrumenter (VAS og ODI). I tillegg ble kliniske tester benyttet for å gjøre mer avgrensede målinger av smerteprovokasjon og funksjon i bekkenrelaterede strukturer. Materialet ble analysert med tanke på å framstille eventuelle sammenhenger mellom disse ulike måle metodene for registrering av kvinnenenes bekkenplager.

Resultatene viste at forholdet mellom de kliniske testene og egenrapportert smerte og funksjon endret seg fra pretest til posttest. Ved pretest framkom relativt jevnt høy skår på VAS og ODI ved både negativ og positiv test (fig. 8-11). Kvinnene rapporterte å være mye plaget selv om den enkelte kliniske test var negativ. Ved posttest var dette endret slik at de med negative tester gjennomsnittlig skåret lavest på VAS og ODI, og det var gjennomgående samsvar mellom negative tester og reduserte plager. Kvinner med positive tester hadde da høyere smerteintensitet og mer nedsatt funksjon enn de med negative tester. Forskjellene mellom pretest og posttest var altså størst for negative tester og minst for positive tester. Det var gjennomsnittlig lavere skår på VAS og ODI ved posttest sammenlignet med pretest, men spredningen i skår var stor på begge tidspunkt.

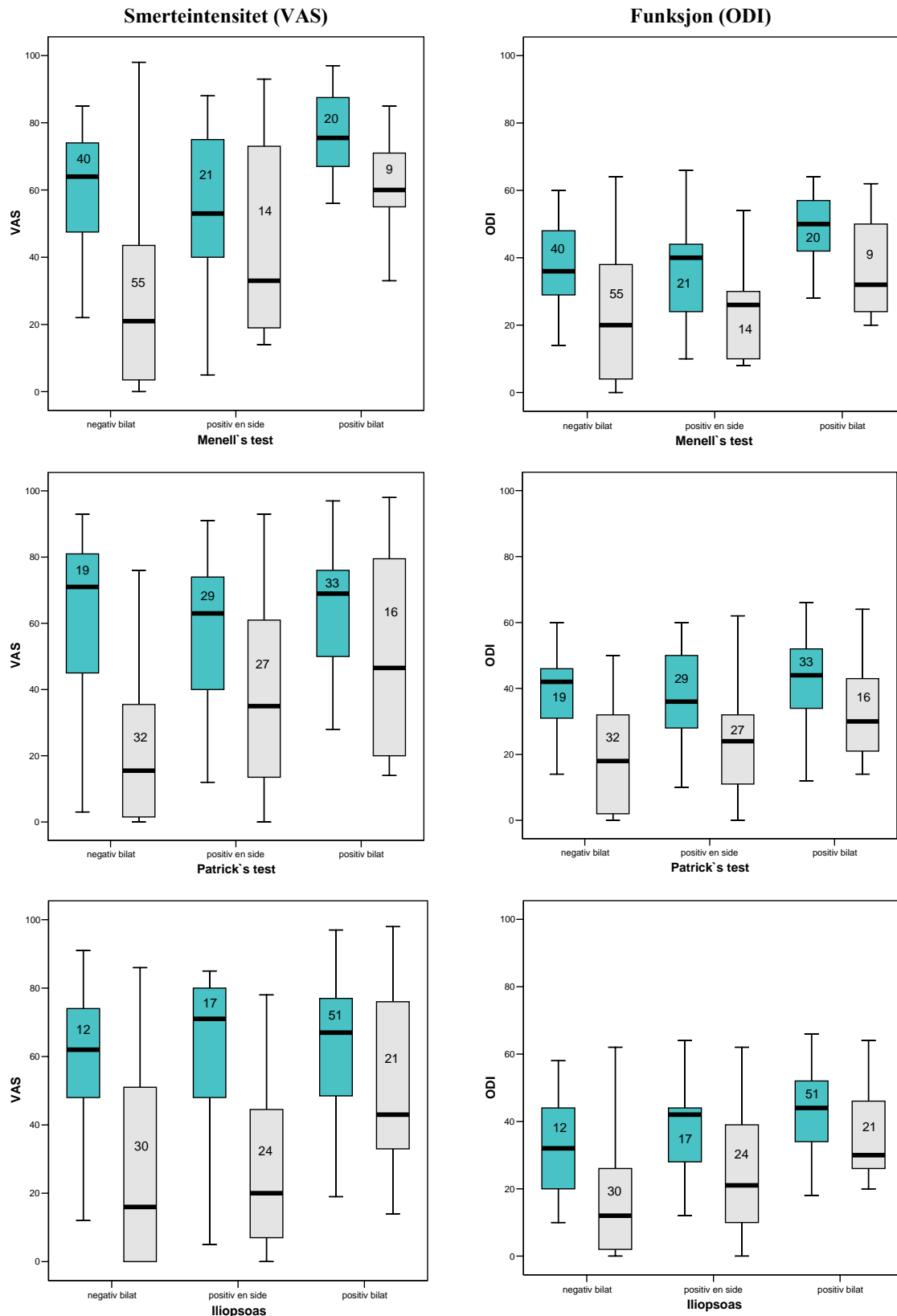
##### ***Bilaterale tester***

Ved pretest framkom relativt jevnt høy skår på VAS og ODI ved både negativ, ensidig positiv og bilateral positiv test (fig. 8 og 9). Ved posttest var dette endret slik at gruppen med bilateralt negative tester skåret lavest på både VAS og ODI. Resultatene viste også gjennomsnittlig lavere skår for gruppen med ensidig positive tester sammenlignet med gruppen for bilateralt positive tester.

Av de bilaterale testene var det kun Menell's test som viste statistisk signifikant sammenheng med både VAS og ODI ved pretest. Palpasjon av iliopsoas og LDL viste en svak (til moderat) korrelasjon med ODI. De øvrige bilaterale testene (P4, sacrotuberale ligament og Patrick's test) viste ingen statistisk signifikant korrelasjon med VAS eller ODI ved pretest. Ved posttest viste de 6 bilaterale testene moderat korrelasjon (fra  $r = 0.23$  til  $r = 0.46$ ) med både VAS og ODI (tab. 3).



**Fig. 8: Sammenheng mellom smerteintensitet målt ved VAS (0-100 mm) (venstre kolonne), funksjon målt ved ODI (0-100 poeng) (høyre kolonne) og bilaterale tester ved pretest ■ og posttest □. Antall og median (horizontal strek) er angitt i hver boks. Boksens nedre grense markerer 25 % fraktile, øvre grense 75 % fraktile. Endepunktene viser høyeste og laveste skår for gruppen.**



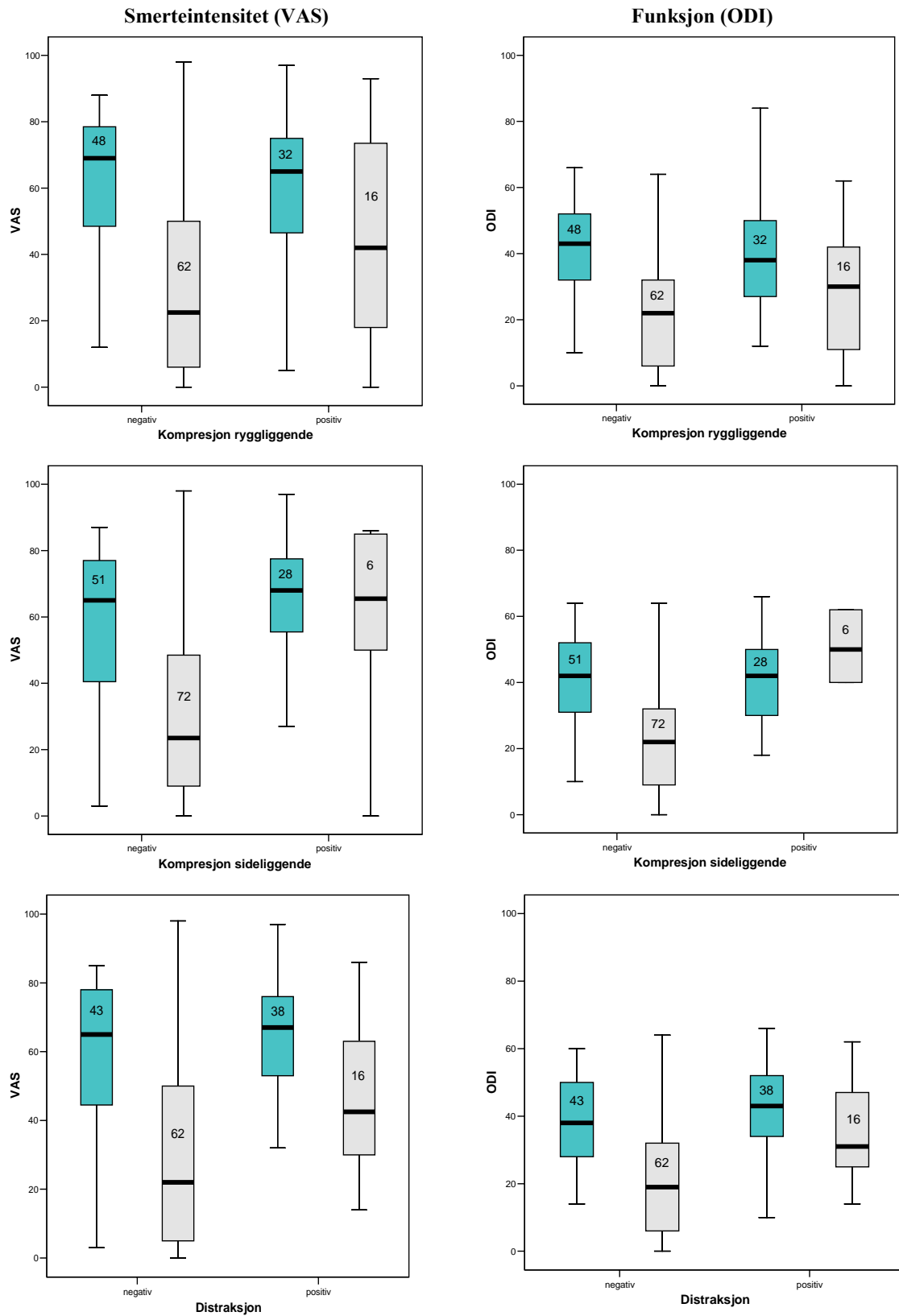
**Fig. 9:** Sammenheng mellom smerteintensitet målt ved VAS (0-100 mm) (venstre kolonne), funksjon målt ved ODI (0-100 poeng) (høyre kolonne) og bilaterale tester ved pretest ■ og posttest □ (forts. fra fig. 8). Antall og median (horisontal strek) angitt i hver boks. Boksenes nedre grense markerer 25 % fraktil, øvre grense 75 % fraktil. Endepunktene viser høyeste og laveste skår for gruppen.

**Tabell 3: Korrelasjon mellom VAS, ODI og kliniske tester ved pretest og posttest.  $r$  = Pearson korrelasjonskoeffisient,  $p < 0.05$  er vist med uthevet skrift, \* viser signifikant korrelasjon ved  $p < 0.01$**

	Pretest (n = 81)			Posttest (n = 78)		
	n	VAS	ODI	n	VAS	ODI
ASLR	81	<b><math>r = 0.35</math> <math>p = 0.001^*</math></b>	<b><math>r = 0.41</math> <math>p &lt; 0.001^*</math></b>	78	<b><math>r = 0.51</math> <math>p &lt; 0.001^*</math></b>	<b><math>r = 0.55</math> <math>p &lt; 0.001^*</math></b>
P4	81	$r = 0.06$ $p = 0.59$	$r = 0.18$ $p = 0.11$	78	<b><math>r = 0.44</math> <math>p &lt; 0.001^*</math></b>	<b><math>r = 0.39</math> <math>p &lt; 0.001^*</math></b>
LDL	79	$r = 0$ $p = 1.0$	<b><math>r = 0.23</math> <math>p = 0.04</math></b>	78	<b><math>r = 0.45</math> <math>p &lt; 0.001^*</math></b>	<b><math>r = 0.48</math> <math>p &lt; 0.001^*</math></b>
Sacrotuberale ligament	81	$r = 0.03$ $p = 0.77$	$r = 0.20$ $p = 0.07$	78	<b><math>r = 0.23</math> <math>p = 0.045</math></b>	<b><math>r = 0.27</math> <math>p = 0.02</math></b>
Kompresjon ryggliggende	80	$r = -0.01$ $p = 0.94$	$r = -0.1$ $p = 0.38$	78	<b><math>r = 0.23</math> <math>p = 0.046</math></b>	$r = 0.18$ $p = 0.13$
Kompresjon sideliggende	79	$r = 0.19$ $p = 0.09$	$r = 0.09$ $p = 0.43$	78	<b><math>r = 0.27</math> <math>p = 0.02</math></b>	<b><math>r = 0.33</math> <math>p = 0.003^*</math></b>
Distraksjon	81	$r = 0.12$ $p = 0.28$	$r = 0.16$ $p = 0.15$	78	$r = 0.24$ <b><math>p = 0.04</math></b>	$r = 0.33$ <b><math>p = 0.003^*</math></b>
Menell's test	81	$r = 0.22$ $p = 0.05$	<b><math>r = 0.30</math> <math>p = 0.007</math></b>	78	<b><math>r = 0.38</math> <math>p = 0.001^*</math></b>	<b><math>r = 0.28</math> <math>p = 0.01</math></b>
Patrick's test	81	$r = 0.05$ $p = 0.64$	$r = 0.18$ $p = 0.11$	75	<b><math>r = 0.38</math> <math>p = 0.001^*</math></b>	<b><math>r = 0.33</math> <math>p = 0.004^*</math></b>
Iliopsoas	80	$r = 0.03$ $p = 0.79$	<b><math>r = 0.31</math> <math>p = 0.006^*</math></b>	75	<b><math>r = 0.35</math> <math>p = 0.002^*</math></b>	<b><math>r = 0.46</math> <math>p &lt; 0.001^*</math></b>
Symfysis pubis	81	$r = -0.08$ $p = 0.49$	$r = 0.02$ $p = 0.89$	77	<b><math>r = 0.27</math> <math>p = 0.02</math></b>	<b><math>r = 0.37</math> <math>p = 0.001^*</math></b>
Diastase	79	$r = -0.22$ $p = 0.06$	<b><math>r = -0.28</math> <math>p = 0.01</math></b>	78	$r = -0.10$ $p = 0.39$	$r = -0.05$ $p = 0.64$
VAS	81		<b><math>r = 0.45</math> <math>p &lt; 0.001^*</math></b>	78		<b><math>r = 0.79</math> <math>p &lt; 0.001^*</math></b>

### Sideuavhengige tester

Trenden var som beskrevet over også for sideuavhengige tester (fig. 10 og 11). Disse testene (kompresjon ryggliggende og sideliggende, distraksjon, provokasjon av symfysis pubis og diastase-mål) viste ingen statistisk signifikante korrelasjoner med VAS eller ODI ved pretest, unntatt mål av diastase som viste svak negativ korrelasjon med ODI (tab. 3). Ved posttest var dette endret slik at kompresjon sideliggende, distraksjon og test av symfysis pubis viste statistisk signifikant korrelasjon med både VAS og ODI. Kompresjon ryggliggende viste statistisk signifikant sammenheng med VAS, men ingen sammenheng med ODI. Mål av diastase skilte seg også her ut fra de andre testene, resultatet viste manglende sammenheng med både VAS og ODI.



**Fig. 10: Sammenheng mellom smerteintensitet målt ved VAS (0-100 mm) (venstre kolonne), funksjon målt ved ODI (0-100 poeng) (høyre kolonne) og sideuavhengige tester ved pretest ■ og posttest □. Antall og median (horisontal strek) er angitt i hver boks. Boksenes nedre grense markerer 25 % fraktil, øvre grense 75 % fraktil. Endepunktene viser høyeste og laveste skår for gruppen.**



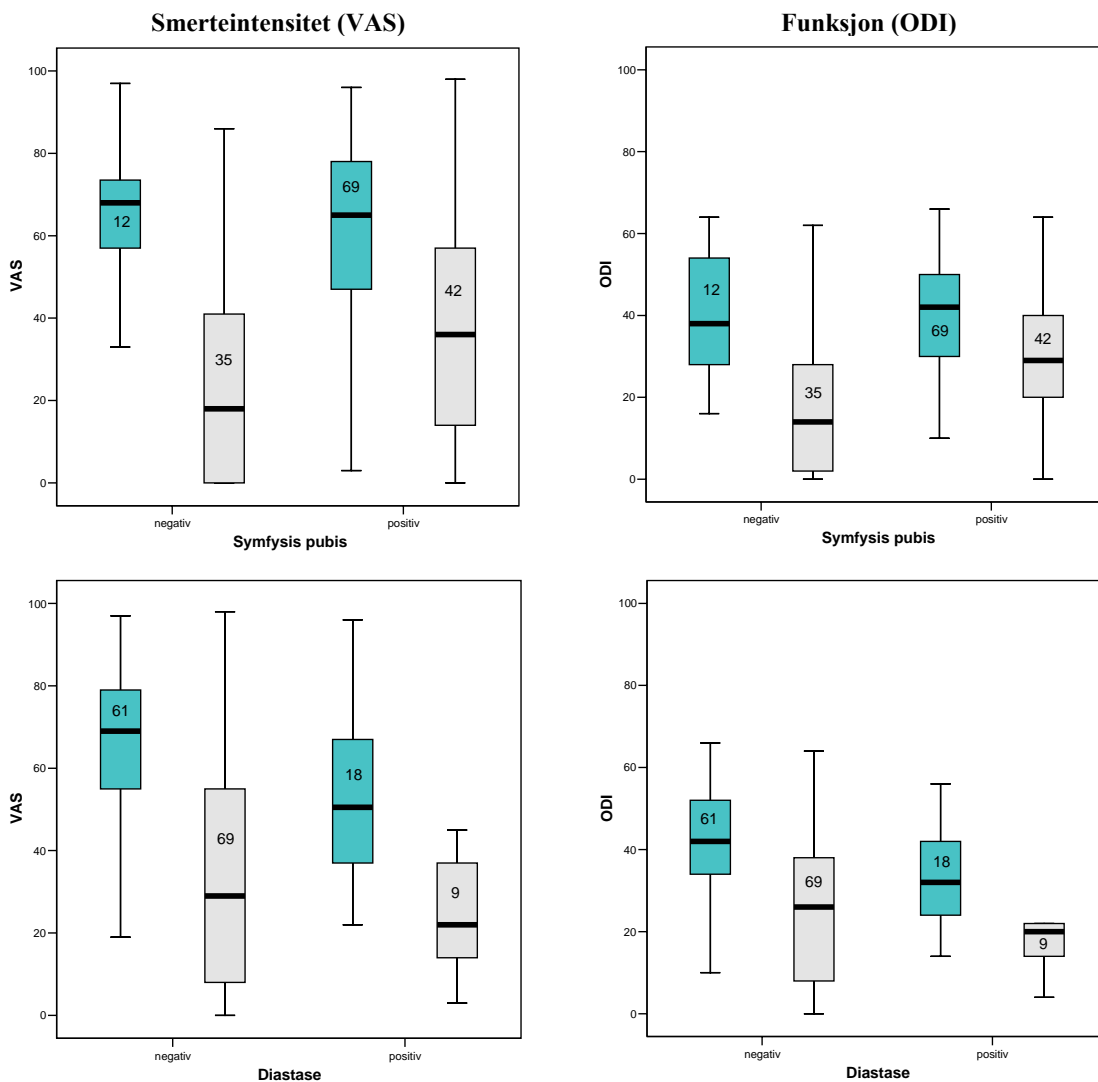


Fig. 11: Sammenheng mellom smerteintensitet målt ved VAS (0-100 mm) (venstre kolonne), funksjon målt ved ODI (0-100 poeng) (høyre kolonne) og sideuavhengige tester ved pretest ■ og posttest □ (forts. fra fig. 10). Antall og median (horisontal linje) er angitt i hver boks. Boksenes nedre grense markerer 25 % fraktil, øvre grense 75 % fraktil. Endepunktene viser høyeste og laveste skår for gruppen.

## ASLR

ASLR var skåret fra 0 til 10, men viste et lignende skåringsmønster som beskrevet for de øvrige testene (fig. 12). Også her var det for alle tallskår høyere smerteintensitet og mer nedsatt funksjon ved pretest enn ved posttest. Samtidig viste også pretest en svakt stigende trend med økende tallskår på ASLR, slik at kvinnene med lav skår på ASLR rapporterte å være noe mindre plaget enn de med høy skår også ved pretest.

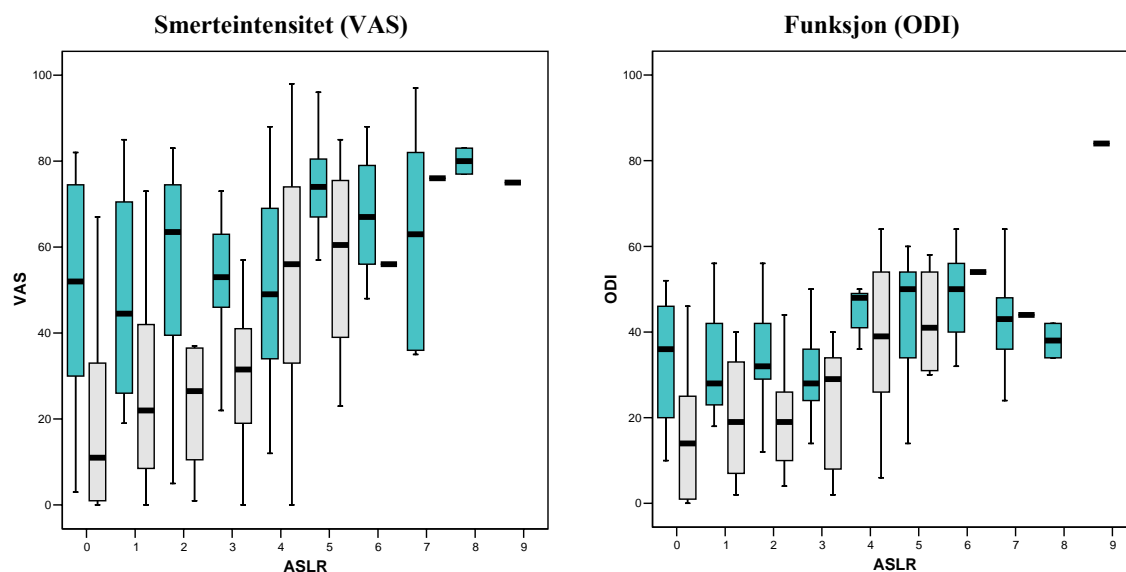


Fig. 12: Sammenheng mellom smerteintensitet målt ved VAS (0-100 mm) (venstre kolonne), funksjon målt ved ODI (0-100 poeng) (høyrekolonne) og skår ved ASLR (0-10) ved pretest ■ og posttest □. Fordeling av antall er vist i fig. 4.

ASLR viste statistisk signifikant sammenheng ( $p < 0.001$ ) med VAS og ODI både ved pretest og posttest. Ved pretest var det moderat korrelasjon med VAS ( $r = 0.35$ ) og ODI ( $r = 0.41$ ), ved posttest god korrelasjon både med VAS ( $r = 0.51$ ) og ODI ( $r = 0.55$ ) (tab. 3).

#### 4.4.3 Testenes selvstendige forklaringsverdi for smerte og funksjon

Med bakgrunn i det sammensatte bildet av sammenhenger i de presenterte resultatene, var det av interesse å undersøke nærmere hvilke kliniske tester som hadde sterkest assosiasjon til kvinnenes egen rapportering av smerte og funksjon. En regresjonsanalyse ville gi en indikasjon på det relative bidraget hver enkelt test kunne gi til forklaring av variasjonen i VAS og ODI. Kun posttest ble valgt for regresjonsanalyse på grunn av at korrelasjonsanalysene viste flest sammenhenger på dette tidspunktet og flere tester var benyttet som inklusjonstester. Det ble gjort baklengs trinnvis multipl regressjon med henholdsvis VAS og ODI som avhengige variabler og 10 kliniske tester som uavhengige variabler. Antallet tester ble redusert fra 14 til 10 på følgende måte: De to kompresjonstestene ble slått sammen. Trendelenburgs test ble tatt ut på grunnlag av at alle kvinner med unntak av én skåret negativt på denne testen allerede ved pretest. Provokasjonstest av symfysis pubis viste positiv skår hos samtlige som hadde symfysesmerter i stående. Da dette i tillegg var to ulike tester av de samme anatomiske strukturene, ble registrering av symfysesmerter i stående utelatt. Mål av diastase i rectus abdominis ble utelatt etter forutgående analyser som viste at det var lavt

antall med diastase allerede ved pretest, og ved posttest viste denne testen manglende sammenheng med alle de andre testene, unntatt svak negativ korrelasjon med ASLR.

Korrelasjonsanalysene for posttest indikerte ingen multikollinearitet. Korrelasjon med VAS var over 0.30 for 6 av 10 tester, korrelasjon med ODI var over 0.30 for 7 av 10 tester (tab. 3). De øvrige testene viste korrelasjon mellom 0.21 og 0.27. Høyeste korrelasjon mellom de uavhengige variablene var 0.55. Outliers ble undersøkt ved inspeksjon av scatterplot og normalplot for standardiserte residualer. Mahalanobis distance ble også analysert, og samtlige skår var lavere enn grensene for kritiske verdier. Forutsetningene for å gjennomføre baklengs trinnvis multipl regressjon ble på bakgrunn av dette vurdert som akseptable.

**Tabell 4: Assosiasjon mellom VAS, ODI og kliniske tester ved posttest (n = 75). B = ustandardisert regresjonskoeffisient,  $\beta$  = standardisert regresjonskoeffisient, analysert ved baklengs trinnvis multipl regressjon**

Uavhengige variabler	Avhengige variabler							
	VAS				ODI			
	B	95 % CI for B	$\beta$	p	B	95 % CI for B	$\beta$	p
<b>ASLR</b>	5.7	2.4, 9.0	0.35	0.001	4.4	2.2, 6.5	0.40	< 0.001
<b>LDL</b>	8.9	1.8, 16.0	0.26	0.015	6.3	1.6, 11.0	0.27	0.009
<b>Patrick`s</b>	10.7	3.7, 17.7	0.83	0.003	5.5	0.9, 10.2	0.22	0.02

ASLR, LDL og Patrick`s test viste statistisk signifikant assosiasjon både med smerte og funksjon (tab. 4). For VAS kunne 40 % av den totale variansen forklares ut fra disse tre kliniske testene (R square = 0.40). 42 % av variansen i ODI kunne forklares ut fra de samme testene (R square = 0.42).

Ut fra denne analysen ser vi at ett poeng endring i skår på ASLR i gjennomsnitt vil føre til ca 6 mm endring i VAS og ca 4 poeng endring i ODI. Konfidensintervallene er imidlertid ganske brede og det er derfor rimelig å forvente at ett poengs endring i skår på ASLR vil føre til mellom 2.4 og 9 mm i endring i VAS, og mellom 2.2 og 6.5 poeng i endring i ODI. For LDL og Patrick`s er det enda større konfidensintervaller. Resultatene viser at en enhets endring i skår på LDL (for eksempel fra bilateral positiv til ensidig positiv test) ville føre til 1.8 til 16 mm endring i VAS og 1.6 til 11 poeng endring i ODI. Tilsvarende endring i skår på Patrick`s test ville føre til 3.7 til 17.7 mm endring i VAS og 0.9 til 10.2 poeng endring i ODI.

Tilsvarende regresjonsanalyser med VAS og ODI som avhengige variabler ble også utført med ASLR dikotomisert slik at skår 0 tilsvarte negativ test og 1-10 positiv test. Dette resulterte i at ASLR ikke lenger viste statistisk signifikant assosiasjon med VAS eller ODI. Resultatene for LDL og Patrick`s test var uendret, men samtidig viste Menell`s test assosiasjon med VAS ( $p < 0.05$ ) og palpasjon av iliopsoas tilsvarende med ODI ( $p < 0.05$ ).

#### 4.4.4 Kombinasjon av kliniske tester – assosiasjon med smerte og funksjon

Ulike kombinasjoner av kliniske tester har blitt anbefalt ved undersøkelse av bekkenplager (kap. 2.2.2). I denne studien ble det benyttet et relativt høyt antall tester med varierende grad av pålitelighet (kap. 2.2.1). Med dette utgangspunktet var det av interesse å se i hvilken grad en sumskår av testene samsvarte med pasientenes egen rapportering av plager. 11 tester ble inkludert i denne summerte framstillingen av skår på kliniske tester. Trendelenburgs test, registrering av diastase-mål og symfysesmerter i stående ble utelatt i samsvar med tidligere analyser. Ved pretest var det ingen kvinner som hadde mindre enn 4 positive tester (fig. 13). Ved posttest var dette endret slik at 24 % av kvinnene skåret positivt på færre enn 4 kliniske tester. Sumskår var redusert fra gjennomsnittlig 8 til 5 positive tester, dette var en statistisk signifikant reduksjon ( $p < 0.001$ ) fra pretest til posttest.

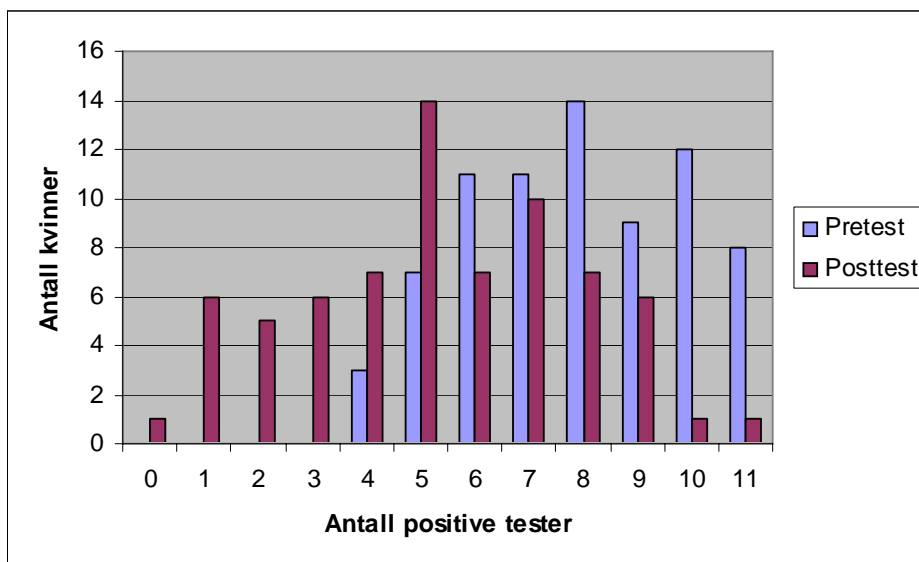


Fig. 13: Fordeling av antall positive tester ved pretest (n = 75) og posttest (n = 71)

Ved pretest var det ingen sammenheng mellom smerteintensitet og antall positive tester ( $r = 0.11$ ,  $p = 0.3$ ) (fig. 14). Derimot var det en sammenheng ved posttest ( $r = 0.62$ ,  $p < 0.001$ ). Figuren viser også at forholdet mellom smerter og antall positive tester endret seg fra før til etter behandling. Ved et gitt antall positive tester under 7, var smerteskår lavere ved posttest enn ved pretest.

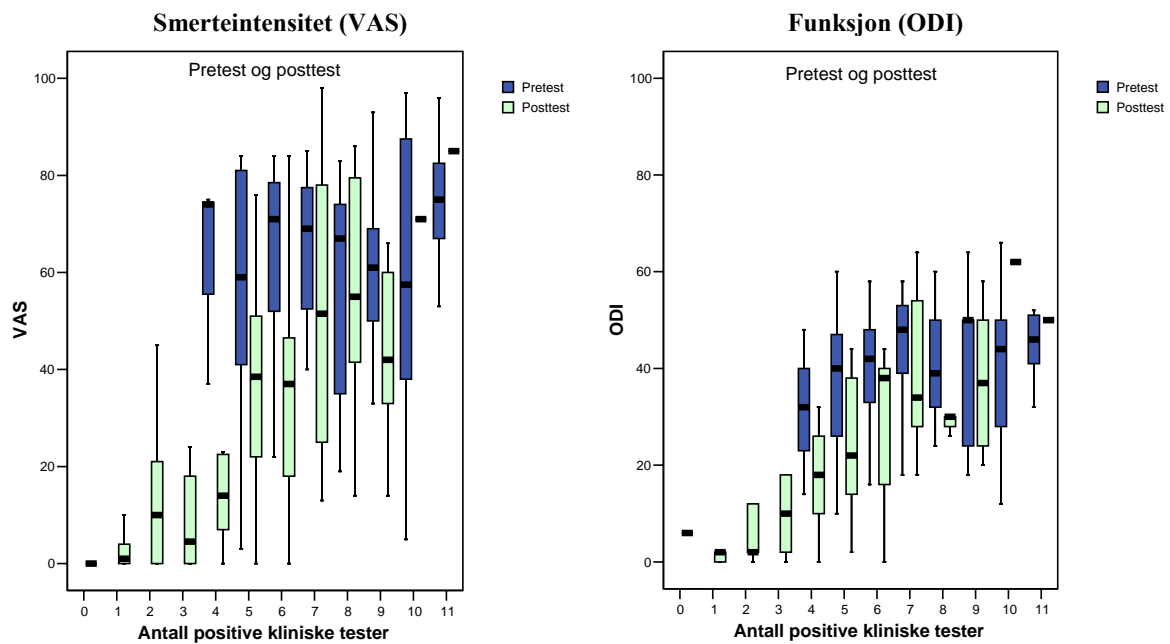


Fig. 14: Fordeling av antall positive tester ved sumskår av 11 tester relatert til skår på VAS (0- 100) og ODI (0- 100) ved pretest og posttest

Kvinnene rapporterte også betydelig grad av nedsatt funksjon ved pretest, men det var heller ikke her sammenheng mellom funksjon og antall positive tester ( $r = 0.21$ ,  $p = 0.6$ ). Ved posttest var dette endret slik at sammenhengen var god ( $r = 0.66$ ,  $p < 0.001$ ). Det var da tendens til en stigende trend for skår på ODI fra 0 til 7 positive tester (fig. 14). Ved høyere antall positive tester var det ingen ytterligere stigning i skår på ODI. Som for smerte, var også medianen for egenrapportert funksjon redusert fra pretest til posttest ved likt antall positive tester, men denne forskjellen ble utjevnet ved høyere antall positive tester. Ved posttest skåret samtlige kvinner med 4 eller færre positive tester lavere enn 20 poeng på ODI. Denne framstillingen viste samsvar mellom høyt antall positive tester og moderat til betydelig funksjonsnedsettelse målt ved ODI uavhengig av testtidspunkt. Få eller ingen positive tester viste også samsvar med minimal funksjonsnedsettelse.

## 4.5 Oppsummering av resultater

1. Ved pretest viste 7 av 14 kliniske tester over 75 % positive skår, de øvrige varierte fra 1 til 51 %. 11 tester viste statistisk signifikant reduksjon i positive skår fra pretest til posttest. Palpasjon av sacrotuberale ligament viste ikke signifikant reduksjon.
2. Kvinnene i dette materialet rapporterte statistisk signifikant reduksjon i smerteintensitet (målt ved VAS) og funksjon (målt ved ODI) fra pretest til posttest.

3. ASLR viste moderat korrelasjon med VAS og ODI ved pretest og god korrelasjon ved posttest. For de øvrige kliniske testene var det gjennomgående svak korrelasjon med VAS og ODI ved pretest. Ved posttest var dette endret slik at testene viste moderat korrelasjon med både VAS og ODI.
4. Det var et stort antall statistisk signifikante sammenhenger mellom de kliniske testene ved posttest. Det var ingen gjennomgående sammenheng mellom ulike provokasjonstester av antatt samme anatomiske strukturer.
5. Forholdet mellom kliniske tester og smerte og funksjon endret seg fra pretest til posttest. Kvinnene rapporterte lavere smerteintensitet og bedret funksjon ved likt antall positive tester ved posttest sammenlignet med pretest.
6. Det var samsvar mellom få positive kliniske tester og kvinnenenes egen rapportering av lite plager. Det var statistisk signifikant sammenheng mellom økende antall positive tester og økende smerte og funksjonsnedsettelse ved posttest.
7. ASLR, LDL og Patrick`s test kunne forklare ca 40 % av den totale variansen i VAS og ODI ved posttest, men usikkerheten i estimatet var relativt stor.

## 5. Diskusjon av metodiske valg

Diskusjon av studien er delt i to hovedavsnitt. Først diskuteres metodiske valg og hvilken betydning disse kan ha for tolkning av resultatene. Deretter diskuteres resultatene av den gjennomførte studien i lys av presentert teorigrunnlag og klinisk erfaring.

### 5.1 Design

Datainnsamlingen for denne studien ble utført som del av en effektstudie. Det ble i ettertid valgt å gjennomføre en kohortestudie for å gi en framstilling av hvordan hele pasientgruppen skåret på kliniske tester sammenlignet med egenrapportert smerte og funksjon. Dersom opprinnelig design hadde vært en kohortestudie, ville det være mer hensiktsmessig å planlegge inklusjon av alle som fødte barn i et område i en viss periode for å sammenligne kvinner med og uten plager.

I perioden for gjennomføring av effektstudien forelå det ingen planer om å bruke resultatene av skår på kliniske tester til særskilte analyser. Undersøker var i testsituasjonen blindet for fordeling til de to behandlingsgruppene. Med dette utgangspunktet ble det vurdert at undersøker ikke var under spesiell påvirkning med tanke på utfall av kliniske tester under testprosedyrene. I tillegg var veiledere sterkt inne med rådgivning i forhold til statistiske analyser, og disse hadde ikke deltatt i de kliniske undersøkelsene. På dette grunnlaget ble det betraktet som akseptabelt å selv gjennomføre de aktuelle statistiske analysene.

### 5.2 Utvalg

Kvinnene som ble rekruttert til studien var alle bosatt i et urbant miljø. Gjennomsnittsalder var 32 år mens landsgjennomsnittet for fødende kvinner i samme periode var 29 år (norgeshelsa.no, 2007). Det var en større andel førstegangsfødende (50 %) i dette utvalget enn på landsbasis (27 %) i samme tidsperiode.

Ved inklusjonstidspunktet for effektstudien hadde 34 % av norske kvinner i samme aldersgruppe høyskoleutdanning, i Oslo var andelen nesten 50 % (norgeshelsa.no, 2007). I utvalget for effektstudien hadde 77 % utdanning på høyskole- eller universitetsnivå. Over 70 % av kvinnene

hadde drevet regelmessig trening siste året før svangerskapet. Rekruttering ble gjort ved at kvinner selv meldte interesse for å delta i prosjektet etter informasjon i primærhelsetjenesten.

Faktorer som utdanningsnivå, treningsaktivitet og initiativ til selv å oppsøke hjelp for aktuelle plager, kan indikere at materialet representerte en gruppe ressurssterke kvinner. Angående treningsaktivitet før svangerskapet, synes det å foreligge lite kunnskap om hvilken innvirkning dette har for bekkenplager etter fødsel. En studie har vist at prevalens for bekkenplager var like høy for toppidrettsutøvere som for andre kvinner både i og etter svangerskap (Bo & Backe-Hansen, 2006). Det er videre ingen holdepunkter for at trening i løpet av svangerskapet innvirker på slike plager (Ostgaard et al., 1994b). I tidligere studier er det vurdert som usikkert om kvinnenes alder har betydning, men det finnes sterk dokumentasjon for at tidligere bekkenplager er en risikofaktor for plager i svangerskapet (Wu et al., 2004). Noe høyere alder og andel førstegangs fødende vurderes likevel ikke å være faktorer som svekker generaliserbarheten av resultatene i studien i betydelig grad siden samtlige inkluderte hadde plager.

### 5.3 Mål for smerte og funksjon

VAS er betraktet som en endimensjonal målemetode for smerteintensitet, men er hyppig brukt i studier av mange ulike diagnosegrupper. Den er regnet for å være mer krevende i bruk enn numeriske smerteskalaer (Grotle et al., 2004). Gruppen som skulle måles i denne kohortestudien var relativt unge kvinner, de viste seg også å ha svært høyt utdanningsnivå. Bruk av VAS som måleskala burde derfor være ukomplisert å administrere for disse kvinnene.

En studie av Steingrimsdottir og medarbeidere (2004) viste at mer enn to måletidspunkt var nødvendig for pålitelige smertemålinger både ved måling av behandlingseffekt og ved kartlegging av helseplager over tid (Steingrimsdottir, Vollestad, Roe, & Knardahl, 2004). Effektstudien til Stuge og medarbeidere (2004) ble planlagt med smertemåling ved inklusjonstidspunkt og etter avsluttet behandling. Dette medførte to smertemålinger med flere måneders mellomrom. På den ene siden medførte det at kvinnene ved posttest hadde liten mulighet til å huske tidligere skåring og dermed mindre mulighet til å bevisst påvirke resultatet. Et bidrag til mer pålitelige skåringer var også at de fysiske omgivelsene og den muntlige informasjonen var lik i begge testsituasjonene. På den andre siden var VAS sannsynligvis en ukjent rapporteringsform for mange ved pretest, mens de ved posttest var kjent med prosedyren. Det er også vist at rapporteringen kan påvirkes av humør, smertenivå under utfylling, forventninger og redsel (Gifford et al., 1997; Steingrimsdottir et al.,



---

2004; Williamson et al., 2005). Slike faktorer kan være med å svekke påliteligheten av de aktuelle smertemålingene.

De inkluderte kvinnene ble spurt om å markere på VAS-skala ”nåværende smerte på det verste” både for morgen og kveld. Smerte om kvelden ble valgt for videre analyser. Klinisk erfaring indikerte at kvinner med bekkenplager etter fødsel ofte hadde økende smerter utover dagen. Det er også vist at denne pasientgruppen har økende smerter ved en rekke daglige aktiviteter (Mens et al., 1996). Smertenivå om kvelden ble derfor vurdert å være en relevant indikator på hvor store plager de opplevde og ble valgt i stedet for gjennomsnittlig smerte. Samtidig var spørsmålet om ”smerte på det verste” et forsøk på å redusere tolkningsmulighetene og dermed standardisere målingen. De kliniske testene ble dels utført på formiddag og dels på ettermiddag. Det kan ses som en svakhet i forhold til sammenheng at skår på smerteintensitet om kvelden analyseres opp mot svar på kliniske tester utført på dagtid. Samtidig finnes lite kunnskap om hvor sensitive provokasjonstester er for døgnvariasjoner i smerte.

”Nåværende smerte” var en formulering som ga rom for noe tolkning hos kvinnene som skulle svare. Noen vil kunne oppfatte dette som smertene i dag, noen ville oppfatte det som smertene nå for tiden eller de siste dagene. Denne skåringen på VAS ble gjort rett etter utfylling av smertetegning der spørsmålet var ”Hvor har du vondt nå for tiden?”. I denne sammenhengen kunne det være rimelig å anta at ”nåværende smerte” ble tolket som smerte nå for tiden. Flere artikler har diskutert smerte *nå* kontra gjennomsnittlig smerte over tid. Grotle og medarbeidere (2004) viste at VAS var egnet til å måle smerter hos pasienter med akutte rygg smerter, men at målemetoden hadde svakheter ved rapportering av kroniske smerter. Bolton og Wilkinson (1998) viste at skåring på gjennomsnittlig smerte hadde større evne til å avdekke klinisk relevante endringer enn smertenivå i øyeblikket. Det er også vist at VAS var mindre sensitiv for kliniske endringer når smerte i øyeblikket ble skåret sammenlignet med smerte over de siste 24 timer (Scrimshaw & Maher, 2001). Sett i lys av disse studiene vurderes spørsmålsformuleringen om smerte og konteksten den ble presentert i som akseptabel.

Det var i effektstudien ønskelig å bruke et mål for kvinnenes egen rapportering av funksjon, og det måtte derfor gjøres et valg ut fra tilgjengelig litteratur om måleredskaper sammen med en skjønnsmessig vurdering. I andre sammenlignbare studier er det benyttet ulike spørreskjemaer som QBPDS, DRI, RMDQ og ODI (Mens et al., 2002a; Laslett et al., 2003; Nilsson-Wikmar et al., 2005; O'Sullivan et al., 2006; Gutke et al., 2006). Felles for disse er at de er utviklet for registrering av ryggplager, og det er ikke kjent at det er gjennomført noen validering av spørreskjemaene for bruk ved bekkenplager (kap. 2.3.2). Det finnes altså ingen gullstandard for denne type målinger.

Forfatterne i de nevnte sammenlignbare studiene har likevel argumentert for at spørreskjemaene for ryggplager er relevante funksjonsmål ved bekkenplager. ODI, DRI og RMDQ er også sammenlignet med tanke på evne til å avdekke kliniske viktige endringer for en gruppe med akutte og kroniske lumbale smerter. Denne sammenligningen viste ingen forskjeller i funksjonsstatus målt med de ulike skjemaene (Grotle et al., 2004).

Det er utarbeidet ulike versjoner av ODI, og forfatterne av spørreskjemaet har anbefalt versjon 2.0 (Fairbank et al., 2000). I effektstudien ble ODI Chiropractic Version benyttet (Stuge et al., 2004). Valget ble gjort på grunnlag av at seksjonen om seksualliv var utelatt i denne versjonen og erstattet av spørsmål om endring av smertenivå. Kvinnene som skulle inkluderes i studien hadde født barn 6-16 uker før de skulle svare på spørreskjemaet. Mål for endring av smerter ble derfor vurdert som mer relevant enn seksualfunksjon på det aktuelle tidspunktet. Fairbank (2000) kritiserte denne versjonen av ODI, men seinere studier har vist at også Chiropractic Version er godt validert og inneholder en hensiktsmessig blanding av ulike funksjonsområder (Grotle et al., 2005).

Klinisk erfaring så vel som forskning viser at det ofte kan foreligge en kombinasjon av lumbosacrale plager og bekkenplager hos den aktuelle pasientgruppen (Wu et al., 2004). Symptomer kan være overlappende og til dels sammenfallende ved patologi og funksjonsforandringer i nedre lumbal og bekken, og det er i begge tilfelle ofte økende plager ved en rekke daglige aktiviteter (Mens et al., 1996; Wu et al., 2004). Det er på denne bakgrunn sannsynlig at ODI kan være et relevant funksjonsmål ved bekkenplager. Gutke og Østgaard benyttet dette som funksjonsmål for gravide og hevdet ODI var bedre egnet enn RMDQ på grunn av høyere sensitivitet ved større grad av funksjonssvikt (Gutke et al., 2006). De argumenterte også for at ODI var bedre egnet enn DRI med begrunnelse i at DRI inneholder spørsmål om lite egnede aktiviteter som tunge løft, tungt arbeid og løping. Videre analyseresultater i denne studien må likevel ses i lys av at spørreskjemaet ODI ikke er noen gullstandard for mål av funksjonsnivå ved bekkenplager.

## 5.4 Eksklusjonskriterier

Valg av eksklusjonskriterier er vesentlig for å få en mest mulig homogen gruppe inkludert i studier. I effektstudien ble det vektlagt å rekruttere kvinner som hadde debut av bekkenplager i forbindelse med siste svangerskap. De som hadde kontinuerlige plager fra tidligere svangerskap eller vedvarende ryggplager ble forsøkt ekskludert. Ved å ikke tillate inklusjon de første seks ukene etter fødsel ble det også forsøkt å unngå rekruttering av pasienter i en periode med stort potensial for spontan bedring. Tidligere arbeider har indikert at smertereduksjonen er særlig stor den første måneden etter fødsel

(Ostgaard & Andersson, 1992; Albert et al., 2001). Krav om språkferdigheter i norsk var et bidrag for å sikre god kommunikasjon og dermed mest mulig pålitelige svar i undersøkelsessituasjonen.

Både i klinisk arbeid og forskning er det en utfordring å differensiere mellom tilstander som primært har sin årsak i lumbalcolumna og de som primært er bekkenplager. Nerverotkompresjon er betraktet som et uttrykk for alvorligere grad av lumbal patologi, og nervestrekktester vil ofte gi utstrålende smerter ved slike tilstander (Kirkesola et al., 1998; Kerr et al., 2002; Bogduk, 2005). Denne type tester er benyttet som sentrale eksklusjonskriterier i mange vitenskapelige studier relatert til bekkenplager (Kristiansson et al., 1996a; Albert et al., 2000; Mens et al., 2002a; Laslett et al., 2005). Nervestrekktester alene er likevel ikke tilstrekkelig for å skille ut pasienter med lumbal patologi som primær årsak til plagene.

Nervestrekktester ble benyttet sammen med smertetegning som eksklusjons/ inklusjonskriterier i effektstudien. Lokalisering av smerte ved smertetegning eller intervju er også benyttet til dette formålet i en rekke andre studier (Kristiansson et al., 1996a; Ostgaard et al., 1997; Hansen et al., 1999; Mens et al., 2001; Laslett et al., 2003; Laslett, 2005; Mousavi et al., 2007). Etter utfyllt smertetegning ble kvinnene intervjuet og bedt om å vise avmerket smertelokalisering på egen kropp for å kontrollere korrelasjon mellom tegningen og aktuell smerte (Fortin et al., 1994; Albert et al., 2000).

Det er mange arbeider som beskriver at bekkenplager ikke har smerteutbredelse proksimalt for L5-S1 (Ostgaard et al., 1994a; Mens et al., 2002a; Vleeming et al., 2002; O'Sullivan et al., 2002; Hungerford et al., 2004). Det er samtidig kjent at lumbal og lumbosacral patologi ofte kan ofte ha smerteutbredelse til bekkenregionen og være årsak til somatisk refererte smerter til både sete og lyske (Bogduk & Twomey, 1987; Bogduk, 2005). Bruk av nervestrekktester og smertetegning kan derfor ikke betraktes som tilstrekkelig for å utelukke lumbale plager som medvirkende årsak til bekkensmertene. Det må på dette grunnlaget tas med i vurderingen at en del av de inkluderte pasientene kunne ha smerter med opprinnelse både fra lumbal og bekken. Samtidig det er vesentlig at plagene hadde debutert i løpet av eller kort tid etter svangerskapet.

14 kvinner ble ekskludert fra effektstudien, de fleste på grunn av ryggplager med indikasjon på radikulopati. Disse kvinnene hadde selv vurdert at de hadde bekkenplager. Dette kan være en indikasjon på at eksklusjonskriteriene fanget opp alvorligere ryggplager. Utfordringen med seleksjon av bekkenplager er felles med de fleste andre studier på dette området, og er så langt ikke løst tilfredsstillende. Dette har også betydning når gyldigheten av resultatene skal vurderes.

## 5.5 Kliniske tester

Det er i teorikapittelet i denne oppgaven vist til mange ulike arbeider for dokumentasjon av validitet og reliabilitet for de aktuelle kliniske testene. Disse viste varierende og til dels motstridende resultater. En svakhet ved valg av flere kliniske tester var at aktuell dokumentasjon var begrenset til undersøkelser av gravide. Dette var tilfelle for testene P4, provokasjonstest av symfyser og symfysermerter ved Trendelenburg's test (Ostgaard et al., 1994a; Albert et al., 2000). Ved denne gjennomgangen framkom også vansker med å sammenligne resultater mellom ulike studier. Beskrivelse av testutførelse er gjort med svært varierende nøyaktighet. Der beskrivelser forelå, viste disse både ulik utførelse av flere tester og ulike definisjoner av positiv testskår.

Eksempler på vesentlige ulikheter i testutførelse finnes i dokumentasjonen av ASLR, kompresjonstester og Patrick's test. For ASLR ble både pasientens utførelse og poengskala for skåring endret underveis i utviklingen av testen (Mens et al., 1999; Mens et al., 2001). Andre har endret kriteriene og definert enhver subjektiv eller objektiv vanskelighet med å løfte beinet som positiv ASLR uten å benytte den beskrevne poengskalaen (Rost et al., 2004). Kompresjonstester er utført med ulike utgangsstillinger og håndplassering (Kristiansson et al., 1996a; Laslett et al., 2005). Albert benyttet "separasjon" om testen sideliggende kompresjon, og hennes "kompresjonstest" var identisk beskrevet med det andre benevnte som distraksjon (Albert et al., 2000; Laslett et al., 2005). Patrick's test er også beskrevet ulikt både med hensyn til utførelse og kriterier for positiv skår. Positiv Patrick's test er registrert ved ulik lokalisering av smerteprovokasjon, ved nedsatt bevegelighet og ved kombinasjoner av disse (van Deursen et al., 1990; Strender et al., 1997; Albert et al., 2000).

I noen studier er all smerte i det aktuelle området registrert som positiv test, i andre arbeider er dette begrenset til konkordant smerte (Robinson et al., 2007). Flere forfattere har også vist til at manglende standardisering for anvendt kraft og tid ved provokasjonstester har betydning for reproducerbarheten (Laslett, 2005; Levin et al., 2005). Utførelsen av testene i den foreliggende kohortestudien er forsøkt beskrevet og standardisert. Det må likevel utvises forsiktighet ved sammenligning mellom resultater fra dette arbeidet og en del andre studier siden flere tester er beskrevet og utført på ulike måter.

I dokumentasjonen av de valgte kliniske testene er det i mange tilfeller lite nøyaktige beskrivelser av hvilke vevsstrukturer og funksjoner som påvirkes av testen. ASLR er beskrevet som en uspesifikk funksjonstest. Med utgangspunkt i anatomiske sammenhenger kan testen påvirkes av årsaksforhold både i iliopsoas, thoracolumbal og lumbosacral overgang, IS-ledd og bekkenorganenes forbindelse til muskel-skjelettsystemet. Testen kan dermed involvere både muskel-skjelettstrukturer og viscerale

---

strukturer i området Th11 til S4 (Gray, 1989). P4 er beskrevet som en uspesifikk test for differensiering mellom lumbale smerter og bekkensmerter. Samtidig vil utførelse av denne testen inkludere kompresjon og glidning i hoftelrådet, og foroverrotasjon av ilium vil kunne påvirke nedre lumbalcolumna gjennom stramming av iliolumbale ligamenter. LDL er beskrevet til å ha sammenhenger med flere fascie- og ligamentstrukturer både i thoracal, lumbal og bekken (Vleeming et al., 1996). Smerte i LDL kan dermed tenkes å ha sin årsak i omliggende strukturer. Både distraksjon og kompresjonstestene vil gi strekk på ligament- og kapselstrukturer samtidig med kompresjon av leddflater i IS-leddene, og kan dermed ikke differensiere mellom patologi i leddene og irritasjon av ligamentstrukturene. Patrick`s test involverer de samme strukturene i og rundt IS-leddene, og vil i tillegg ved rotasjon av ilium påvirke lumbosacral overgang. Menell`s test vil også ved sagittal bevegelse i IS-ledd samtidig påvirke lumbal-columna. Dersom det legges til grunn at differensiering mellom ulike vevsstrukturer er vesentlig for å skille lumbal og bekken som smertekilde, vil de aktuelle testene være lite presise.

Fem av de kliniske testene ble bruk som inklusjonskriterier sammen med smertetegning i effektstudien. P4 eller ASLR måtte være positiv sammen med en av symfysetestene eller LDL. Disse testene er også benyttet i ulike kombinasjoner i en rekke andre studier (Pool-Goudzwaard et al., 2005; Nilsson-Wikmar et al., 2005; O'Sullivan et al., 2006). Det er ikke kjent at studier i ettertid har vist til andre og mer pålitelige inklusjonskriterier for bekkenplager. De aktuelle inklusjonstestene er også anbefalt for diagnostikk av bekkenplager i retningslinjene fra European Cost Commission (Vleeming et al., 2004). Samtidig kan det på grunnlag av det anatomiske og biomekaniske perspektivet som er presentert over, argumenteres for at testene er mangelfulle som redskap for differensiering av ulike strukturer som kan gi plager i bekkenet og columna.

## 5.6 Gjennomføring av testprosedyre

Gjennomføring av en klinisk undersøkelse vil alltid inneholde elementer av subjektiv vurdering både fra pasient og undersøker. I testprosedyren ble det lagt vekt på å standardisere både muntlig informasjon om og praktisk utførelse av testene for å styrke påliteligheten av registreringene. For den enkelte kvinnes vedkommende ville testsvar kunne påvirkes av ulike faktorer som blant annet forventninger, engstelse eller tidligere smerteerfaringer (Gifford et al., 1997; Brodal, 2005). Grundig og mest mulig entydig informasjon i forkant av undersøkelsen ble forsøkt gjennomført for å skape en trygg testsituasjon. Mange av testene var provokasjonstester der skåringen ble gjort på grunnlag av et ja- eller nei-svar på spørsmål om smerte. I en del tilfeller var pasientene i tvil om de skulle angi det

de kjente som smerter. Også her var det subjektive vurderinger som ga rom for redusert pålitelighet av testene. Pasientene ble i tvilstilfeller ved bilaterale tester bedt om å sammenligne og vurdere i forhold til samme test på motsatt side. Det finnes i andre studier eksempler på at smerteprovokasjonstester er skåret med gradering av smerte i flere kategorier (Vleeming et al., 2002). Dette kunne vært en mulighet for å få fram mer nyanserte registreringer for provokasjonstestene, men de aktuelle graderingene er ikke reliabilitetstestet.

Samme undersøker utførte testing på samtlige pasienter ved både pre- og posttest. Dette ble gjort for å sikre mest mulig lik gjennomføring av tester og registreringer gjennom prosjektet. Andre arbeider har vist at intratester-reliabilitet er høyere enn reliabilitet mellom ulike undersøkere (Van der Wurff et al., 2000a). Samtidig representerte dette en fare for at eventuelle feil eller svakheter i utførelsen ble gjennomgående i hele datamaterialet. Undersøker stilte med ulik klinisk erfaring relatert til de ulike testene. Noen av testene var benyttet regelmessig i klinisk arbeid over mange år, andre tester var ikke benyttet tidligere (ASLR og palpasjon av LDL). For ”nye” tester som aktivt skulle utføres av undersøker, kunne det tenkes at økende erfaring underveis i prosjektet kunne gi noe endring i utførelsen og dermed påvirke resultatet av testene. Ved palpasjonstester for smerteprovokasjon (LDL, sacrotuberale ligament og iliopsoas) er det også usikkerhet forbundet med presisjonsnivået for palpasjon og mulighet for at dette kunne påvirke resultatet.

En betydelig utfordring ved studier av kliniske tester er påvirkningen selve testen har på den aktuelle tilstanden. Det er lite sannsynlig at en rekke tester kan utføres på et område med smerter og funksjonsforandringer uten at dette faktisk endrer tilstanden i det aktuelle området. Det er derfor mulig at samme testprosedyre utført i en annen rekkefølge kunne gitt andre resultater.

## 5.7 Statistiske analyser

De kliniske testene var registrert som 14 ulike variabler. Med unntak av ASLR resulterte alle disse i kategoriske data. Målene for smerte og funksjon representerte kontinuerlige data. Dette utgangspunktet la føringer for hvilke analysemåter som kunne være aktuelle for datamaterialet.

I innledende tabellanalyser for sammenheng mellom testene, var det en svakhet at kravet om forventet antall i en vesentlig andel av cellene ikke var oppfylt. Det ble korrigert for dette ved å bruke Fischer`s eksakte test, men resultatene må likevel tolkes med varsomhet.

Trinnvis multippel regresjon ble benyttet for å analysere testenes selvstendige forklaringsverdi for smerte og funksjon. Antall uavhengige variabler ble i denne analysen redusert fra 14 til 10, men var fortsatt høyt relatert til antallet inkluderte pasienter. Analysen kan kritiseres ut fra at noen kliniske tester er fjernet eller slått sammen. Denne reduksjonen er begrunnet i klinisk vurdering sammenholdt med skåringsresultatene både ved pre- og posttest.

Med utgangspunkt i forskningsspørsmålene som ble presentert innledningsvis, ble det valgt å ikke analysere skår på bilaterale kliniske tester separat for høyre og venstre side. Når det gjelder sammenheng mellom testene, kan det likevel hevdes at en slik analyse kunne være av interesse. Det er tidligere vist at asymmetrisk bevegelse i IS-leddene har prediktiv verdi for bekkenplager etter fødsel (Damen et al., 2002). I praksis er også sidelokalisering av aktuelle positive testskår vesentlig for den kliniske resonneringen som gjøres ved slike undersøkelser.

## 6. Diskusjon av resultater

Resultatene i denne studien er presentert for å beskrive en gruppe kvinner med bekkenplager og for å avdekke eventuelle sammenhenger og endringer relatert til de aktuelle kliniske testene, smerteintensitet og funksjon. Diskusjonen av resultater konsentreres om noen framtrepende trekk i beskrivelsen av datamaterialet samt forskningsspørsmålene som er presentert i kapittel 2.5.

### 6.1 Beskrivelse av bekkenplagene

Lokalisering av smerter er vurdert å ha betydning for alvorlighetsgrad av bekkenplager (Albert et al., 2001). Fordeling av smertelokalisasjon i materialet for kohortestudien var relativt lik kartlegginger som tidligere er gjort blant gravide. 12 % av de inkluderte kvinnene markerte kun symfysesmerter, 25 % hadde smerter fra alle tre bekkenledd. En nederlandsk studie av gravide viste at 18 % hadde kun symfysesmerter og vel 30 % hadde smerter fra alle tre bekkenledd (Rost et al., 2004). En annen studie viste også at 12 % hadde kun fremre smerter (Mogren & Pohjanen, 2005). Smerter lokalisert kun til symfyseområdet er vist å ha god prognose (Albert et al., 2001). Blant de inkluderte i den foreliggende studien var det altså en liten andel (12 %) som tilhørte denne gruppen.

Det er tidligere vist at rundt 20 % av gravide har hatt kroniske ryggplager før svangerskapet (Ostgaard et al., 1992; Ostgaard et al., 1997). Kvinner med kroniske ryggplager før aktuelle graviditet ble forsøkt ekskludert fra materialet i effektstudien. Halvparten av de inkluderte kvinnene rapporterte likevel episoder med vondt i korsryggen før siste svangerskap. Dette samsvarer med hyppigheten av lumbale plager i den norske befolkningen forøvrig (Natvig, Nessioy, Bruusgaard, & Rutle, 1994). Tidligere lumbale smerter er vurdert som risikofaktor for lumbal- og bekkenplager i svangerskap (Wu et al., 2004; Albert, Godskesen, Korsholm, & Westergaard, 2006). Blant friske gravide er det vist at kun 20 % hadde hatt rygg smerter før svangerskapet (Kristiansson, Svardsudd, & von, 1996b), mens blant gravide med bekkenplager hadde 60 % hatt rygg smerter tidligere (Rost et al., 2004).

Materialet for denne kohortestudien viste altså stor grad av samsvar med tidligere publiserte data når det gjelder smerteutbredelse og tidligere ryggplager hos kvinner med bekkenplager etter fødsel.



### 6.1.1 Smerte og funksjon

Spredningen i skår på VAS og ODI var stor både ved pretest og posttest. Dette viste stor variasjon i smerteintensitet og funksjon innenfor en gruppe som var forsøkt gjort så homogen som mulig. Resultatene av smerte- og funksjonsmålingene ved pretest ga likevel klare indikasjoner på at inklusjonskriteriene medvirket til rekruttering av kvinner med betydelige plager relatert til bekkenet. Gjennomsnittlig skår på VAS ved pretest var 61 mm, dette tilsvarer sterke smerter (Collins et al., 1997). Gravide med bekkenplager rapporterte i en studie fra 1994 gjennomsnittlig smerteskår på 44 mm (Ostgaard et al., 1994b). I en studie med alle gravide i et område av Göteborg inkludert, viste skår på VAS i underkant av 40 mm tre måneder etter fødsel (Ostgaard et al., 1997). En annen studie viste at gravide med kun bekkenplager skåret 26 mm på VAS, ved kombinasjon av lumbal- og bekkenplager var skår 36 mm (Gutke et al., 2006). Resultatene viste dermed høyere smerteintensitet for de inkluderte i denne kohortestudien enn i lignende studier av gravide med bekkenplager. De skåret også høyere enn et gjennomsnittsmateriale (med og uten plager) målt på tilsvarende tidspunkt etter fødsel (Ostgaard et al., 1997). Dette må likevel ses i lys av at gruppen har svært stor spredning i smerterapportering (SD 22, CI 56,66).

Smertereduksjonen fra pretest til posttest målt ved VAS var gjennomsnittlig 26 mm. Dette var noe større reduksjon enn det som er angitt å være variabilitet ved bruk av VAS som smertemål (Rosier et al., 2002). Endringen var statistisk signifikant, men lå i grenseland for det som er antydning å være klinisk signifikant smertereduksjon (Kropmans et al., 1999; Williamson et al., 2005).

De inkluderte kvinnene hadde også større vansker med funksjon enn ved tilsvarende målinger i sammenlignbare studier. Gjennomsnittlig skår på ODI ved pretest var 40 poeng, dette tilsvarte grensen mellom moderate og betydelige funksjonsproblemer. Over halvparten av kvinnene hadde i følge denne rapporteringen betydelige funksjonsproblemer. Men også for ODI var det stor spredning i rapporteringen (SD 14, CI 37,44). Til sammenligning skåret en gruppe gravide svenske kvinner i gjennomsnitt 14 poeng på ODI ved bekkenplager og 18 poeng ved kombinerte lumbal- og bekkenplager, de symptomfrie skåret 4 poeng (Gutke et al., 2006).

Reduksjon i skår på ODI var gjennomsnittlig 26 poeng fra pretest til posttest. Klinisk signifikant endring er hevdet å være mellom 4 og 15 poeng, flere studier viser til rundt 11 poeng som en minimumsgrense (Fairbank et al., 2000; Grotle et al., 2003; Holm et al., 2003). Reduksjonen i skår på ODI må betraktes som en betydelig bedring i funksjon slik kvinnene selv rapporterte dette.

Kohortestudien viste samsvar mellom høy smerteintensitet og betydelige funksjonsproblemer. Når det gjelder forholdet mellom disse variablene, er det også tidligere vist høy korrelasjon mellom smerteskår og funksjon hos friske gravide (Kristiansson et al., 1996b).

Resultatene i denne kohortestudien kan tolkes slik at endringen i funksjon er mer overbevisende enn reduksjon i smerteintensitet. Smerte og funksjon er i praksis faktorer som gjensidig påvirker hverandre. Lav smerteintensitet kan oppnås hos kvinner som velger å være lite aktive for å unngå smerter, og dette vil da ikke entydig være uttrykk for en bedret situasjon. Men når redusert smerteintensitet her følges av betydelig redusert ODI, uttrykker dette at funksjonsnivået er klart øket samtidig som smertene er redusert. Det kan ut fra et slikt resonnement tenkes at ODI er et mer vesentlig mål enn smerteintensitet når det gjelder endring av tilstanden for denne pasientgruppen. Det kan også tolkes slik at variablene må ses i sammenheng dersom hensikten er å presentere pålitelige mål for graden av plager. Et slikt resonnement samsvarer med en studie av von Korff og medarbeidere (2000), som viste til at mål for smerteintensitet sammen med smertens påvirkning på aktiviteter kunne gi en mer valid og reliabel smertere registrering.

### **6.1.2 Kliniske tester**

Kvinnene som ble inkludert i denne studien er forsøkt beskrevet ut fra hvordan de skåret på kliniske tester. I den videre diskusjonen ses denne skåringsprofilen i lys av testresultater i andre relevante studier og erfaringer fra klinisk praksis.

ASLR viste ved pretest gjennomsnittlig skår på 4 poeng, men hadde spredning over tilnærmet hele skalaen (0-9). Dette er omtrent sammenfallende med resultater fra en sammenlignbar pasientgruppe der gjennomsnittlig skår var 4.6 med spredning over hele måleskalaen (Mens et al., 2001) og en gruppe gravide med gjennomsnittlig skår på 3.9 (SD=2) (de Groot M., Pool-Goudzwaard, Spoor, & Snijders, 2006). Den høye andelen positive skår ved pretest (ca. 90 %) antas å ha sammenheng med at ASLR ble benyttet som inklusjonstest, og det var krav om at enten ASLR eller P4 måtte være positiv. 88 % av kvinnene skåret positivt både på ASLR (positiv test ved sumskår lik 1 eller høyere) og P4. En tidligere studie av en tilsvarende pasientgruppe med plager etter fødsel viste at i underkant av 60 % hadde positiv test på både ASLR og P4 når testene ikke var benyttet som inklusjonskriterier (Vleeming et al., 1998). Blant gravide er det vist at to tredjedeler (66 %) skåret positivt på ASLR, men da med en alternativ definisjon av positiv testskår (Rost et al., 2004).

P4 viste størst andel bilateralt positive skår (75 %) ved pretest og lavest andel bilateralt negative. Ut fra klinisk erfaring kunne det være rimelig å vente hyppig smertesvar ved en uspesifikk test av bakre

---

strukturer rundt IS-leddene hos denne pasientgruppen. Det er likevel verd å merke seg at tre fjerdedeler av kvinnene hadde smerter på begge sider ved denne testen.

LDL var definert som en mer lokal provokasjonstest, men den hadde et tilnærmet likt skåringsmønster som P4 med 73 % bilateralt positive og omtrent 10 % bilateralt negative skår. Begge testene ble brukt som inklusjonstester. På tross av at test av LDL ble utført ved direkte trykk mot ligamentet, kan det være usikkerhet rundt årsaken til smertesvar. Det er vist at laterale greiner av S1-S3 har forløp under ligamentet, og at kompresjon av disse kan være årsak til smerteprovokasjon ved test (Willard, Carreiro, & Manko, 1998). Det er også vist signifikant assosiasjon mellom smerter ved palpasjon av LDL og lumbale smerter (Carreiro et al., 1998). Lumbale plager kunne ut fra dette være en mulig underliggende årsak til positiv test av LDL.

Trendelenburgs test ble benyttet i testprosedyren da det var aktuelt å registrere eventuelle symfysesmerter ved stående på ett bein. Klinisk erfaring kunne tilsa at positiv test ville opptre relativt sjelden, men andre studier har vist gangvansker og en sviktfølelse i underekstremitetene hos kvinner etter fødsel (Sturesson et al., 1997; Rost et al., 2004). I dette materialet var det kun én pasient som skåret positivt på Trendelenburg's test. Dette ga en klar indikasjon på at test av muskelstyrke i gluteus medius ikke fanget opp aktuelle bekkenplager. I ettertid synes det relevant å kunne ha registrert eventuelle smerter lokalisert til IS-leddene ved denne testen. Test ved stående på ett bein ville kunne gi informasjon om plager ved overføring av krefter fra columna til underekstremiteter i vekt bærende stilling. Slik informasjon ble ikke fanget opp ved de øvrige testene. Eventuelle resultater ville være relevant å sammenholde med ASLR som er hevdet å gi lignende informasjon ved ikke vekt bærende stilling (Mens et al., 1999).

Menell's test hadde også klar overvekt av negative testsvar. Dersom testen skulle forventes å avdekke plager fra området i og rundt iliosacralleddene, kunne testresultatene indikere at den ikke er godt egnet da over 90 % av kvinnene samtidig rapporterte bakre bekkenplager. Albert og medarbeidere (2000) har anbefalt Menell's test for subgruppering av bekkenplager hos gravide. De mente den sagitale bevegelsen som påføres bekkenet ved testen er sensitiv for smerter i IS-leddene. Positiv testskår kan synes å være definert mer avgrenset i Alberts studie. Det kunne også tenkes at Menell's test var bedre egnet under graviditet enn etter fødsel. Kohortestudien viste ikke resultater for Menell's test som er sammenfallende med Alberts konklusjon.

Patrick's test var positiv på en eller begge sider hos vel tre fjerdedeler av kvinnene i den foreliggende studien. Dette var omtrent samsvarende med resultater for samme test utført på gravide (72 %), men her var ikke kriterier for positiv testskår beskrevet (Rost et al., 2004). I effektstudien til Stuge og

medarbeidere (2004), og videre også i denne kohortestudien, ble Patrick's test definert positiv både ved smerteprovokasjon og nedsatt bevegelighet. Valget ble begrunnet med klinisk relevans, da både smerteprovokasjon av fremre og bakre strukturer i bekkenet og nedsatt bevegelighet i området kunne tenkes å være uttrykk for bekkenplager.

Kompresjonstest ble utført med kvinnene i både sideliggende og ryggliggende stilling, resultatene viste noen flere positive skår i ryggliggende. De fleste i denne pasientgruppen hadde plager relatert til bakre strukturer i bekkenet. Det kunne tenkes at sideliggende for disse var en mindre provoserende utgangsstilling og dermed kunne eliminere noen falske positive testsvar. De relativt like resultatene kan tolkes som en indikasjon på at det er unødvendig å utføre testen i både ryggliggende og sideliggende stilling.

Symfysesmerter ved provokasjonstest skåret positivt hos mange flere kvinner enn de som rapporterte symfysesmerter ved stående på ett bein. Provokasjonstesten var samtidig positiv hos alle som hadde symfysesmerter i stående. Test i stående ga likevel informasjon om plager i vekt bærende stilling som kunne være verdifull i en klinisk resonneringsprosess.

### **6.1.3 Endringer i testskår**

De inkluderte kvinnene rapporterte betydelig redusert smerteintensitet og bedret funksjon fra pretest til posttest. Det var derfor av interesse å se om det kliniske bildet endret seg i samsvar med dette. Analysene viste at fleste kliniske testene hadde statistisk signifikant reduksjon i antall positive skår i denne perioden. Unntaket var Trendelenburgs test, palpasjon av sacrotuberale ligament og måling av diastase. Trendelenburgs test hadde ikke potensial for endring da alle unntatt en kvinne skåret negativt ved pretest. Ved måling av diastase i rectus abdominis skåret også over tre fjerdedeler av de inkluderte negativt ved pretest, og potensialet for endring var lite. Samtidig er det også tidligere dokumentert at testen ikke viste diskriminerende evne mellom gravide kvinner med og uten plager (Wormslev et al., 1994).

Palpasjon av sacrotuberale ligament skilte seg ut fra de andre provokasjonstestene ved å ikke vise statistisk signifikant reduksjon av positive skår fra pretest til posttest. Testen var fortsatt smerteprovoserende hos en stor andel av kvinnene som selv rapporterte bedret funksjon og redusert smerteintensitet. Positiv test var ikke gradert relatert til smerte, og kunne derfor bli angitt som positiv selv ved moderat palpasjonsømheter. Men resultatet var ikke tilsvarende for LDL på tross av at utfordringene var de samme i skåringsprosedyren. Det kunne tenkes at test av sacrotuberale ligament er bedre egnet enn de øvrige til å avdekke mer beskjedne plager, men graden av nøyaktighet ved

utførelse av palpasjonstester kan også ha betydning for resultatet. Det er ikke kjent hvordan et materiale av friske kvinner ville skåre på denne testen.

Ved ASLR skåret kvinnene gjennomsnittlig 2 poeng lavere ved posttest sammenlignet med pretest. Reduksjonen var da i mange tilfelle kun ett poeng på hver side. Det er usikkert i hvilken grad en slik reduksjon i skår gir uttrykk for en meningsfull og klinisk relevant endring i funksjon.

Blant bilaterale tester var det P4 som viste størst reduksjon i positive testskår ( $p < 0.001$ ). Symfyseprovokasjon viste størst endring av sideuavhengige tester ( $p < 0.001$ ). Korrelasjonen med smerte og funksjon ved posttest var svak til moderat på lik linje med de fleste øvrige testene. Det er likevel av interesse at tester som provoserte symfyssen og bakre strukturer rundt IS-leddene var de som viste størst endring hos en gruppe som selv rapporterte reduserte bekkenplager.

## 6.2 Sammenhenger

Resultatene viste statistisk signifikant korrelasjon mellom mange av de kliniske testene, samtidig var disse sammenhengene gjennomgående moderate eller svake. Det var vanskelig å avdekke noe klart mønster for sammenhenger relatert til de aktuelle strukturene som ble testet. Dette kan tolkes som et uttrykk for at testene var for lite presist definert eller utført. Det kunne også være at provokasjon av de samme strukturene i ulike retninger må forventes å gi forskjellig testsvar. I ettertid kan det synes interessant å gjennomføre analyser av testene basert på kunnskap om biomekaniske sammenhenger og ”funksjonskjeder” som er kjent fra anatomiske studier og kliniske resonnement. Leddflaters form, ligamentstrukturer og muskulatur danner funksjonelle kjeder for stabilitet og bevegelse. Studier av funksjon har for eksempel vist at LDL ikke kan ses isolert fra sacrotuberale ligament, glutealmuskulatur, multifider, erector spinae og biceps femoris (Vleeming, Snijders, Stoeckart, & Mens, 1997; Vleeming et al., 2002). Strukturene har gjensidig påvirkning på hverandre (Pool-Goudzwaard, Vleeming, Stoeckart, Snijders, & Mens, 1998; van Wingerden, Vleeming, Buyruk, & Raissadat, 2004). Sett i lys av dette kan det være av interesse at korrelasjonsanalysene viste statistisk signifikant sammenheng mellom LDL og sacrotuberale ligament. Strukturene rundt iliosacralleddene ble også uspesifikt provosert ved P4-testen, og denne viste statistisk signifikant sammenheng med både ASLR, LDL, distraksjon, Menell's og Patrick's test.

ASLR, P4 og LDL var alle brukt som inklusjonstester. Sammenheng mellom disse ble analysert ved posttest. ASLR og LDL viste moderat korrelasjon ( $r = 0.43$ ), P4 og LDL viste også statistisk signifikant korrelasjon ( $\chi^2 = 12.5$ ). Vleeming og medarbeidere (2002) har vist til lignende resultater.

P4 og LDL analysert separat for høyre og venstre side viste i deres studie signifikant lav korrelasjon ( $r=0.35$  og  $0.41$ ), det samme var tilfelle for P4 og LDL ( $r = 0.33$  og  $0.41$ , LDL analysert med smertegradering). De viste også til at utgangsstilling for testene har betydning for testresultatet. ASLR og P4 utføres i ryggliggende stilling med fleksjon av hoften, dette medfører konternutasjon som gir en annen belastning av LDL enn ved test i framliggende stilling. Korrelasjon mellom ASLR og P4 i den foreliggende kohortestudien var  $0.46$ . En tidligere studie har vist korrelasjon på  $0.27$  mellom disse testene, og den lave korrelasjonen var begrunnet med at testene sannsynligvis målte ulike aspekter av bekkenplager (Mens et al., 2001).

Resultatene viste statistisk signifikant korrelasjon mellom ryggliggende og sideliggende utgangsstilling for kompresjonstesten. Dette styrker oppfatningen av at de ulike stillingene ved utførelse av testen gir lik informasjon. Samtidig var det manglende sammenheng mellom ulike tester som var antatt å provosere de samme anatomiske strukturene. I dette materialet viste analysene ingen sammenheng mellom resultatene for P4 og kompresjonstestene. Alle disse testene var beskrevet til å provosere bakre strukturer rundt IS-leddene, men med ulik kraftretning. Samtidig viste innledende tabellanalyser at samtlige kvinner som skåret positivt på kompresjonstest i sideliggende også skåret positivt på P4 på en eller begge sider både ved pretest og posttest. Tilnærmet samme resultat viste seg også for ryggliggende kompresjon. Det kunne stilles spørsmål om P4 er bedre egnet enn kompresjonstestene for å undersøke ligamentstrukturene rundt IS-leddene. Det kan også betraktes som en svakhet for alle disse provokasjonstestene at de med stor sannsynlighet samtidig komprimerer deler av leddflatene og strekker kapsel- og ligamentstrukturer i andre deler av leddforbindelsen (Laslett et al., 1994). Dersom det legges til grunn, vil disse provokasjonstestene ikke bidra til differensiering mellom ulike vevsstrukturer som årsak til smerte.

Forholdet mellom de kliniske testene og egenrapportert smerte og funksjon endret seg fra pretest til posttest. Flere momenter kunne være av betydning for å forklare denne endringen. Sammensetningen av pasientgruppen må antas å være vesentlig. Ved pretest var det forsøkt rekruttert en mest mulig homogen gruppe av kvinner med bekkenplager. Ved posttest hadde den samme gruppen endret status der en del fortsatt var plaget i varierende grad, og en del var helt eller tilnærmet symptomfri. Det ble dermed større spredning i materialet ved posttest og større muligheter for å påvise sammenhenger.

Et annet moment er at skår på smerteprovokasjonstester ikke var egnet for gradering av plagene på samme måte som måleskalaene for VAS og ODI. Provokasjonstestene var negative kun når pasienten selv registrerte smertefrihet ved testen. Resultatet ville da bli positiv skår enten det var beskjeden eller betydelig grad av smerteprovokasjon. I den praktiske gjennomføringen var det også behov for

ulik grad av provokasjon før smerte ble utløst ved test, og denne variasjonen ble ikke registrert. Provokasjonstestene fanget derfor ikke opp intensiteten i smerten på samme måte som VAS.

ASLR viste ved pretest moderat og ved posttest god korrelasjon med VAS og ODI. Korrelasjon med smerte og funksjon syntes altså å være noe sterkere for denne testen enn for de øvrige kliniske testene. En mulig medvirkende årsak til dette kunne være ulike skåringsskalaer for de kliniske testene. Innledende analyser med ulike cut-off-verdier av ASLR viste tilnærmet like resultater, men dersom dikotomisering ble gjort mellom 0 og 1 (cut-off-verdien med best balanse mellom sensitivitet og spesifisitet), var korrelasjonen svakere og ved pretest da ikke statistisk signifikant.

ASLR er beskrevet som en funksjonstest som avdekker evnen til overføring av krefter fra overkroppen til underekstremitetene (Mens et al., 1999). Mens og medarbeidere (2002a) konkluderte med at ASLR var et valid mål for alvorlighetsgrad av bekkenplager. Deres studie viste høy korrelasjon ( $r = 0.70$ ) mellom ASLR og funksjon målt ved QBPDS hos kvinner etter fødsel. Resultatene fra den foreliggende studien er svakere, men ASLR viste moderat til god korrelasjon med både VAS og ODI og kunne slik til en viss grad understøtte deres konklusjon.

En studie av gravide med bekkenplager viste ikke samsvar mellom alvorlighetsgrad av bekkenplagene og vanlig brukte kliniske tester (Rost et al., 2004). Det var utarbeidet et eget spørreskjema for å måle alvorlighetsgrad. Dette ble målt ved varighet av smertefrihet ved daglige aktiviteter, og blant testene var ASLR og Patrick's test. Både målemetode og resultater er forskjellige fra den foreliggende studien der ASLR ved posttest viste god og Patrick's test moderat korrelasjon med både funksjon og smerte.

De fleste kliniske testene benyttet ved undersøkelse av bekken er primært utviklet for å verifisere plager i iliosacralledd og symfyse, og har dermed ikke til hensikt å avdekke smerteintensitet og funksjonsnivå. VAS og ODI kan ikke betraktes som gullstandard ved validering av kliniske tester, men er i denne studien benyttet som subjektive uttrykk for kvinnenens tilstand. Svak til moderat korrelasjon med de kliniske testene indikerer at testene kan ha en viss grad av validitet for smerteintensitet og funksjon knyttet til bekkenplager.

## 6.3 Kliniske testers bidrag til å forklare smerte og funksjon

Materialet i studien representerte etter behandling en gruppe sammensatt av kvinner både med og uten plager. De som da rapporterte lav smerteintensitet og god funksjon, skåret negativt på de aller

fleste kliniske testene. Resultatene viste også samsvar mellom økende antall positive tester og økende smerteintensitet og funksjonsnedsettelse, selv om variasjonen var stor for hvert enkelt individ. Dette samsvaret gir en styrket indikasjon på at bekkenplagene hos de inkluderte var forårsaket funksjonsforandringer i ulike vevsstrukturer med anatomisk tilknytning til bekkenet. Dersom årsak til plagene primært var av annen opprinnelse, ville det være mindre sannsynlig å finne slikt samsvar mellom kliniske tester for bekken, smerteintensitet og funksjonsnivå.

Ved analyser av antall positive kliniske tester relatert til VAS og ODI var resultatene ulike for pretest og posttest. Kvinnene rapporterte lavere smerteintensitet og bedre funksjon ved et gitt antall positive tester etter behandling sammenlignet med pretest. Korrelasjon mellom økende antall positive kliniske tester og økende smerteintensitet og funksjonsnedsettelse var statistisk signifikant ved posttest. Det aktuelle samsvaret var sammenfallende med resultater fra noen tidligere studier gjennomført for gravide. Kristiansson og medarbeidere (1996a) viste signifikant korrelasjon mellom smerteskår og de kliniske testene som ble benyttet, med økende tendens til positive tester ved økende smerte.

Wormslev og medarbeidere (1994) fant også signifikant korrelasjon mellom samlet antall positive tester og smerteintensitet og funksjonsnivå (målt ved 11 punkts boksskala). I begge disse studiene var gravide både med og uten plager inkludert. Et tilsvarende materiale ble undersøkt av Röst og medarbeidere (2004), men resultatene av deres studie viste at alvorlighetsgrad av bekkenplagene ikke var relatert til skår på kliniske tester.

I den undersøkte gruppen var det også slik at plagene var omtrent like store enten de hadde 7 eller flere positive kliniske tester. En slik tendens er ikke rapportert i sammenlignbare studier (Wormslev et al., 1994; Kristiansson et al., 1996a). Ved planlegging av effektstudien ble det valgt å benytte et relativt stort antall kliniske tester for kartlegging av bekkenplagene. Dette var i tråd med vanlig klinisk praksis, men det ville samtidig være en fordel både for pasient og undersøker dersom man i utredning kunne bruke et lavere antall tester som var mer pålitelige. Det ville i så fall føre til en mindre belastende prosedyre for pasienten, det ville også være mindre tidkrevende.

Denne kohortestudien fokuserte på om de kliniske testene kunne gi informasjon om smerteintensitet og funksjonsnivå, og viste at ASLR, LDL og Patrick's test samlet kunne forklare rundt 40 % av variansen i VAS og ODI. Usikkerheten i assosiasjonen var stor, men av de undersøkte testene var det disse tre som samlet kunne gi informasjon om kvinnenes smerteintensitet og funksjonsnivå. For ASLR er dette et betydelig svakere resultat enn det som er publisert i en tidligere studie. Mens og medarbeidere (2002b) viste at 49 % av variansen i QBPDS kunne forklares av ASLR alene, dette relateres da til et spørreskjema for funksjon som har mange fellestrekk med ODI. Test av LDL eller nærliggende strukturer er tidligere vist å ha høy prediktiv verdi for bekkenplager (Kristiansson et al.,



---

1996a), men det er ikke kjent at LDL og Patrick`s test tidligere er undersøkt for sammenheng med rapportering av smerteintensitet og funksjon.

## 6.4 Videre utfordringer ved diagnostikk av bekkenplager

En klinisk undersøkelse for diagnostikk av bekkenplager har som mål å avdekke patologi og funksjonsforandringer, og dermed være retningsgivende for behandlingstiltak. Smertelokalisering og ulike funksjonsmål er tidligere benyttet som sentrale kriterier for å definere plager ved validitetsstudier av kliniske tester. Det synes å være et mønster at gjentatt bruk av enkelte spørreskjema og kliniske tester i studier på dette feltet fører til aksept for å benytte disse som måleredskap og inklusjonskriterier. Dette gjøres uten ytterligere spørsmål om de er godt validert for den aktuelle pasientgruppen. På den bakgrunn gjenstår det fortsatt betydelige utfordringer med å etablere tilstrekkelig gode standarder for evaluering av kliniske tester.

Mange studier har benyttet en kombinasjon av anamnese og smerteutbredelse for å kategorisere plager. I klinikerens hverdag framtrer anamnesen ofte som helt sentral for å fange opp aktuelle symptomer, endringer og årsakssammenhenger. En bedret systematisering av informasjonen i en anamnese ville være vesentlig også i vitenskapelige studier. Det er mulig å tenke seg at kategorisering av informasjon i anamnesen ved hjelp av kvalitative forskningsmetoder kunne bidra til en bedret differensiering av diagnosegrupper som skal studeres.

Studier som er presentert i teorikapittelet i denne oppgaven viser varierende presisjonsnivå, og i noen tilfeller mangelfulle beskrivelser av hvordan kliniske tester er utført. Det er også vist til behov for bedret standardisering av trykk-kraft og retning for testutførelse. Manglende reliabilitet i en del studier er antatt å ha sammenheng med at behandlere med ulik utdanning og erfaringsbakgrunn utfører de samme testene på ulike måter. Det er sannsynlig at øket kvalitet på praksisundervisning og ferdighetstrening for klinikere ville ha betydning for resultat i validitets- og reliabilitetsstudier.

Definisjonen av bekkenplager fra European Cost Commission har lagt vekt på smerteutbredelse, funksjonskapasitet og differensiering fra lumbale plager (Vleeming et al., 2004). Denne definisjonen har også til hensikt å avgrense plagene fra patologiske tilstander innen gynekologi og urologi. Bekkenplager etter fødsel befinner seg i skjæringspunktet mellom mange ulike profesjoner som kan representere aktuell kompetanse og vurderer slike plager ut fra ulike faglige ståsted. Avgrensningen kan være en svakhet med tanke på å få gjennomført en samlet vurdering av de ulike anatomiske strukturene som kan relateres til bekkensmerter.

Det er tidligere vist til at smerteutbredelse i bekkenet kan ha sitt utspring i området fra Th 11 til S4 samt fra bekkenbunn og urogenetalia (kap. 2.2.3). Ved fødsel utsettes bindevevsstrukturene i bekkenet for betydelig mekanisk belastning som i mange tilfeller fører til vevsskade. Studier har vist at over halvparten av kvinner med lumbal- og bekkenplager etter fødsel har dysfunksjon i form av øket aktivitet i bekkenbunnsmuskulatur, og at plagene har sammenheng med forstyrrelser i blærekontroll (Pool-Goudzwaard et al., 2005). Videre er det funnet sannsynlig at bekkenplager kan ha sammenheng med skader på bekkenets ligamenter og gi vedvarende nedsatt stabilitet og sensitivisering med potensial for kroniske smerter (O'Sullivan et al., 2006). Nedsatt passiv fleksjon og innadrotasjon av hofter er også vist å ha sammenheng med alvorlighetsgrad av bekkenplager (Rost et al., 2004). De fleste kliniske testene som er hevdet å være relevante ved bekkenplager, har til hensikt å undersøke iliosacralledd og symfyse samt noen aktuelle ligamenter. Det synes rimelig at både bekkenets indre organer, iliolumbale ligamenter og ligamenter og bindevevsstrukturene i bekkenbunnen også kan være relevante som årsak til smerter i bekkenet (Pool-Goudzwaard et al., 2005). For å bedre differensiering av bekkenplager kan det være av betydning å fokusere mer på sammenheng mellom tilstand i disse ulike anatomiske strukturene og aktuell smerteutbredelse.

Med denne studien er det forsøkt å bidra til en utvidet beskrivelse av slike plager og belyse mulige sammenhenger mellom de kliniske testene og kvinnenens egen rapportering av smerteintensitet og funksjon. Presentasjon av aktuelt teorigrunnlag har også synliggjort mangelfull dokumentasjon for kliniske tester som daglig benyttes i klinisk praksis. Det synes å være behov for betydelig mer kunnskap for å oppnå tilstrekkelig gode diagnostiske prosedyrer for bekkenplager.

## 7. Konklusjon

Resultatene for studien indikerer at de benyttede kliniske testene har sammenheng med bekkenplagene hos denne gruppen kvinner. Dette er vist ved å presentere endringer i skår på de kliniske testene og egenrapportert smerteintensitet og funksjon fra ca 3 måneder etter fødsel til etter en 20 ukers behandlingsperiode. Det er også vist ved sammenligning av resultater målt samtidig etter behandling.

Det var stor spredning i andel positive skår på de 14 kliniske testene ved begge måletidspunkt. Etter behandling rapporterte kvinnene lavere smerteintensitet og bedret funksjon, og andelen positive skår var klart redusert for de aller fleste kliniske testene sammenlignet med før behandling.

Sammenhengen mellom de kliniske testene og kvinnenes egenrapportering av smerteintensitet og funksjon var gjennomgående svak til moderat. De tre testene ASLR, LDL og Patrick`s test ble funnet å ha en klar sammenheng med smerteintensitet og funksjon. Resultatene indikerer at det er disse tre av de 14 testene som har størst grad av validitet for egenrapportert smerteintensitet og funksjon ved bekkenplager. Det kan ikke konkluderes med manglende validitet for de øvrige kliniske testene.

## Kildeliste

- Albert, H., Godskesen, M., & Westergaard, J. (2001). Prognosis in four syndromes of pregnancy-related pelvic pain. *Acta Obstetrica et Gynecologica Scandinavica*, 80, 505-510.
- Albert, H., Godskesen, M., & Westergaard, J. (2000). Evaluation of clinical tests used in classification procedures in pregnancy-related pelvic joint pain. *European Spine Journal*, 9, 161-166.
- Albert, H. B., Godskesen, M., Korsholm, L., & Westergaard, J. G. (2006). Risk factors in developing pregnancy-related pelvic girdle pain. *Acta Obstetrica et Gynecologica Scandinavica*, 85, 539-544.
- Altman, D. (1999). *Practical statistics for medical research*. (9th ed.) London: Chapman & Hall.
- Andresen, E. M. (2000). Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine and Rehabilitation*, 81, S15-S20.
- Baker, D. J., Pynsent, P. B., & Fairbank, J. C. T. (1990). The Oswestry Disability Index revisited: its reliability, repeatability and validity, and a comparison with the St Thomas's Disability Index. In M.O.Roland & J. R. Jenner (Eds.), *Back Pain New approaches to rehabilitation and education* (pp. 174-186). Manchester University Press.
- Bastiaanssen, J. M., de Bie, R. A., Bastiaenen, C. H., Essed, G. G., & van den Brandt, P. A. (2005). A historical perspective on pregnancy-related low back and/or pelvic girdle pain. *Eur.J Obstet.Gynecol.Reprod.Biol.*, 120, 3-14.
- Benestad, H. B. & Laake, P. (2004). Forskning: Metode og planlegging. In H.B.Benestad & P. Laake (Eds.), *Forskningsmetode i medisin og biofag* (1st ed., pp. 83-113). Oslo: Gyldendal Norsk Forlag AS.
- Beurskens, A. J., de Vet, H. C., & Koke, A. J. (1996). Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain*, 65, 71-76.
- Bo, K. & Backe-Hansen, K. L. (2006). Do elite athletes experience low back, pelvic girdle and pelvic floor complaints during and after pregnancy? *Scand.J Med.Sci.Sports*.
- Bogduk, N. (2005). *Clinical Anatomy of the Lumbar Spine and Sacrum*. (4th ed.) Edinburgh London New York Oxford Philadelphia St Louis Sydney Toronto: Elsevier Churchill Livingstone.
- Bogduk, N. & Twomey, L. T. (1987). *Clinical Anatomy of the Lumbar Spine*. (1st ed.) Melbourne Edinburgh London and New York: Churchill Livingstone.
- Boissonnault, J. S. & Blaschak, M. J. (1988). Incidence of diastasis recti abdominis during the childbearing year. *Physical Therapy*, 68, 1082-1086.
- Bolton, J. E. & Wilkinson, R. C. (1998). Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients. *J Manipulative Physiol Ther.*, 21, 1-7.
- Broadhurst, N. A. & Bond, M. J. (1998). Pain provocation tests for the assessment of sacroiliac joint dysfunction. *Journal of Spinal Disorders*, 11, 341-345.
- Brodal, P. (2005). [The neurobiology of pain]. *Tidsskr.Nor Laegeforen.*, 125, 2370-2373.

---

Carreiro, J. E., Dery, M. A., Carroll, G., Devlin, S., Michaud, J., Manko, W. et al. (1998). The relationship between function of the long posterior sacroiliac ligament and low back pain. In A. Vleeming, V. Mooney, H. Tilscher, T. Dorman, & C. Snijders (Eds.), *3rd interdisciplinary World Congress on Low Back and Pelvic Pain* (pp. 199-206). European Conference Organizers Rotterdam.

Cibulka, M. T. & Koldehoff, R. (1999). Clinical usefulness of a cluster of sacroiliac joint tests in patients with and without low back pain.[see comment]. *Journal of Orthopaedic & Sports Physical Therapy*.29(2):83-9; discussion 90-2.

Collins, S. L., Moore, R. A., & McQuay, H. J. (1997). The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain*, 72, 95-97.

CRAM, R. H. (1953). A sign of sciatic nerve root pressure. *Journal of Bone and Joint Surgery.British Volume*, 35-B, 192-195.

Damen, L., Buyruk, H. M., Guler-Uysal, F., Lotgering, F. K., Snijders, C. J., & Stam, H. J. (2002). The prognostic value of asymmetric laxity of the sacroiliac joints in pregnancy-related pelvic pain. *Spine*, 27, 2820-2824.

de Groot M., Pool-Goudzwaard, A. L., Spoor, C. W., & Snijders, C. J. (2006). The active straight leg raising test (ASLR) in pregnant women: Differences in muscle activity and force between patients and healthy subjects. *Man.Ther.*

Deville, W. L., van der Windt, D. A., Dzaferagic, A., Bezemer, P. D., & Bouter, L. M. (2000). The test of Lasegue: systematic review of the accuracy in diagnosing herniated discs. *Spine*, 25, 1140-1147.

Deyo, R. A., Battie, M., Beurskens, A. J., Bombardier, C., Croft, P., Koes, B. et al. (1998). Outcome measures for low back pain research. A proposal for standardized use. *Spine*, 23, 2003-2013.

Domholdt, E. (1993). *Physical Therapy Research Principles and Applications*. Philadelphia, Pennsylvania: W.B.Saunders Company.

Dreyfuss, P., Michaelsen, M., Pauza, K., McLarty, J., & Bogduk, N. (1996). The value of medical history and physical examination in diagnosing sacroiliac joint pain. *Spine*, 21, 2594-2602.

Endresen, E. H. (1995). [Pelvic floor relaxation--a condition with many names and uncertain criteria]. *Tidsskr.Nor Laegeforen.*, 115, 3271-3273.

Estridge, M. N., Rouhe, S. A., & Johnson, N. G. (1982). The femoral stretching test. A valuable sign in diagnosing upper lumbar disc herniations. *J Neurosurg.*, 57, 813-817.

Fairbank, J. (2000). Revised Oswestry Disability questionnaire. *Spine*, 25, 2552.

Fairbank, J. C., Couper, J., Davies, J. B., & O'Brien, J. P. (1980). The Oswestry low back pain disability questionnaire. *Physiotherapy*, 66, 271-273.

Fairbank, J. C. & Pynsent, P. B. (2000). The Oswestry Disability Index. *Spine*, 25, 2940-2952.

Finch, E., Brooks, D., Stratford, P. W., & Mayo, N. E. (2002). *Physical Rehabilitation Outcome Measures A guide to Enhanced Clinical Decision Making*. (2nd ed.) Hamilton, Ontario: BC Decker Inc.

Forsberg, J. G. (1988). *Bäckenets og bäckenorganens anatomi*. (3rd ed.) Lund: Studentlitteratur.

Fortin, J. D., Aprill, C. N., Ponthieux, B., & Pier, J. (1994). Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique. Part II: Clinical evaluation. *Spine*, 19, 1483-1489.

Fortin, J. D., Dwyer, A. P., West, S., & Pier, J. (1994). Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique. Part I: Asymptomatic volunteers. *Spine*, *19*, 1475-1482.

Freburger, J. K. & Riddle, D. L. (2001). Using published evidence to guide the examination of the sacroiliac joint region.[see comment]. [Review] [31 refs]. *Physical Therapy*.*81(5):1135-43*.

Gifford, L. S. & Butler, D. S. (1997). The integration of pain sciences into clinical practice. *Journal of Hand Therapy*, *10*, 86-95.

Gilleard, W. L. & Brown, J. M. (1996). Structure and function of the abdominal muscles in primigravid subjects during pregnancy and the immediate postbirth period. *Physical Therapy*, *76*, 750-762.

Gray, H. (1989). *Gray's Anatomy*. (37th ed.) Edinburgh London Melbourne and New York: Churchill Livingstone.

Grotle, M., Brox, J. I., & Vollestad, N. K. (2003). Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. *J.Rehabil.Med.*, *35*, 241-247.

Grotle, M., Brox, J. I., & Vollestad, N. K. (2004). Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. *Spine*, *29*, E492-E501.

Grotle, M., Brox, J. I., & Vollestad, N. K. (2005). Functional status and disability questionnaires: what do they assess? A systematic review of back-specific outcome questionnaires. *Spine*, *30*, 130-140.

Gutke, A., Ostgaard, H. C., & Oberg, B. (2006). Pelvic girdle pain and lumbar pain in pregnancy: a cohort study of the consequences in terms of health and functioning. *Spine*, *31*, E149-E155.

Hansen, A., Jensen, D. V., Wormslev, M., Minck, H., Johansen, S., Larsen, E. C. et al. (1999). Symptom-giving pelvic girdle relaxation in pregnancy. II: Symptoms and clinical signs. *Acta Obstetrica et Gynecologica Scandinavica*, *78*, 111-115.

Hicks, C. M. (1999). *Research methods*. (3rd ed.) Edinburgh: Churchill Livingstone.

Holm, I., Friis, A., Storheim, K., & Brox, J. I. (2003). Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires: a reliability study. *Spine*, *28*, 828-833.

Hoppenfeld, S. (1976). *Physical Examination of the Spine and Extremities*. New York: Appleton-Century-Crofts, A Publishing Division of Prentice-Hall, Inc.

Hudson-Cook, N., Tomes-Nicholson, K., & Breen, A. (1989). A Revised Oswestry Disability Questionnaire. In M.O.Roland & J. R. Jenner (Eds.), *Back pain, New approaches to rehabilitation and education* (pp. 187-204). Manchester and New York: Manchester University Press.

Hungerford, B., Gilleard, W., & Lee, D. (2004). Altered patterns of pelvic bone motion determined in subjects with posterior pelvic pain using skin markers. *Clinical Biomechanics.*, 456-464.

Kerr, J. M., Vujnovich, A. L., & Bradnam, L. (2002). Changes in alpha-motoneuron excitability with positions that tension neural tissue. *Electromyography and Clinical Neurophysiology*, *42*, 459-471.

Kirkesola, G. & Solberg, A. (1998). *Klinisk undersøkelse av ryggen*. Kristiansand: Høyskoleforlaget AS - Norwegian Academic Press.

Knardahl, S. (2002). *Kropp og sjel Psykologi, biologi og helse*. (2nd ed.) Oslo: Gyldendal Norsk Forlag AS.

- 
- Kogstad, O. & Biornstad, N. (1990). [Pelvic girdle relaxation. Pathogenesis, etiology, definition, epidemiology]. *Tidsskr.Nor Laegeforen.*, 110, 2209-2211.
- Kokmeyer, D. J., Van der, W. P., Aufdemkampe, G., & Fickenscher, T. C. (2002). The reliability of multitest regimens with sacroiliac pain provocation tests. *Journal of Manipulative & Physiological Therapeutics*.25(1):42-8.
- Kreitz, B. G., Cote, P., & Yong-Hing, K. (1996). Crossed femoral stretching test. A case report. *Spine*, 21, 1584-1586.
- Kristiansson, P. & Svardsudd, K. (1996a). Discriminatory power of tests applied in back pain during pregnancy. *Spine*.21(20):2337-43; discussion 2343-4.
- Kristiansson, P., Svardsudd, K., & von, S. B. (1996b). Back pain during pregnancy: a prospective study. *Spine*, 21, 702-709.
- Kropmans, T. J., Dijkstra, P. U., Stegenga, B., Stewart, R., & de Bont, L. G. (1999). Smallest detectable difference in outcome variables related to painful restriction of the temporomandibular joint. *Journal of Dental Research*, 78, 784-789.
- Kuchera, W. A. & Kuchera, M. L. (1993). *Osteopathic Principles in Practice*. (2nd revised ed.) Ohio: Original Works Greyden Press Original Workbooks Columbus.
- Larsen, E. C., Wilken-Jensen, C., Hansen, A., Jensen, D. V., Johansen, S., Minck, H. et al. (1999). Symptom-giving pelvic girdle relaxation in pregnancy. I: Prevalence and risk factors. *Acta Obstetrica et Gynecologica Scandinavica*, 78, 105-110.
- Laslett, M. (2005). *Diagnostic Accuracy of the Clinical Examination Compared to Available Reference Standards in Chronic Low Back Pain Patients*. Faculty of Health Sciences, Linköpings universitet.
- Laslett, M., Aprill, C. N., McDonald, B., & Young, S. B. (2005). Diagnosis of Sacroiliac Joint Pain: Validity of individual provocation tests and composites of tests. *Man.Ther.*, 10, 207-218.
- Laslett, M. & Williams, M. (1994). The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine*, 19, 1243-1249.
- Laslett, M., Young, S. B., Aprill, C. N., & McDonald, B. (2003). Diagnosing painful sacroiliac joints: A validity study of a McKenzie evaluation and sacroiliac provocation tests. *Australian Journal of Physiotherapy*.49(2):89-97.
- Levin, U., Nilsson-Wikmar, L., Harms-Ringdahl, K., & Stenstrom, C. H. (2001). Variability of forces applied by experienced physiotherapists during provocation of the sacroiliac joint. *Clin.Biomech.(Bristol, Avon.)*, 16, 300-306.
- Levin, U., Nilsson-Wikmar, L., & Stenstrom, C. H. (2005). Variability within and between evaluations of sacroiliac pain with the use of distraction testing. *J Manipulative Physiol Ther.*, 28, 688-695.
- Levin, U., Nilsson-Wikmar, L., Stenstrom, C. H., & Lundeberg, T. (1998). Reproducibility of manual pressure force on provocation of the sacroiliac joint. *Physiotherapy Research International*, 3, 1-14.
- Levin, U. & Stenstrom, C. H. (2003). Force and time recording for validating the sacroiliac distraction test. *Clinical Biomechanics*.18(9):821-6.
- Maigne, J. Y., Aivaliklis, A., & Pfefer, F. (1996). Results of sacroiliac joint double block and value of sacroiliac pain provocation tests in 54 patients with low back pain. *Spine*, 21, 1889-1892.

McCombe, P. F., Fairbank, J. C., Cockersole, B. C., & Pynsent, P. B. (1989). 1989 Volvo Award in clinical sciences. Reproducibility of physical signs in low-back pain. *Spine*, *14*, 908-918.

Melzack, R. (2001). Pain and the neuromatrix in the brain. *Journal of Dental Education*, *65*, 1378-1382.

Mens, J. M., Damen, L., Snijders, C. J., & Stam, H. J. (2006). The mechanical effect of a pelvic belt in patients with pregnancy-related pelvic pain. *Clin.Biomech.(Bristol., Avon.)*, *21*, 122-127.

Mens, J. M., Vleeming, A., Snijders, C. J., Koes, B. W., & Stam, H. J. (2001). Reliability and validity of the active straight leg raise test in posterior pelvic pain since pregnancy. *Spine*, *26*, 1167-1171.

Mens, J. M., Vleeming, A., Snijders, C. J., Koes, B. W., & Stam, H. J. (2002a). Validity of the active straight leg raise test for measuring disease severity in patients with posterior pelvic pain after pregnancy. *Spine*, *27*, 196-200.

Mens, J. M., Vleeming, A., Snijders, C. J., Ronchetti, I., Ginai, A. Z., & Stam, H. J. (2002b). Responsiveness of outcome measurements in rehabilitation of patients with posterior pelvic pain since pregnancy. *Spine*, *27*, 1110-1115.

Mens, J. M., Vleeming, A., Snijders, C. J., Stam, H. J., & Ginai, A. Z. (1999). The active straight leg raising test and mobility of the pelvic joints. *Eur.Spine J*, *8*, 468-474.

Mens, J. M., Vleeming, A., Stoeckart, R., Stam, H. J., & Snijders, C. J. (1996). Understanding peripartum pelvic pain. Implications of a patient survey. *Spine*, *21*, 1363-1369.

Mogren, I. M. & Pohjanen, A. I. (2005). Low back pain and pelvic pain during pregnancy: prevalence and risk factors. *Spine*, *30*, 983-991.

Mousavi, S. J., Parnianpour, M., & Vleeming, A. (2007). Pregnancy related pelvic girdle pain and low back pain in an Iranian population. *Spine*, *32*, E100-E104.

Natvig, B., Nessioy, I., Bruusgaard, D., & Rutle, O. (1994). [Musculoskeletal complaints in a population. Occurrence and localization]. *Tidsskr.Nor Laegeforen.*, *114*, 323-327.

Netter, F. H. (1991). *Atlas of Human Anatomy*. Basle: Ciba-Geigy Limited.

Nilsson-Wikmar, L., Holm, K., Oijerstedt, R., & Harms-Ringdahl, K. (2005). Effect of three different physical therapy treatments on pain and activity in pregnant women with pelvic girdle pain: a randomized clinical trial with 3, 6, and 12 months follow-up postpartum. *Spine*, *30*, 850-856.

Noren, L., Ostgaard, S., Johansson, G., & Ostgaard, H. C. (2002). Lumbar back and posterior pelvic pain during pregnancy: a 3-year follow-up. *European Spine Journal*, *11*, 267-271.

norgeshelsa.no (2007). Norgeshelsa, inntekt/ utdanning, utdanningsnivå, prosent. <http://www.norgeshelsa.no/norgeshelsa/> [On-line].

O'Sullivan, P. B. & Beales, D. J. (2006). Changes in pelvic floor and diaphragm kinematics and respiratory patterns in subjects with sacroiliac joint pain following a motor learning intervention: A case series. *Man.Ther.*.

O'Sullivan, P. B., Beales, D. J., Beetham, J. A., Cripps, J., Graf, F., Lin, I. B. et al. (2002). Altered motor control strategies in subjects with sacroiliac joint pain during the active straight-leg-raise test. *Spine*, *27*, E1-E8.

Ostgaard, H. C. & Andersson, G. B. (1992). Postpartum low-back pain. *Spine*, *17*, 53-55.



- 
- Ostgaard, H. C., Andersson, G. B., & Karlsson, K. (1991). Prevalence of back pain in pregnancy. *Spine, 16*, 549-552.
- Ostgaard, H. C., Roos-Hansson, E., & Zetherstrom, G. (1996). Regression of back and posterior pelvic pain after pregnancy. *Spine, 21*, 2777-2780.
- Ostgaard, H. C., Zetherstrom, G., & Roos-Hansson, E. (1997). Back pain in relation to pregnancy: a 6-year follow-up. *Spine, 22*, 2945-2950.
- Ostgaard, H. C., Zetherstrom, G., & Roos-Hansson, E. (1994a). The posterior pelvic pain provocation test in pregnant women. *European Spine Journal.3(5):258-60*.
- Ostgaard, H. C., Zetherstrom, G., Roos-Hansson, E., & Svanberg, B. (1994b). Reduction of back and posterior pelvic pain in pregnancy. *Spine, 19*, 894-900.
- Pallant, J. (2004). *SPSS Survival manual, A step by step guide to data analysis using SPSS for Windows (Versions 10 and 11)*. (5th ed.) Glasgow: Bell & Bain Ltd.
- Pool-Goudzwaard, A. L., Sliker ten Hove, M. C., Vierhout, M. E., Mulder, P. H., Pool, J. J., Snijders, C. J. et al. (2005). Relations between pregnancy-related low back pain, pelvic floor activity and pelvic floor dysfunction. *Int.Urogynecol.J Pelvic.Floor.Dysfunct., 16*, 468-474.
- Pool-Goudzwaard, A. L., Vleeming, A., Stoeckart, R., Snijders, C. J., & Mens, J. M. (1998). Insufficient lumbopelvic stability: a clinical, anatomical and biomechanical approach to 'a-specific' low back pain. *Man.Ther., 3*, 12-20.
- Potter, N. A. & Rothstein, J. M. (1985). Intertester reliability for selected clinical tests of the sacroiliac joint. *Physical Therapy, 65*, 1671-1675.
- Revill, S. I., Robinson, J. O., Rosen, M., & Hogg, M. I. (1976). The reliability of a linear analogue for evaluating pain. *Anaesthesia, 31*, 1191-1198.
- Roach, K. E., Brown, M. D., Dunigan, K. M., Kusek, C. L., & Walas, M. (1997). Test-retest reliability of patient reports of low back pain. *J Orthop.Sports Phys.Ther., 26*, 253-259.
- Robinson, H. S., Brox, J. I., Robinson, R., Bjelland, E., Solem, S., & Telje, T. (2007). The reliability of selected motion- and pain provocation tests for the sacroiliac joint. *Man.Ther., 12*, 72-79.
- Roos-Hansson, E. & Zetherstrom, G. (1991). Tillförlitlighet av provokationstest vid bakre bäckensmärtor under graviditet. *Vetenskapligt supplement, 2*, 23-24.
- Rosier, E. M., Iadarola, M. J., & Coghill, R. C. (2002). Reproducibility of pain measurement and pain perception. *Pain, 98*, 205-216.
- Rost, C. C., Jacqueline, J., Kaiser, A., Verhagen, A. P., & Koes, B. W. (2006). Prognosis of women with pelvic pain during pregnancy: a long-term follow-up study. *Acta Obstetricia et Gynecologica Scandinavica, 85*, 771-777.
- Rost, C. C., Jacqueline, J., Kaiser, A., Verhagen, A. P., & Koes, B. W. (2004). Pelvic pain during pregnancy: a descriptive study of signs and symptoms of 870 patients in primary care. *Spine, 29*, 2567-2572.
- Scrimshaw, S. V. & Maher, C. (2001). Responsiveness of visual analogue and McGill pain scale measures. *J.Manipulative Physiol Ther., 24*, 501-504.
- Siconolfi, S. F., Lasater, T. M., Snow, R. C., & Carleton, R. A. (1985). Self-reported physical activity compared with maximal oxygen uptake. *American Journal of Epidemiology, 122*, 101-105.

Slipman, C. W., Jackson, H. B., Lipetz, J. S., Chan, K. T., Lenrow, D., & Vresilovic, E. J. (2000). Sacroiliac joint pain referral zones. *Archives of Physical Medicine and Rehabilitation*, 81, 334-338.

Slipman, C. W., Sterenfeld, E. B., Chou, L. H., Herzog, R., & Vresilovic, E. (1998). The predictive value of provocative sacroiliac joint stress maneuvers in the diagnosis of sacroiliac joint syndrome. *Archives of Physical Medicine and Rehabilitation*, 79, 288-292.

Stankovic, R., Johnell, O., Maly, P., & Willner, S. (1999). Use of lumbar extension, slump test, physical and neurological examination in the evaluation of patients with suspected herniated nucleus pulposus. A prospective clinical study. *Man.Ther.*, 4, 25-32.

Steingrimsdottir, O. A., Vollestad, N. K., Roe, C., & Knardahl, S. (2004). Variation in reporting of pain and other subjective health complaints in a working population and limitations of single sample measurements. *Pain*, 110, 130-139.

Strender, L. E., Sjoblom, A., Sundell, K., Ludwig, R., & Taube, A. (1997). Interexaminer reliability in physical examination of patients with low back pain. *Spine*, 22, 814-820.

Stuge, B., Laerum, E., Kirkesola, G., & Vollestad, N. (2004). The efficacy of a treatment program focusing on specific stabilizing exercises for pelvic girdle pain after pregnancy: a randomized controlled trial. *Spine*.29(4):351-9.

Sturesson, B., Uden, G., & Uden, A. (1997). Pain pattern in pregnancy and "catching" of the leg in pregnant women with posterior pelvic pain. *Spine*, 22, 1880-1883.

To, W. W. & Wong, M. W. (2003). Factors associated with back pain symptoms in pregnancy and the persistence of pain 2 years after pregnancy. *Acta Obstetricia et Gynecologica Scandinavica*, 82, 1086-1091.

Van der Wurff, P., Hagmeijer, R. H., & Meyne, W. (2000a). Clinical tests of the sacroiliac joint. A systemic methodological review. Part 1: reliability. [Review] [36 refs]. *Manual Therapy*.5(1):30-6.

Van der Wurff, P., Meyne, W., & Hagmeijer, R. H. (2000b). Clinical tests of the sacroiliac joint. [Review] [30 refs]. *Manual Therapy*.5(2):89-96.

van Deursen, L. L. J. M., Patijn, J., Ockhuysen, A. L., & Vortman, B. J. (1990). The value of some clinical tests of the sacroiliac joint. *J Manual Medicine*, 5, 96-99.

van Wingerden, J. P., Vleeming, A., Buyruk, H. M., & Raissadat, K. (2004). Stabilization of the sacroiliac joint in vivo: verification of muscular contribution to force closure of the pelvis. *European Spine Journal*, 13, 199-205.

Vleeming, A., Albert, H. B., Ostgaard, H. C., Stuge, B., & Sturesson, B. (2004). European Guidelines On The Diagnosis And Treatment Of Pelvic Girdle Pain. In *5th Interdisciplinary World Congress on Low Back & Pelvic Pain Effective Diagnosis and Treatment of Lumbopelvic Pain November 10-13 2004 Melbourne* (pp. 6-49). Melbourne Australia: ECO.

Vleeming, A., de Vries, H. J., Mens, J. M., & van Wingerden, J. P. (2002). Possible role of the long dorsal sacroiliac ligament in women with peripartum pelvic pain. *Acta Obstetricia et Gynecologica Scandinavica*, 81, 430-436.

Vleeming, A., Mens, J., de Vries, H., van Wingerden, J. P., & Pool, A. (1998). Possible role of the long dorsal sacroiliac ligament in peripartum pelvic pain. In A.Vleeming, V. Mooney, H. Tilscher, T. Dorman, & C. Snijders (Eds.), *3rd interdisciplinary World Congress on Low Back and Pelvic Pain* (pp. 149-156). European Conference Organizers Rotterdam.

---

Vleeming, A., Pool-Goudzwaard, A. L., Hammudoghlu, D., Stoeckart, R., Snijders, C. J., & Mens, J. M. (1996). The function of the long dorsal sacroiliac ligament: its implication for understanding low back pain. *Spine, 21*, 556-562.

Vleeming, A., Snijders, C. J., Stoeckart, R., & Mens, J. M. (1997). The role of the sacroiliac joints in coupling between spine, pelvis, leg and arms. In A. Vleeming, V. Mooney, T. Dorman, & et al (Eds.), *Movement, stability and low back pain: The essential role of the pelvis* (1997 ed., pp. 53-71). New York: Churchill Livingstone.

Von Korff, M., Jensen, M. P., & Karoly, P. (2000). Assessing global pain severity by self-report in clinical and health services research. *Spine, 25*, 3140-3151.

WHO (2001). *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization.

Willard, F. H., Carreiro, J. E., & Manko, W. (1998). The long posterior interosseus ligament and the sacrococcygeal plexus. In A. Vleeming, V. Mooney, H. Tilscher, T. Dorman, & C. Snijders (Eds.), *3rd interdisciplinary World Congress on Low Back and Pelvic Pain* (pp. 207-209). European Conference Organizers Rotterdam.

Williamson, A. & Hoggart, B. (2005). Pain: a review of three commonly used pain rating scales. *Journal of Clinical Nursing, 14*, 798-804.

Wormslev, M., Juul, A. M., Marques, B., Minck, H., Bentzen, L., & Hansen, T. M. (1994). Clinical examination of pelvic insufficiency during pregnancy. An evaluation of the interobserver variation, the relation between clinical signs and pain and the relation between clinical signs and physical disability. *Scandinavian Journal of Rheumatology, 23*, 96-102.

Wu, W. H., Meijer, O. G., Uegaki, K., Mens, J. M., Van Dieen, J. H., Wuisman, P. I. et al. (2004). Pregnancy-related pelvic girdle pain (PPP), I: Terminology, clinical presentation, and prevalence. *Eur. Spine J.*

Zelle, B. A., Gruen, G. S., Brown, S., & George, S. (2005). Sacroiliac joint dysfunction: evaluation and management. *Clinical Journal of Pain, 21*, 446-455.