

Diagnostisk nøyaktighet ved fire kliniske
provokasjonstester for subacromialt
impingement



Manuellterapi teori: MANT395

Masterprogram i helsefag – Klinisk masterstudium i
manuellterapi for fysioterapeuter
Institutt for samfunnsmedisinsk fag

Kandidatnummer: 209103

Antall ord: 10288

(inkludert referanser i tekst, ekskludert figurer/tabeller)

Universitetet i Bergen

Høsten 2012

Innholdsfortegnelse

SAMMENDRAG	1
ABSTRACT	2
1.0 INTRODUKSJON	3
1.1 BAKGRUNN	3
1.2 RELEVANT TEORI	4
1.2.1 SKULDERENS FUNKSJONELLE ANATOMI	4
1.2.2 GENERELT OM SUBACROMIALT IMPINGEMENTSYNDROM (SIS)	5
1.2.3 DIAGNOSTISK ULTRALYD	6
1.2.4 SMERTE	7
1.2.5 BRUK AV MÅLEMETODER	7
1.2.6 VALIDITET OG RELIABILITET	8
1.2.7 TIDLIGERE FORSKNING PÅ OMRÅDET	11
2.0 HENSIKT OG PROBLEMSTILLING	15
2.1 HENSIKT	15
2.2 PROBLEMSTILLING	15
2.2.1 BEGREPSSPESIFISERING	15
3.0 METODE	16
3.1 VALG AV FORSKNINGSDESIGN	16
3.2 UTVALG	16
3.3 UTVELGELSESPROSESSEN	18
3.4 PROSEDYRE OG DATAINNSAMLING	19
3.5 ANALYSE AV DATAMATERIALET	22
3.6 ETISKE ASPEKTER	22
4.0 RESULTATER	24
4.1 BASELINEDATA	24
4.2 KRYSSABELLER	24
5.0 DISKUSJON	26
5.1 TIDLIGERE STUDIER	27
5.2 KONFOUNDERE	29
5.3 EKSTERN VALIDITET	33
5.4 VIDERE FORSKNING	34
6.0 KONKLUSJON	34
REFERANSELISTE	35
VEDLEGG	42
VEDLEGG 1: ANONYMT SKÅRINGSKJEMA	42
VEDLEGG 2: FORESPØRSEL OM DELTAGELSE I STUDIEN/SAMTYKKE	43
VEDLEGG 3: GODKJENNING FRA REK	47
VEDLEGG 4: BASELINEDATA OG TESTSCORE	49

**TITTEL: DIAGNOSTISK NØYAKTIGHET VED FIRE KLINISKE
PROVOKASJONSTESTER FOR SUBACROMIALT IMPINGEMENT
- En tverrsnittsstudie av 33 deltagere i Norge.**

**Klinisk masterstudium i manuellterapi for fysioterapeuter, institutt for
samfunnsmedisinsk fag, Universitetet i Bergen.**

ÅR: 2012

Sammendrag

Bakgrunn

Skuldersmerter er en av de vanligste plagene i befolkningen. For å kunne gi optimal behandling er det viktig å diagnostisere skulderlidelsene med tester som er valide og reliable. Målet med denne studien var derfor å evaluere den diagnostiske nøyaktigheten til fire av de mest anvendte impingementtester for skulderen.

Problemstilling

”Hvor diagnostisk nøyaktig er fire kliniske provokasjonstester for subacromialt impingement ved bruk av ultralyd som referanseramme?”

Metode

Denne studien er en tverrsnittsstudie av diagnostisk nøyaktighet. Halvparten av deltagerne ($n=17$, 53% menn, gjennomsnittsalder: 47) ble fortløpende rekruttert etter diagnostisering ved hjelp av ultralyd. I tillegg ble 16 deltagere uten skulderplager rekruttert. Sensitivitet, spesifisitet, positiv- og negativ prediktiv verdi, positiv- og negativ og nøyaktigheten av testene ble beregnet i en 2x2 tabell.

Resultater

Neers, Hawkins, Jobe og Painful Arc hadde henholdsvis 64%, 71%, 82% og 64% nøyaktighet (sensitivitet 77%, 92%, 77%, 70%, spesifisitet 25%, 0%, 100%, 50%) sammenliknet med annen skulderpatologi. Ved sammenlikning med friske hadde Neers, Hawkins, Jobe og Painful Arc henholdsvis 90%, 90%, 90% og 86% nøyaktighet (sensitivitet 77%, 92%, 77%, 70%, spesifisitet 100%, 88%, 100%, 100%).

Konklusjon

Neers, Hawkins og Painful Arc test indikerte ved sammenlikning mot annen skulderpatologi liten sannsynligheten av impingement ved negativ test, mens Jobe test indikerte at sjansen for impingement var meget stor ved positiv test. Jobe test viste også den høyeste nøyaktigheten med 82% treffsikkerhet. Ved sammenlikning mot friske indikerer Neers, Jobe og Painful Arc at sjansen for impingement var stor ved positiv test. Hawkins test indikerte i motsetning at sjansen for impingement var liten ved negativ test.

Nøkkelord

Diagnostisk nøyaktighet, subacromialt impingementsyndrom, kliniske tester, diagnostisk ultralyd.

DIAGNOSTIC ACCURACY OF FOUR PHYSICAL EXAMINATIONS MANEUVERS FOR SUBACROMIAL IMPINGEMENT

- A cross section study with 33 participants in Norway

Master in Health Sciences- Manuel therapy, Department of Public Health and Primary Health Care, The university of Bergen

YEAR: 2012

Abstract

Objective

Shoulder pain is a common complaint in the population. For optimal treatment, it is important with valid and reliable tests. The aim of the present study was to evaluate the discriminate ability of four physical tests used to assess subacromial impingement syndrome.

Research question

“How is the discrimination ability of four physical tests in diagnosing subacromial impingement syndrome confirmed by diagnostic ultrasound?”

Methods

The study is a cross sectional study assessing the diagnostic accuracy of four physical shoulder tests. The participants ($n=17$, Male 53%, mean age: 47) were recruited after being diagnosed by ultrasound. 16 participants without shoulder pain were also recruited. Sensitivity, specificity, positive and negative predictive values and diagnostic accuracy of the tests were examined.

Result

The diagnostic accuracy of Neers, Hawkins, Jobe and Painful Arc tests was 64%, 71%, 82% and 64% respectively (sensitivity was 77%, 92%, 77%, 70% and specificity was 25%, 0%, 100%, 50%) compared with other shoulder pathology. Compared to the asymptomatic participants, the diagnostic accuracy of Neers, Hawkins, Jobe og Painful Arc Tests was 90%, 90%, 90% and 86% respectively (sensitivity was 77%, 92%, 77%, 70% and spesificity 100%, 88%, 100%, 100%).

Conclusion

Compared to other shoulder pathology, Neers, Hawkins and Painful Arc Tests demonstrated little probability of impingement with negative test results, while Jobe Test indicated a high chance of impingement with positiv test result. The test with the highest diagnostic accuracy was Jobe test. Compared to the asyphmatic participants, Neers, Jobe and Painful Arc Tests indicated high likelihood for impingement with a positiv test result. In contrast, Hawkins Test indicated little probability of impingement with a negative test result.

Keywords

Diagnostic Accuracy, subacromial impingement syndrome, physical examination, diagnostic ultrasound.

1.0 Introduksjon

1.1 Bakgrunn

Skulderproblematikk er en alminnelig problematikk i klinisk praksis. Arbeidstilsynet rapporterte uspesifikke skulderlidelser som en av de hyppigste meldte diagnosene i perioden 2006-2010 (Svendsen, 2011). Dette samsvarer med empiriske studier som viser at opp mot 50 % av befolkningen hadde en skulderlidelse sist år (Brox et al. 2011). Omfanget av skulderlidelsene resulterer dermed også i høye økonomiske kostnader. Arbeidstilsynet har estimert at sykefravær knyttet til muskel – og skjelettplager koster det norske samfunnet ca. 40 – 50 milliarder årlig i direkte og indirekte utgifter (Ihlebak & Lærum, 2004). En amerikansk studie fant at de direkte økonomiske kostnadene knyttet til behandling av dysfunksjon i skulder i 2000 beløp seg til 7 milliarder dollar (Meislin et al. 2005). Det er derfor viktig å kartlegge denne pasientgruppe.

De vanligste atraumatiske skulderlidelsene er uspesifikke skuldresmerter (myalgier), subakromiale smerter med eller uten degenerative rupturer, adhesiv kapsulitt og akromioklavikulærleddsartrose (Ekeberg et al. 2009). I en Engelsk studie hvor 392 hadde skuldresmerter fant Van der Windt et al. (1996) at impingement var den vanligste årsaken til skuldresmertene. Dette er i tråd med en annen engelsk undersøkelse av Östör et al. (2005) som fant tegn til impingement hos 74% av pasientene som oppsøkte allmennlege grunnet skuldresmerter. Diagnosen Subacromialt impingement syndrom (SIS) opptrer i det subacromiale rommet når den subacromiale bursa og rotatorcuffen avklemmes mellom caput humeri og acromion eller mot det coracoacromiale ligamentet (Kelly et al. 2010).

Det er viktig å diagnostisere skulderlidelsene for å kunne gi den optimale behandling (Brox et al. 2011, Michener et al. 2009). Det er fremdeles ikke en enhetlig konsensus vedrørende diagnostiske kriterier av impingement (Awerbuch, 2008). Røntgen, artroskopi, ultralyd, MR og MR artrografi blir ofte brukt som supplerende undersøkelse ved skulderplager (Silva et al. 2008, Calis et al. 2000, Kelly et al. 2010, Michener et al. 2009, Juel et al. 2007, Kromer et al. 2010). Forskere mener at artroskopi er den mest optimale undersøkelse for kartlegging av skulderplager (Park et al. 2005, Malhi & Khan, 2005). Dette underbygges av empirisk forskning som demonstrerer at det er 95 % samsvar mellom funn på MR og ultralyd sammenlignet med artroskopi (Malhi & Khan, 2005). Samtidig viser forskning at ultralyd egner seg spesielt godt for skulderundersøkelser fordi det er enkelt å sammenlikne med den

andre skulderen, samt at skulder/arm kan beveges i forskjellige posisjoner for å analysere strukturene i forhold til hverandre (Juel et al, 2007). Ultralyd har også den fordel at det ikke er noen kontraindikasjoner og gir heller ikke smerte sammenliknet med for eksempel artroskopi, hvor en del pasienter kan bli ekskludert grunnet kontraindikasjoner (Itoi et al. 2006, Holtby & Razmjou 2004, MacDonald et al. 2000). Et positivt funn på ultralyd vil på bakgrunn av dette være et av inklusjonskriteriene i denne oppgaven.

Ved anvendelse av kliniske tester som målemetoder i praksis forutsetter dette at testene er valide og reliable. I tillegg er det viktig å inkludere flere forskjellige utfallsmål for å fange opp og bekrefte diagnosen (Juel, 2007). Innen fysioterapien blir det lagt vekt på å bruke standardiserte målemetoder for å undersøke og evaluere klinisk praksis (Gross et al. 2006). Testene brukt i denne studien for impingement er derfor nøye utvalgt. Hawkins og Neers test har blant annet vært gjenstand for kadaverstudier som viser at strukturer i det subacromiale rommet kan gi avklemming (Valadie et al. 2000). I tillegg er kostnaden liten ved bruk av kliniske skuldertester.

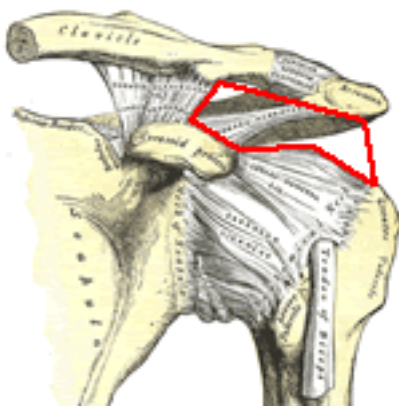
Kelly (2010) mener at det ikke finnes entydige gode resultater av de få klinisk validitets- og reliabilitetsstudiene ved bruk av ultralyd som referanseramme for kliniske tester. Andre studier ved bruk av MR, røntgen og artroskopi viser også varierende resultater (Calis et al. 2000, Park et al. 2005, MacDonald et al. 2000). En metaanalyse indikerte at fysiske tester kan være nyttige ved rotator cuff syndromer (høy sensitivitet), men mindre nyttig for å skille de forskjellige strukturene (lav spesifisitet) (Dinnes et al. 2003). I min praksis som fysioterapeut erkjenner jeg at det er et stort behov for kvalitetssikrede tester. Derfor ønsker jeg å evaluere og drøfte den diagnostiske nøyaktigheten til fire kliniske tester benyttes i diagnostikk av skulderimpingement ved bruk av ultralyd som referanseramme. Studien vil gjøres i samråd med to medforskere som vil analysere inter – og intrareliabiliteten til de samme fire testene. Jeg vil senere i oppgaven gjøre rede for utvelgelsen av de fire testene.

1.2 Relevant teori

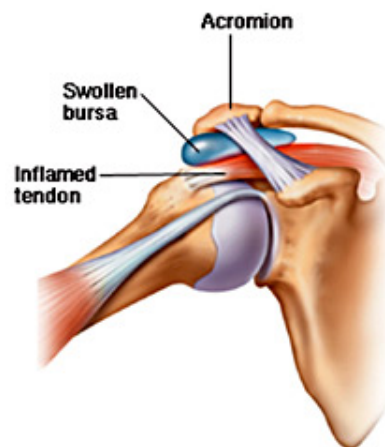
1.2.1 Skulderens funksjonelle anatomi

Grunnlaget for en målrettet funksjonsundersøkelse baseres på kunnskap om anatomi og biomekanikk i skulderen (Juel, 2007). Skulderen består av tre knokler, tre ekte ledd med ligamenter og mer enn 20 muskler. De ekte leddene er glenohumoralleddet,

akromioklavikulærleddet og sternoklavikulærleddet. I tillegg til disse ledd er det en bevegelig forbindelse mellom skulderbladet og thoraxveggen som kalles thoracoskapulærleddet. I glenohumoralleddet stilles det store krav til bløtdelenes stabilisering, da det er dårlig benet stabilitet (Bojsen-Møller, 2012). Bevegelsen av hvordan humerus opptrer i forhold til scapula kalles den humeroscapulære rytmen. Nyere forskning viser at rytmen ikke er konstant og at det er store individuelle forskjeller (Juel, 2007). Ved abduksjon av skulderen mellom 80 og 140 grader er bevegelsen av scapula størst. Dette tilsvarer smertebuen (Painful arc) ved subacrominalt smertesyndrom og tyder på at plassforholdnene subacromialt er dårligst i denne delen av bevegelsesbanen. Smerte gir ofte en reflektorisk påvirkning av bevegelser i skulderen. Dette kan gi opphav til en dysfunksjonell bevegelse og kan forverre eller vedlikeholde smerten (Juel, 2007). På figur 1 og 2 vises det subacromiale rommet i skulderen.



Figur 1: Subacromiale rommet markert med rødt.



Figur 2 viser en inflammasjon i det subacromiale rommet som fører til redusert plassforhold, da bursaen (slimposen) hovner opp og senen blir betent.

1.2.2 Generelt om Subacromialt impingementsyndrom (SIS)

Progresjonen av SIS er vanligvis klassifisert etter inflammasjons eller degenerasjons stadiet av rotator cuffen eller bursaen (Kelly, 2010). Neer (1972) beskrev tre stadier av impingement syndromet. Stadium 1 kjennetegnes av reversibel inflammasjon (ødem og blødninger) i bursa og rotatorcuff og sees typisk hos unge mennesker (figur 2). Stadium 2 kjennetegnes av mer irreversible kroniske forandringer, fibrose og tendinit i cuffen og sees typisk hos mennesker mellom 25 og 40 år. I stadium 3 tilkommer ytterligere kroniske forandringer som partielle eller totale cuffrupturer og pasientene er gjerne over 40 år (Neer, 1972).

Årsak til impingement

Det er flere medvirkende faktorer til utviklingen av impingement syndrom i skulderen inkludert muskel svakhet/ubalanse i cuffen, overbelastning, særlig ved arbeid over skulderens plan, aldersforandringer i cuffen (degenerativ tendinopati), utvikling av artrose i acromioclavicularleddet, med påleiringer som kan buke ned i cuffen, og acromions form (Bigliani et al. 1986, Torstensen et al. 2005).

Bigliani et al. (1986) definerte tre anatomiske varianter av acromions form; (1) flat, (2) buet og (3) krokformet. De fant at full-tykkelse rifter i cuffen (Neer stadium 3 impingement) var hyppigst assosiert med krokformet (type 3) acromion. Forskning tyder på at type 2 og type 3 disponerer for subacromialt impingement (Torstensen et al. 2005).

Differensialdiagnoser

I følge Brox et al. (2010) finnes følgende differensialdiagnoser til Subacromialt Impingementsyndrom: Glenohumeral artrose, inflammatorisk eller infeksøs tenosynovitt eller artritt, leddleppeaffeksjon, bicepsseaffeksjon, multidireksjonal instabilitet, nevritt, nerveskade pga. trykk (entrapment), nerverotsmerter fra nakken, thoraxapertursyndrom, sympatisk refleksdystrofi eller skulder-hånd-syndrom og tumor (Brox et al. 2010). I klinisk sammenheng er det viktig å kartlegge pasienten og dermed utelukke differensialdiagnosene ved SIS. I denne studien vil pasienter med disse diagnosene ekskluderes i SIS gruppen ved ultralyd.

1.2.3 Diagnostisk Ultralyd

Skulderundersøkelse ved hjelp av ultralyd har vært brukt i ca. 30 år med stor teknologisk og faglig utvikling (Moosmayer et al. 2004). Metoden bidrar til en mer presis diagnostikk av bløte vevstrukturer. Undersøkelsen gjøres som regel etter en standard hvor det brukes et visst antall snitt for å kartlegge aktuelle strukturer (Juel, 2007).

Billedannelsen ved ultralyd skjer ved hjelp av lydbølger med forskjellige hastigheter og oppløseligheter reflekteres på grenseflaten mellom ulike vevstyper. Signalene sendes og mottas av høyfrekvente lydholder (Moosmayer et al. 2004). Det er mange fordeler ved å bruke Ultralyd ved undersøkelse av skulder fremfor annen bildediagnostikk. Ultralydundersøkelsen er kostnadseffektivt, plassbesparende, utføres relativt raskt og er tilstrekkelig for mange tilstander (Juel, 2007). Undersøkelsen er også non-invasiv, uten strålebelastning og gir ingen

bivikninger eller kontraindikasjoner. I mange land anbefales ultralyd som første bildeundersøkelse av senevev (Juel, 2007). De bløte strukturene som ofte bidrar til subcromiale smerter kan kartlegges ved ultralyd, for eksempel tendinopatier, kalknedslag, bursareaksjoner og cuffrupturer (Kelly et al. 2010, Juel, 2007).

1.2.4 Smerte

Smerte defineres som en ubehagelig sensorisk og empsjonell opplevelse som opptrer i sammenheng med vevsskade eller truende vevsskade (Brodal, 2007). Symptomene ved subacromialt impingementsyndrom er typisk sviende smerter i fremre ytre del av skulder og yttersiden av øvre del av overarm (Torstensen et al. 2005). Smertene forverres ved bruk av armen over skulder høyde. Plagene er initialt anfallsvis for etter hvert å bli mer konstante (Harvey, 2009). Ved smerter i skulder er det i flere studier funnet innvekst av frie nerveender i senevev hvor nervotransmittere som substans P, glutamat og kalsitoningsenrealtert peptid identifisert (Torstensen et al. 2005, Brodal, 2007). Disse formidler smerte og finnes i både senen og den subacromiale bursaen og kan være en av grunnene til smertefulle tilstander ved SIS (Juel, 2007, Brodahl, 2007). Dersom smertene blir langvarige vil det kunne føre til økt sensitivisering og smerten oppleves kraftigere enn det smerteirritamentet skulle tilsi (Brodahl, 2007). Dette fører til økt sensitivisering i smertebanene og noen pasienter kan få ledsagende utstråling til albue, hånd og eventuelt fingre. Mange kan også plages av prikking, stikking eller nummenhet i finger (Brodahl, 2007).

1.2.5 Bruk av målemetoder

En viktig del av fysioterapien og manuellterapien i klinisk praksis er å kunne vurdere pasienter i rehabilitering (Gross et al. 2006). Det kan for eksempel være å klassifisere, vurdere og evaluere utfall av eventuelle dysfunksjoner, fremgang i rehabilitering eller predikere videre forløp av behandling. Det blir derfor lagt stor vekt på bruken av standardiserte målemetoder for å undersøke og evaluere klinisk praksis (Gross et al. 2006). Det er svært sentralt med gode målemetoder basert på målbare definisjoner av fokusområdet til en målemetode.

Måleredskapene kan brukes som i kommunikasjon med kollegaer, på tvers av faggrupper og i kvalitetssikring og forskning (Gross et al. 2006). Det vil derfor være svært sentralt å undersøke den kliniske nøyaktigheten av skuldertester.

1.2.6 Validitet og reliabilitet

Validitet beskriver en målemetodes evne til å måle det den har til hensikt å måle (Portney & Watkins, 2000, Streiner & Norman, 2008, Ringdal, 2009). Med andre ord innebærer det gyldigheten av testen. En forutsetning for høy validitet er høy reliabilitet (Ringdal, 2009). Med reliabilitet menes pålitelighet, hvilket innebærer at målinger må utføres korrekt og at eventuelle feilmarginer angis (Dalland, 2007). Validitet er det mest generelle av de to begrepene. Med dette menes at reliabilitet er et rent empirisk spørsmål, mens validitet krever i tillegg en teoretisk vurdering. Spørsmålet om validitet må derfor alltid refereres til den teoretiske sammenheng begrepet brukes i (Ringdal, 2009). I denne oppgaven vil jeg konsentrere meg om testvaliditet.

Testvaliditet

Med testvaliditet menes validiteten av tester som skal skille mellom syke og friske personer. Validiteten av en test må vurdere om testen klassifiserer riktig med henblikk på sykdom. Det finnes forskjellige tester for å skille friske fra syke eller for å predikere sykdom eller helseskade (Benestad & Laake 2004).

Ulike verdier fra statistisk analyse kan benyttes for å presentere testenes diagnostiske nøyaktighet herunder sensitivitet, spesifisitet, testens prediktiv evne, likelihood ratio, konfindensintervall og nøyaktighet (Hattam & Smeatham 2010). Dette vil jeg forklare nærmere i de kommende avsnittene.

Sensitivitet og spesifisitet

For å undersøke om skulderimpingement-testene kan brukes til å predikere sykdom er vi interessert i sammenhengen mellom test og sykdom, og test og ultralyd (Tabell 1 og 2). Testene blir mer valide jo flere testene klarer å karakterisere som henholdsvis syke og friske. For å vurdere testens validitet inndeles testpersonene i syke, friske, testpositive og testnegative (Tabell 1 og 2). Testens sensitivitet og spesifisitet er uttrykk for testens validitet og fanger opp de som ble syke og skiller ut de som forble friske (Benestad & Laake 2004).

Diagnostisk sensitivitet kan defineres som testens evne til å identifisere individer med tilstanden (positive) i forhold til antallet med tilstanden identifisert av referansetesten og

uttrykkes ved $\text{Positiv}/(\text{Positiv}+\text{Falsk Negativ})$ (tabell 1 og 2) (Benestad & Laake, 2004). Diagnostisk spesifisitet defineres tilsvarende som testens evne til å ekskludere individer uten tilstanden (negativ) og det vil si $\text{Negativ}/(\text{Falsk positiv}+\text{Negativ})$. Med andre ord er sensitivitet evnen til å avsløre at en person er syk, mens spesifisitet er testens evne til å avsløre om en person er frisk (Løvås, 2004). Sensitivitet og spesifisitet er alltid relative verdier basert på en referanse (Benestad & Laake 2004). Sensitivitet og spesifisitet regnes ut som ratio, men angis også som prosentvise verdier med 95 % konfidensintervall (KI) (Se forklaring nedenfor) (Hattam & Smeatham, 2010). Sensitivitet vil være 100% dersom ingen falske negative og spesifisitet vil være 100% dersom ingen falske positive (Hattam & Smeatham, 2010).

Tabell 1 og 2: Sammenheng mellom test og sykdom (2x2 tabell)

	Gruppe 2: Annen skulderpatologi	Gruppe 1: Impingement
Impingement tester negative	Negative (N)	Falske negative (FN)
Impingement tester positive	Falske positive (FP)	Positive (P)

	Gruppe 3: Friske	Gruppe 1: Impingement
Impingement tester negative	Negative (N)	Falske negative (FN)
Impingement tester positive	Falske positive (FP)	Positive (P)

(Benestad & Laake, 2004)

Konfidensintervall

Et konfidensintervall er en måte å angi feilmarginen av en måling eller en beregning på. Konfidensintervallet angir intervallet som med en spesifisert sannsynlighet inneholder den sanne verdien av variabelen man har målt. Sannsynligheten angis som et 95 % konfidensintervall med en sannsynlighet på 0,95 (Ringdal, 2009).

Testens prediktive evne

Testens prediktive evne vurderer i hvilken grad de testpositive virkelig er syke. Den kan enten

handle om forutsigbarhet ved et positivt testresultat eller ved et negativt testresultat (Benestad & Laake, 2004). Testens prediktive evne ved et positivt resultat (PPV) uttrykkes ved $P/(P+FP)$. Tilsvarende er $N/(N+FN)$ uttrykk for testens prediktive evne ved negativt resultat (NPV). Det er ønskelig at både PPV og NPV er så nær 1 som mulig. Den predikative evnen av en test er avhengig av prevalensen av tilstanden. Argumentene for og imot testing må derfor alltid inneholde en vurdering av hvor mye det vil bety for de som kommer til å bli syke blant de testpositive, å bli oppdaget på forhånd og tatt under behandling (Benestad & Laake, 2004).

Likelihood ratios (Sannsynlighetsandeler)

Likelihood ratios (LR) er statistikk for summering av diagnostisk nøyaktighet og er kanskje den beste måten å vurdere den kliniske nytteverdien av en test på. LR er estimert ved en kombinasjon av sensitivitet og spesifisitet ((LR-: $(1 - \text{Sensitivitet})/\text{Spesifisitet}$, LR+: $\text{Sensitivitet}/(1 - \text{Spesifisitet})$). En LR på 1 indikerer at testresultatet ikke medvirker til å favorisere tilstanden, mens en LR høyere enn 1 øker sjansen for at testresultatene er assosiert med sykdomstilstanden og en LR mindre enn 1 svekker sjansen for sykdomstilstanden. Den beste test definisjonen for å avgjøre tilstanden er en tilstand med høyest positiv LR (LR+), mens den beste definisjonen for å avgjøre ut av en tilstand vil ha den minste negative LR (LR-)(Altmann, 1991). Det vil med andre ord si at LR lengst mulig fra 1 vil gi sterk evidens for sykdom. LR over 10 eller under 0.1 er i de fleste tilfeller forbundet med sterke bevis for å bestemme om testen er diagnostisk representativ (Altmann & Deeks, 2004, Hattam & Smeatham, 2010). Det vil med andre ord si at ved LR+ lik 10 vil det være 10 ganger større sannsynlighet for å få positiv test når man er syk enn når man er frisk. For hver 10. positive test vil det likevel være én falsk positiv. Ved LR- på 0,1 vil sjansen for å ha negativ test for en som er syk være en tiendedel av hva den er for de som ikke er syke. For hver 10. negative test vil det være én falsk negativ.

Nøyaktighet

Nøyaktighet beskrives som antall korrekte klassifikasjoner (Positive+Negative) i forhold til alle klassifikasjonene. Ved en perfekt segmentering vil nøyaktigheten bli 1 (100%) (Lu et al. 2004). I tabell 4 vises en oppsummering av formlene for sensitivitet, spesifisitet, KI, PPV, NPV, LR+, LR- og nøyaktighet.

Tabell 4: Formlen for sensitivitet, spesifisitet, KI, PPV, NPV, LR+, LR- og nøyaktighet

Sensitivitet	$P/(P+FN)$
Spesifisitet	$N/(FP+N)$
KI	$\frac{a}{n} \pm 1.96 \sqrt{\frac{\frac{a}{n}(1 - \frac{a}{n})}{n}}$
PPV	$P/(P+FP)$
NPV	$N/(FN+N)$
LR+	Sensitivitet/ (1-spesifitet)
LR-	(1- sensitivitet)/spesifisitet
Nøyaktighet	$(P+N)/(P+FP+FN+N)$

<i>KI = Konfidensintervall</i>	<i>P = Positiv</i>
<i>PPV = Positiv prediktiv verdi</i>	<i>N = Negativ</i>
<i>NPV = Negativ prediktiv verdi</i>	<i>FP = Falsk positiv</i>
<i>LR+ =Positiv likelihood-ratio</i>	<i>FN = Falsk negativ</i>
<i>LR- = Negativ likelihood-ratio</i>	<i>a = antall positive/negative</i>
	<i>n = total antall</i>

1.2.7 Tidligere forskning på området

Det ble utført et søk i PubMed den 05.07.2012 for å få en oversikt over hvilke studier som finnes med tanke på å undersøke de kliniske impingementtestenes diagnostiske validitet. Dette søket gav 22 treff ved søkestrengen ("Shoulder Impingement Syndrome/ diagnosis" [MAJR] AND "test" OR "tests" OR "physical examination" AND "diagnostic accuracy"). Det ble funnet fem særlig relevante studier på bakgrunn av innhold i abstractet og dette var studiene til Calis et al. (2000), Park et al. (2005), Silva et al. (2008), Michener et al. (2009) og Kelly et al. (2010). Disse artiklene tar eksplisitt for seg den diagnostiske validiteten ved kliniske impingement-tester, ved at de kartlegger samvariasjon mellom ulike kliniske tester i henhold til MR, Ultralyd, røntgen eller artroskopi. I tillegg ønsket jeg å ta med en metaanalyse fra 2012 som vurderte den diagnostiske nøyaktigheten av kliniske tester for subacromialt impingementsyndrom mot artroskopi eller kirurgi (Alqunae et al 2012). Det ble også gjort søk i andre databaser som EMBASE, Cinahl, Cochrane Library, Swemed+ og PEDro. Her ble det i tillegg funnet studier som undersøker diagnostisk validitet for tester som brukes for å kartlegge patologi i strukturer som kan være affisert ved impingement. Disse strukturene, som for eksempel bursa eller rotatorcuff, kan imidlertid være skadet også uten at det foreligger impingement. Derfor ble disse testene ekskludert. Resultater fra de seks særlig relevante studiene er oppsummert i Tabell 3.

Tabell 3: Studier av diagnostisk validitet av kliniske tester brukt for å avdekke subacromialt impingement ved skulderplager. Testresultatene er vurdert opp mot diagnose på grunnlag av funn på MR, røntgen, artroskopi og/eller ultralyd.

Studie	Test	S	SP	PPV	NPV	Nøyaktighet	LR+	LR-
Silva et al. 2008 <i>n</i> = 30 Referanse: MR	Neer	68.4	30	65	33.3	55.2	0.98*	1.05*
	Hawkins	73.7	40	70	44.4	62.1	1.23*	0.66*
	Yocum	79	40	71.4	50	65.5	1.32*	0.53*
	Jobe	73.7	30	66.7	37.5	58.6	1.05*	0.88*
	Patte	57.9	60	73.3	42.9	58.6	1.45*	0.70*
	Gerber	68.4	50	72.2	45.5	62.1	1.37*	0.63*
	Passiv abduksjon	73.7	10	60.9	16.7	51.7	0.82*	2.63*
	Abduksjon mot motstand	57.9	20	57.9	20	44.8	0.72*	2.1*
Calis et al. 2000 <i>n</i> = 120 Referanse: Røntgen og MR	Hawkins	92.1	25	75.2	56.2	72,8	1.23*	0.32*
	Neer	88.7	30.5	75.9	52.3	72	1.28*	0.37*
	Horisontal adduksjon	82	27.7	73.7	38.4	66.4	1.13*	0.65*
	Speeds test	68.5	55.5	79.2	41.6	64.8	1.54*	0.57*
	Yergasons test	37	86.1	86.8	35.6	51.2	2.66*	0.73*
	Painful arc	32.5	80.5	80.5	32.5	46.4	1.67*	0.84*
	Drop arm	7.8	97.2	87.5	29.9	33.6	2.78*	0.96*
Kelly et al. 2010 <i>n</i> = 59 Referanse: Ultralyd	Neer	62.1	0			54.5	0.62*	+∞*
	Hawkins	74.1	50			71	1.48*	0.52*
	Painful arc	29.6	50			31	0.60*	1.41*
	Svakhet ved abduksjon	37.9	50			39.4	0.76*	1.24*
	Smerte ved abduksjon	55.2	75			57.6	2.21*	0.59*
	Svakhet, utadrotasjon	55.2	25			51.5	0.74*	1.79*
	Smerte, utadrotasjon	34.5	100			42.4	+∞*	0.66*
	Svakhet, empty can	51.9	66.7			53.3	1.56*	0.72*
	Smerte, empty can	51.9	33.3			50	0.78*	1.45*
	Svakhet, full can	44.8	75			48.5	1.79*	0.74*
Smerte, full can	34.5	25			33.3	0.46*	2.62*	

Park et al. 2005 <i>n</i> = 913 Referanse: Artroskopi	Neer	68	68.7	80.4	53.2		2.17*	0.46*
	Hawkins	71.5	66.3	79.7	55.7		2.12*	0.43*
	Painful arc	73.5	81.1	88.2	61.5		3.89*	0.33*
	Jobe/Supraspinatus muskeltest	44.1	89.5	88.4	46.8		4.2*	0.63*
	Speed test	38.3	83.3	80.5	42.9		2.30*	0.74*
	Cross-body adduksjonstest	22.5	82	69.3	36.9		1.25*	0.95*
	Drop arm-test	26.9	88.4	81	39.7		2.32*	0.83*
	Infraspinatus muskeltest	41.6	90.1	90.6	45.8		4.2*	0.65*
Michener et al. 2009 <i>n</i> = 55 Referanse: Artroskopi	Hawkins	63	62			62	1.63	0.61
	Neer	81	54			68	1.76	0.35
	Painful arc	75	67			71	2.25	0.38
	Jobe (Empty can)	50	87			69	3.9	0.57
	Utadrotasjon mot motstand	56	87			72	4.39	0.5
Alqunae et al 2012 <i>n</i> = 1684 Referanse: Artroskopi	Hawkins	74	57				1.70	0.46
	Neer	78	58				1.86	0.37
	Jobe (Empty can)	69	62				1.81	0.50
	Drop arm test	21	92				2.62	0.86
	Lift-off test	42	97				16.47	0.59

* = regnet ut på bakgrunn av oppgitt sensitivitet (*S*) og spesifisitet (*SP*)

Referanse = en antatt undersøkelse som måler tilstanden.

S = Sensitivitet

SP = spesifisitet

PPV = Positiv prediktiv verdi

NPV = Negativ prediktiv verdi

LR+ = Positiv likelihood-ratio

LR- = Negativ likelihood-ratio

Silva et al. (2008) evaluerte den diagnostiske nøyaktigheten til 8 skuldertester på 30 personer med subacromialt impingementsyndrom og/eller subacromial bursitt bekreftet ved MR.

Konklusjonen var at de fleste kliniske testene for å identifisere subacromialt impingementsyndrom var diagnostisk nøyaktige, men de har likevel generelt lav spesifisitet

(se Tabell 3). Yocum var den testen som hadde de beste verdiene med henblikk på sensitivitet og presisjon.

En annen studie, Calis et al. (2000) kartla den diagnostiske nøyaktigheten til 7 skuldertester på 130 personer med subacromialt impingementsyndrom opp mot røntgen og MR. Den mest sensitive diagnostiske testen var Hawkins (92,1%), Neer test (88,7%) og Horisontal adduksjon (82%). Testene med høyest spesifisitet var Drop arm test (97,2%), Yergason test (86,1%) og Painful arc (80,5) (Calis et al 2000).

Kelly et al. (2010) evaluerte også den diagnostiske nøyaktigheten til 7 skuldertester på 34 personer med subacromialt impingementsyndrom bekreftet ved diagnostisk ultralyd. Den diagnostiske nøyaktigheten varierte fra test til test. Hawkins test var den mest diagnostisk nøyaktige for å kartlegge subacromialt impingement syndrom (71%). De mest diagnostisk nøyaktige for subkategorier av SIS var smerte ved motstand mot utadrotasjon og svakhet ved Jobe test (63,6%) ved veske subdeltoid, smerte ved motstand mot utadrotasjon ved partiel ruptur (58,8%) og Painful arc ved total ruptur (62,1%).

Park et al. (2005) evaluerte den diagnostiske nøyaktigheten til 8 skuldertester hos deltagere med SIS bekreftet ved artroskopi. Som vi ser i tabell 3 varierte verdiene betydelig i studien. Hawkins, Painful arc og Infraspinatus muskeltest innbragte best sannsynlighet ved posttest av impingementsyndrom (95%). Kombinasjonen av Painful arc, Drop-arm test og Infraspinatus muskel test utgjorde den beste sannsynligheten ved posttest for total rotator cuff ruptur.

Michener et al. (2009) så på den diagnostiske nøyaktigheten og reliabiliteten av individuelle og kliniske tester for SIS opp mot artroskopi. Konklusjonen av studien som vist fra verdiene i tabell 3, var at Painful arc og Neer test er gode for å utelukke SIS. Painful arc er også god til å bekrefte SIS sammen med Jobe og Utadrotasjon mot motstand.

Det nyeste studieutført av Alqunae et al. (2012) er en metaanalyse av kliniske tester for SIS validert opp mot artroskopi eller kirurgi. Ti studier med 1684 pasienter ble inkludert i metaanalysen. Hawkins test, Neers test og Jobe viste seg å være bedre til å utelukke enn å bekrefte SIS med bedre sensitivitetsscore (69%, 78%) enn spesifisitet (59%, 62%). En negativ Neers test reduserer sjansen for SIS fra 45% til 14%.

Ut fra tallene i tabell 3 og fra foreliggende litteratur har jeg i samråd med to medforskere valgt å benytte de fire testene Hawkins, Neers, Painful arc og Jobe (Empty can test) i studien. Hawkins og Neers test beskrives i litteraturen som impingementtester (Magee, 2008, Solberg, 2002, Hattam & Steatham, 2010, Bahr & Mæhlum, 2010). I tillegg kom i følge Kelly et al. (2010) og Calis et al. (2000) Hawkins test ut som den mest diagnostisk nøyaktige for å kartlegge SIS. Den diagnostiske nøyaktigheten varierte likevel fra test til test i alle studiene. Likhetsratioene på Jobe test og Painful arc er noe sprikende. Painful arc og Jobe viste likevel gode resultater i flere studier (Calis et al. 2000; Michener et al. 2009; Park et al. 2005; Kelly et al. 2010; Alqunae et al. 2012). Til tross for sprikende data presenteres både Jobe test (Gross et al. 1996) og painful arc (Bahr & Mæhlum, 2003) i litteratur som impingementtester. Vi har spesielt vektlagt studien til Park et al. (2005) og Alqunae et al (2012) i utvelgelsen av de fire testene. Dette grunnet den store utvalgsstørrelse på 913 og 1684. I studien til Park et al. (2005) kommer også de 4 valgte tester totalt best ut med høy positiv LR og lav negativ LR.

2.0 HENSIKT OG PROBLEMSTILLING

2.1 Hensikt

Hensikten med oppgaven var å undersøke den diagnostiske nøyaktigheten for Hawkins, Neers test, Painful arc og Jobe, da eksisterende studier ikke gir et entydig svar på dette. Dette er fire kliniske tester som benyttes i diagnostikk av skulderimpingement. Jeg vil i tillegg drøfte den diagnostiske validiteten av de samme fire testene opp mot resultater fra tidligere studier.

2.2 Problemstilling

”Hvor diagnostisk nøyaktig er fire kliniske provokasjonstester for subacromialt impingement ved bruk av ultralyd som referanseramme?”

2.2.1 Begrepsspesifisering

Subacromialt impingement opptrer i det subacromiale rommet når den subacromiale bursa og rotatorcuffen avklemmes mellom caput humeri og acromion eller mot det coracoacromiale ligamentet (Kelly et al. 2010).

Referanseramme defineres i denne oppgaven som en antatt gullstandard som fanger opp tilstanden. I denne studien vil dette være ultralyd.

3.0 METODE

3.1 Valg av forskningsdesign

Denne studien er en tverrsnittsstudie av diagnostisk nøyaktighet på deltagere som har vært til ultralydscanning av skulderen. Dette gir mulighet for statistisk beskrivelse av populasjonen utvalget er hentet utfra, generaliserbarhet (Ringdal, 2009). Studien kan videre beskrives som et statisk design og er egnet til å kartlegge grad av samvariasjon (Halvorsen, 1996). Dette samsvarer med studiens hensikt som vil være å evaluere den diagnostiske nøyaktigheten til fire skuldertester på personer med subacromialt impingementsyndrom bekreftet ved diagnostisk ultralyd. Studien er et kvantitativt design, hvilket innebærer at dataene kan uttrykkes i tall eller andre mengdetermer (Ringdal, 2009).

3.2 Utvalg

I studien brukes tre kliniske utvalg; Et utvalg består av pasienter med diagnostisert impingement på ultralyd, et utvalg med pasienter som har annen skulderproblematikk diagnostisert med ultralyd og et utvalg av personer uten skuldersmerter. Rekrutteringen til utvalgene foregikk på et stort fysikalsk instiutt og hos en lege som er spesialist i ultralyddiagnostikk av skulder. Legen har undersøkt og diagnostisert deltagerne med positiv skulderimpingement og deltagerne som har annen skulderpatologi til studien. Han er i tillegg allmennlege og idrettslege og har de siste 10 årene hatt ca. 2000 pasienter årlig til skulderdiagnostikk ved ultralyd. Deltagerne som ikke hadde skuldersmerter ble tilfeldig rekruttert fra det fysikalske instituttet og treningscenteret hvor testerne jobber til daglig. Definisjonen av friske vil i denne studien derfor være fravær av skuldersmerter. Det vil dermed si at deltagerne eventuelt kunne ha plager/smerter fra andre steder på kroppen. Deltagerne ble invitert til studien hvis de inngikk under følgende definerte inklusjons- og eksklusjonskriterier:

Gruppe 1: Impingement	
Inklusjonskriterier:	Eksklusjonskriterier:
<ul style="list-style-type: none"> • Diagnostisert med impingement ved ultralyd • Skuldersmerter med varighet mer enn én måned • Skuldersmertene må være pasientens aktuelle hovedproblem • Personer i alderen 18-60 år • Må kunne gi et informert samtykke 	<ul style="list-style-type: none"> • Ved annen kjent patologi i skulder • Tidligere traume mot skulder • Tidligere skulderoperert • Inflammatoriske leddlidelser • Smerter ved bevegelse av nakken • Nevropatiske smerter i nakke eller overekstremitet • Røde flagg som kan mistenke annen alvorlig patologi

Gruppe 2: Annen skulderpatologi	
Inklusjonskriterier:	Eksklusjonskriterier:
<ul style="list-style-type: none"> • Diagnostisert med annen skulderpatologi ved ultralyd • Skuldersmerter med varighet mer enn én måned • Skuldersmertene må være pasientens aktuelle hovedproblem • Personer i alderen 18-60 år • Må kunne gi et informert samtykke 	<ul style="list-style-type: none"> • Skulderimpingement • Tidligere traume mot skulder • Tidligere skulderoperert • Inflammatoriske leddlidelser • Smerter ved bevegelse av nakken • Nevropatiske smerter i nakke eller overekstremitet • Røde flagg som kan mistenke annen alvorlig patologi

Gruppe 3: Friske	
Inklusjonskriterier:	Eksklusjonskriterier:
<ul style="list-style-type: none"> • Ingen smerter i skulderområdet siste 3 måneder. • Personer i alderen 18-60 år • Må kunne gi et informert samtykke 	<ul style="list-style-type: none"> • Smerter i skulderområdet siste 3 måneder. • Tidligere traume mot skulder • Tidligere skulderoperert • Inflammatoriske leddlidelser • Smerter ved bevegelse av nakken • Nevropatiske smerter i nakke eller overekstremitet • Røde flagg som kan mistenke annen alvorlig patologi

Vi ønsket å inkludere menn og kvinner i aldersgruppen 18 - 60 år som snakker og forstår norsk, da all informasjon og testing foregikk på norsk. Ved å sette maxsgrensen på 60 år unngikk vi eldre pasienter som kunne ha annen komorbiditet, da dette øker med alderen. Vi ønsket ikke å ha medisinbruk som et eksklusjonskriterium, da pasienter med smerter ofte

bruker smertestillende. Det var likevel viktig å registrere medisinbruken under studien. I henhold til retningslinjene fra helsedirektoratet beskrives røde flagg som for eksempel unormalt stort vekttap de siste måneder, konstante smerter, infeksjon etc. (Lærum et al, 2007).

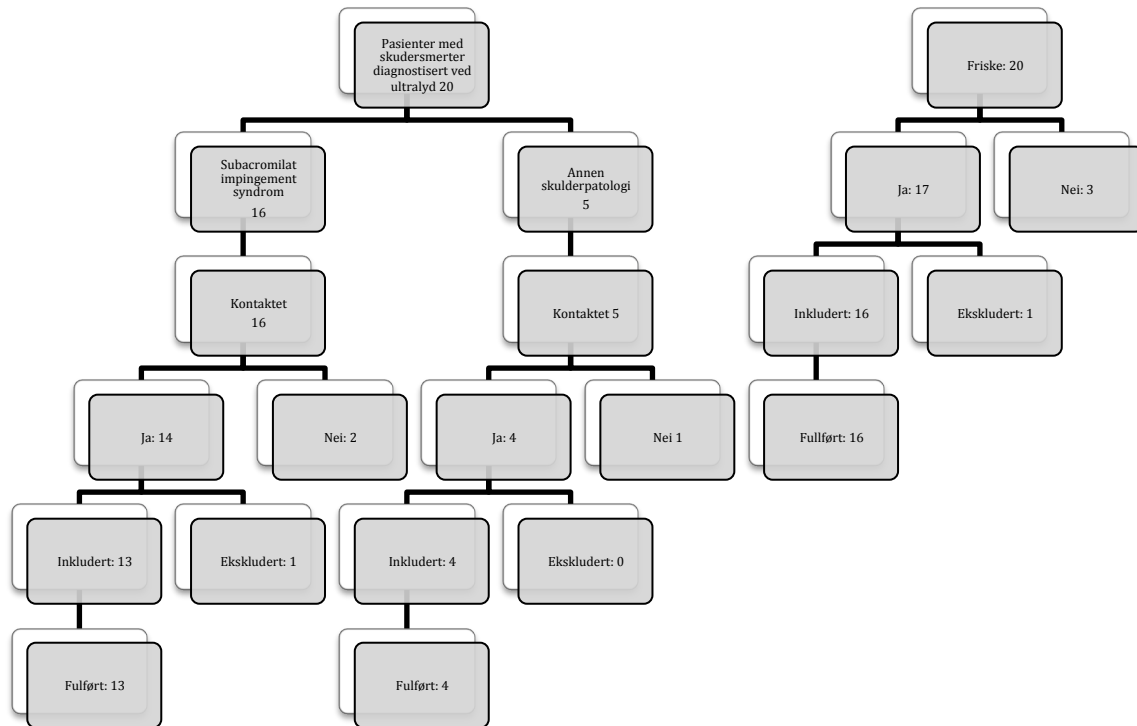
Ved diagnostisering av impingement med ultralyd ble følgende vurderinger lagt til grunn:

- Grad av depresjon av humerus, så vel som sprett eller tannhjulsbevegelse av rotatorcuffsener eller gradvis utvidelse av bursa mens pasientens skulder ble ledet i aktiv fleksjon og abduksjon
- Skader og rifter samt avrivning av rotatorcuff
- Skade på bursa

3.3 Utvelgelsesprosessen

Ut fra de 20 deltagerne som ble diagnostisert ved ultralyd ble 16 stykker diagnostisert med impingement og fem med annen skulderpatologi. Disse ble diagnostisert som irritasjon anterolateralt eller frozen shoulder. Til sammen takket tre nei, da de allerede hadde begynt i et behandlingsforløp. I impingementgruppen ble en ekskludert, da deltageren ikke oppfylte inklusjons og eksklusjonskriteriene likevel.

De første 20 deltagerne som ikke hadde skuldersmerter ble tilfeldig valgt ut. Av disse gjennomførte 16, tre takket nei og en ble ekskludert da deltageren ikke innfridde inklusjonskriteriene (Figur 3).



Figur 3: Utvelgelsesprosessen til forsøket

I tillegg til vurdering av utvalgets sammensetning må utvalgenes størrelse defineres. Det var viktig at så mange som mulig deltok i studien for at resultatene skulle bli så valide som mulig og representere hele populasjonen (Lund, 2002). Terwee et al. (2007) anbefaler minimum 50 pasienter for å få tilstrekkelig sikkerhet og statistisk power i estimatene. På bakgrunn av dette ville det vært hensiktsmessig med minimum 50 deltagere i denne studien. Dessverre lot det seg ikke gjøre å rekrutere så mange som ønsket. Det ville vært lavere usikkerhet forbundet med punkttestimatene og konfidensintervallet ville blitt smalere hvis utvalget hadde vært større (Straus et al. 2005).

3.4 Prosedyre og datainnsamling

Min studie er en del av en tredelt studie som vil gjøres i samarbeide med to andre medforskere (fysioterapeuter/manuellterapeutstudenter). Disse vil bidra til å gjennomføre de fire kliniske testene for subacromialt impingementsyndrom og vil henholdsvis undersøke intertester og test-retest reliabiliteten av disse testene. Selve invitasjonen for å delta i studien, det vil si rekrutteringsprosessen, administreres av en uavhengig person som vil gi informasjon og forespørre personene om deltakelse i studien (Vedlegg 1).

Datainnsamlingen foregikk på et stort fysikalsk institutt. Deltagerne i studien ble satt sammen i tilfeldig rekkefølge slik at fysioterapeutene som utførte de kliniske testene (Tester 1 og 2) var blindet for hvem som hadde eller ikke hadde diagnosen skulder impingement. Det ble gjort kvantitative målinger på impingementtestene i skulderen, og det ble registrert om den gitte testen fremprovoserer smerte i skulderen eller ikke. Datainnsamlingen foregikk i løpet av sommeren 2012 med analyse av dataene og skriving av masteroppgaver høsten 2012. Tidsfristen for ferdig prosjekt var 1. november 2012. De kliniske testene som ble brukt var Neers test, Hawkins test, Jobses test og painful arc.

Moen et al. (2010) har gjort rede for hvordan Neers test, Hawkins test og Jobses test skal utføres og legges til grunn for testprosedyren. Painful arc er standardisert som beskrevet nedenfor.

Testprosedyren

Neers test

Pasienten i sittende, terapeuten står bak. Terapeuten fikserer scapula på samme side for å hindre protraksjon av skulder. Pasientens arm flekteres passivt i sagittalplanet. Positiv test hvis bevegelsen utløser smerter (Moen et al. 2010).



Hawkins test

Pasienten i sittende, med overarmen 90 grader flektert. Scapula på samme side fikseres for å hindre rotasjon av scapula. Overarmen roteres passivt innover til end-feel eller til eventuelle smerter oppstår (Solberg, 2002).



Jobes test

Pasienten i sittende, skulderen 90 grader abduisert, 30 grader horisontal addusert og maksimalt innadrotert. Terapeuten plasserer hånd på oversiden av pasientens overarm. Pasient skal holde overarmen i ro mens terapeut appliserer et trykk nedover mot pasientens arm. Positiv test hvis smerter (Moen et al. 2010).



Painful arc

Pasienten i sittende. Park et al. (2005) instruerer pasienten til å føre armen i fleksjon i scapulas plan/scaption. Testen er positiv hvis pasienten opplever smerter mellom 60 og 120 grader abduksjon (Kessel & Watson 1977, Bahr og Mæhlum 2003, Park et al. 2005, Michener et al. 2009, Kelly et al. 2010).



Skåring av testene

Testene ble skåret med utgangspunkt i dikotomien positiv eller negativ test. Dette ble gjort fordi alle de relevante studiene (Calis et al. 2000, Park et al. 2005, Silva et al. 2008, Michener et al. 2009, Kelly et al. 2010 og Alqunaee et al. 2012) skårte smerteprovoserende tester basert på dikotomien smerter/ikke smerter. I praksis kan det likevel oppstå ulik grad og lokalisering av smerter og ubehag ved faktisk test-gjennomføring. Generelt gjelder at hvor en velger å legge avkuttingspunktet for positiv test vil være av betydning for sensitiviteten, da sensitiviteten øker med høyere avkuttingspunkt (Jamtvedt et al. 2003). Testene skåres som positiv hvis den primært gir kjente smerter i skulderområdet. Pasienten må oppleve relativt distinkte smerter i skulder for at testen skal skåres som positiv, andre opplevelser som tøyning, strekk, diffuse ubehag og lignede skåres som negativ test.

Standardisering

Standardiseringen foregikk ved at terapeutene leste gjennom prosedyrene som grunn for testutførelse. Annen standardisering mellom terapeutene ble ikke gjennomført, for i større grad søke å bevare den økologiske validiteten. Med økologisk validitet menes en generalisering hvor en ønsker at de eksperimentelle resultatene skal gjelde i det virkelige liv

(Kruuse, 2007), og et studiedesign hvor terapeuter har en omfattende trening sammen i forkant av testingen, kan medføre at standardiseringen blir mer omfattende enn hva som er sannsynlig i klinisk praksis. Terapeutene som gjennomførte studien er fysioterapeuter og i gang med et klinisk masterstudium i manuellterapi. De har relativt lik yrkesbakgrunn og klinisk erfaring, ca. 5 år, innenfor muskel/skjelett-problematikk.

Faktisk utførelse

Når testene ble utført var det kun den ene testeren og deltageren som var inne i testlokalet. Hver test ble utført med to minutters mellomrom først en gang av terapeut nr. 1. Deretter ble deltageren testet på nytt av terapeut nr. 2 etter tre minutters pause. Det ble valgt to minutters pause mellom hver test og tre minutter mellom hver av de to terapeutene for å hindre at den første testen påvirker svarene ved test nummer to. Terapeutene alternerte i å teste pasientene først og sist og deltagerne trakk testrekkefølgen for å unngå systematiske feil.

Skåringsdataene ble registrert og anonymisert (Vedlegg 1) i en database. Identifiserbare person-opplysninger samt sykdomsstatus og hvilke av utvalgene pasienten tilhører ble nedtegnet og opplysningene sammen med koblingsnøkkel ble oppbevart på et sikkert og innelåst sted.

3.5 Analyse av datamaterialet

Etter endt studie ble resultatene analysert. Datamaterialet ble analysert ved ”Statistical Package for the Social Sciences” (SPSS), versjon 19 for mac. Skåringen ble deretter satt inn i en 2 x 2 tabell. Det ble både sett på verdiene mellom SIS og annen skulderpatologi, samt SIS og friske. Utfra tabellene ble testenes diagnostiske validitet beregnet ved at testenens sensitivitet, spesifisitet, negativ- og positiv likelihood-ratio, nøyaktighet, samt positiv- og negativ prediktiv verdi ble utregnet. Grunnet begrensningen av denne masterstudien ble kun terapeut 1 sine verdier brukt til utregning av den diagnostiske validiteten.

3.6 Etiske aspekter

Deltagerne skulle innen studiens start underskrive et informert samtykke om prosjektets formål, forløp, taushetsplikt, anonymisering og forpliktelsene (Vedlegg 2). Det er frivillig å

delta, og deltageren kunne derfor til enhver tid trekke seg uten konsekvenser. En uavhengig person spurte aktuelle kandidater om deltagelse for å ikke påvirke et avhengighetsforhold, samt at deltagerne ikke følte seg forpliktet til å delta hvis de ikke hadde lyst. Deltagerne ble også informert om konsekvenser og eventuelle risikoer forbundet med intervensjonen (Vedlegg 2).

I følge Dalland (2007) er det viktig å tenke igjennom hvem som kan ha nytte av prosjektet og om det kan bidra på samfunnsnivå. Formålet med dette prosjektet er å bidra med supplerende forskning på området. Det antas at studien vil ha stor betydning da det er essensielt med gode standardiserte tester for å kunne behandle pasienten mest mulig korrekt. Det er som tidligere skrevet mange, opp mot 50 %, som har hatt en skulderlidelse sist år (Brox et al 2011). Samfunnsmessig vil studien også være av betydning da denne pasientgruppe koster samfunnet betydelig i direkte kostnader. I tillegg er det ingen samfunnsmessige omkostninger forbundet med studien. Det antas at studien kan bidra med ny viten og optimering av impingementtestene. På bakgrunn av dette sikres intervensjonsgruppen mot unødvendige belastninger uten relevant formål. Ut fra dette vurderes studien til å ha en verdi for den enkelte pasient med skuldersmerter, fysioterapeut/manuellterapeut og samfunnet generelt.

Regionale Etske Komite (REK) skal godkjenne hvert studie som involverer personer og persondata hvor forskningsprosjektet ikke er av en slik art at det regnes som en del av vanlig etablert behandlingsprosedyre. REKs hovedoppgave er å vurdere om medisinsk og beslektede forskningsprosjekt på mennesker oppfyller Helsinki-deklarasjonen og andre relevante etiske normsett for forskning (Ruyter et al., 2000). Dette studiet ble derfor godkjent av REK –Vest våren 2012 (Vedlegg 3).

Generelle etiske overveielser følger reglene og anbefalingene som er utarbeidet av Helsinkideklarasjonen (Ruyter et al. 2000). I deklarasjonen legges det vekt på at hensynet til forsøkspersonenes helse og integritet må gå foran hensynet til forskning og samfunn (Dalland, 2007). Dette innebærer at data vil bli aidentifisert og oppbevart innelåst hvor kun forskerne har tilgang. I tillegg kan deltagerne trekke seg når som helt uten å oppgi grunn.

4.0 Resultater

4.1 Baselinedata

Baselinedata for deltagerne i studien er presentert i tabell 4. Kjønn og aktuell side testet var jevnt fordelt mellom alle deltagerne, ved at det ble testet 19 kvinner og 14 menn hvorav 18 testet høyre skulder og 15 testet venstre skulder. Gruppene var relativt homogene når det gjelder gjennomsnitt i alder ved henholdsvis 47, 44 og 40 år. Gjennomsnittet for varighet av smerte i SIS-gruppen ble høy (43 mnd) grunnet en enkelt deltager som hadde hatt smerter i 3 år. Medianen (12 mnd) gir derfor i dette tilfellet en mer nøyaktig måling. Innenfor tidsrammen var det ikke mulig å rekruttere flere deltagere til gruppen annen skulderpatologi. Baselinedata registrert for hver enkelt deltager er presentert i vedlegg 4.

Tabell 4: Alder, kjønn, og sykdomsvarighet hos friske deltagere (n=16) og deltagere diagnostisert via diagnostisk ultralyd som referanse ramme (n=17).

	Gruppe 1/ Impingement	Gruppe 2/Annen skulderpatologi	Gruppe 3/Friske
Antall (% av total)	13/33 (39)	4/33 (12)	16/33 (48)
Kjønn (% av total)			
Kvinner	6 (32)	2 (11)	11 (57)
Menn	7 (50)	2 (14)	5 (36)
Alder (år)			
gjennomsnitt	47	44	40
(SD)	(13)	(9)	(16)
Variasjonsbredde	18 - 60	33-45	21-60
Sykdomsvarighet			
Gjennomsnitt, median	39*,	16,	-
(SD)	12	16	
Variasjonsbredde	(97)	(7)	
(mnd)	1-360	9-24	
Aktuell side testet			
(% av total)			
Høyre	3 (17)	4 (22)	11 (61)
Venstre	10 (67)	0 (0)	5 (33)

*Gjennomsnitt på 39 grunnet en deltagers verdi på 360 mnd

SD = Standardavvik

Mnd = Måneder

4.2 Krysstabeller

I tabell 5 vises en krysstabell for test positive, test negative, falske negative og falske positive for deltagerne i de tre gruppene ved de fire kliniske testene for impingement. Utfra tabellen illustreres at alle testene ga falske negative verdier i SIS gruppen. Hawkins test var den testen

som ga færrest falske negative ved kun 1 deltager og Painful Arc ga flest ved fire deltagere. I gruppen med annen skulderpatologi ga alle testene utenom Jobe test falske positive svar. Hawkins test ga alle deltagere falske positive resultater. I gruppen uten skuldersmerter testet alle negativt utenom Hawkins test hvor to deltagere av totalt 16 fikk falskt positivt svar. Tallene i tabellen legges til grunn for utregning av den diagnostiske nøyaktigheten av de fire kliniske skuldertestene for å diagnostisere impingement. Testscorene for hver enkelt deltager er presentert i vedlegg 4.

Tabell 5: test positive, test negative, falske negative og falske positive for ultralyd positive, ultralyd negative og friske deltakere

Klinisk test	Antall	Ultralyd positive Gruppe 1		Ultralyd negative# Gruppe 2		Friske Gruppe 3	
		Test positive	Falske negative	Falske positive	Test Negative	Falske positive	Test Negative
Neers	33*	10	3	3	1	0	16
Hawkins	33*	12	1	4	0	2	14
Jobe	33*	10	3	0	4	0	16
Painful arc	33*	9	4	2	2	0	16

*13 i gruppe 1: Impingement, 4 i gruppe 2: Annen skulderpatologi og 16 i gruppe 3: Friske

Ultralyd negative = gruppen annen skulderpatologi som ble diagnostisert negative for diagnosen SIS ved ultralyd.

Tabell 6 og 7 viser sensitivitet, spesifisitet, positiv prediktiv verdi, negativ prediktiv verdi, positiv likelihood ratio, negativ likelihood ratio og nøyaktigheten for testene Neers, Hawkins, Jobe og Painful Arc. Verdiene er regnet ut på bakgrunn av fordelingen vist i tabell 5. I tabell 6 er SIS gruppen sammenliknet med gruppen annen skulderpatologi og i tabell 7 er SIS gruppen sammenliknet med gruppen friske deltagere. Tabell 6 illustrerer at testene Neers, Hawkins og Painful Arc har høyere sensitivitet enn spesifisitet. Jobe test har derimot høyere spesifisitet enn sensitivitet. I tillegg har Jobe test både den høyeste positive prediktive verdien og negative prediktive verdien. Jobe test ville videre hatt en høy positiv LR, men denne kan ikke utregnes grunnet den gode og høye spesifisiteten (1.0). I motsetning kan den negative LR verdien til Hawkinstest ikke utregnes grunnet en spesifisitet på 0. Jobe test har den høyeste nøyaktigheten (0.82).

Tabell 6: Gruppe 1/Impingement sammenliknet med gruppe 2/Annen skulderpatologi

Test	Sensitivitet (KI)	Spesifisitet (KI)	PPV	NPV	+LR	-LR	Nøyaktighet
Neers	0.77 (0.54, 1.0)	0.25 (0.0, 0.67)	0.77	0.25	1.03	0.92	0.64
Hawkins	0.92 (0.78, 1.0)	0	0.75	0	0.92	-	0.71
Jobe	0.77 (0.54, 1.0)	1.00 (-)	1.00	0.57	-	0.23	0.82
Painful Arc	0.69 (0.44, 0.94)	0.50 (0.01, 0.99)	0.81	0.33	1.40	0.60	0.64

PPV = Positiv prediktiv verdi

NPV = Negativ prediktiv verdi

LR+ = Positiv likelihood-ratio

LR - = Negativ likelihood-ratio

KI = Konfidensintervall

Tabell 7 viser at Hawkins test har fått høyest sensitivitet og at denne er høyere enn spesifisiteten. Testene Neers, Jobe og Painful Arc har fått høyest spesifisitet og PPV (1.0). Den meget høye spesifisiteten gjør at konfidensintervallet eller den positive LR ikke kan regnes ut. Hawkins test har fått den laveste negative LR (0.09) og den høyest NPV (0.93). Både testene Neers, Hawkins og Jobe har den høyeste verdien av nøyaktighet (0.90).

Tabell 7: Gruppe 1/Impingement sammenliknet med gruppe 3/friske

Test	Sensitivitet	Spesifisitet	PPV	NPV	+LR	-LR	Nøyaktighet
Neers	0.77 (0.54, 1.0)	1.00 (-)	1.00	0.84	-	0.23	0.90
Hawkins	0.92 (0.78, 1.0)	0.88 (0.71, 1.00)	0.86	0.93	7.67	0.09	0.90
Jobe	0.77 (0.54, 1.0)	1.00 (-)	1.00	0.84	-	0.23	0.90
Painful Arc	0.69 (0.44, 0.94)	1.00 (-)	1.00	0.80	-	0.31	0.86

PPV = Positiv prediktiv verdi

NPV = Negativ prediktiv verdi

LR+ = Positiv likelihood-ratio

LR - = Negativ likelihood-ratio

KI = Konfidensintervall

5.0 Diskusjon

For effektiv behandling og prognose av pasienter med skulderpatologi er det nødvendig med tester som skiller mellom syke og friske og mellom ulike diagnosegrupper. I denne studien ble den diagnostiske validiteten av fire kliniske tester for SIS med ultralyd som referanseramme vurdert. Resultatene er presentert i tabell 6 og 7. Neers, Hawkins og Painful

Arc har høyere sensitivitet enn spesifisitet ved sammenlikning av SIS og annen skulderpatologi. Dette indikerer at sjansen for å ha SIS ved negativ test er liten. Hawkins test er også positiv på alle med annen skulderpatologi. Dette medfører at Hawkins test i denne studien ikke differensierte skulderpatologi, hvilket resulterte i 0 % spesifisitet. I motsetning til de andre testene har Jobe test høyere spesifisitet enn sensitivitet, hvilket indikerer at sjansen for SIS er meget stor ved positiv test. Jobe test har også den høyeste nøyaktigheten (0.82). Ved sammenlikning av SIS og friske hadde Neers, Jobe og Painful Arc tester høyere spesifisitet enn sensitivitet. Dette indikerer at sjansen for å ha SIS er stor ved positiv test sammenliknet med de som er friske. Hawkins test hadde i motsetning til de andre tre testene høyere sensitivitet enn spesifisitet, og sjansen for SIS var dermed liten hvis testen var negativ. Hawkins test hadde en negativ sannsynlighetsandel (-LR) på under 0,1 (0,09). Dette indikerer at sjansen for å ha negativ test for SIS vil være ca. en tiendedel av hva den er for de som er friske. Med andre ord, vil det for hver tiende negative være en falsk negativ. På bakgrunn av dette er derfor Hawkins test den mest diagnostisk representative testen. Den positive LR indikerte derimot at det var nesten 8 ganger større sannsynlighet for å få positiv test når man har SIS enn når man er frisk. I den friske gruppen var det ingen deltagere som var falske positive på testene Neers, Jobe og Painful Arc. Dette gjør at de fikk en spesifisitet på hele 1.0 eller 100% og det var dermed ikke mulig å regne ut den positive LR grunnet dette. Dersom en deltager hadde vært falsk positiv ville den positive LR for Neer og Jobe vært hele 12.83, noe som ville tilsagt en diagnostisk representativ test.

I de følgende avsnitt vil jeg først diskutere resultatene i studien mot tidligere studier, for deretter å diskutere faktorer som kan påvirke resultatet (konfoundere), samt styrker og svakheter ved denne studien. Til slutt vil jeg drøfte den eksterne validiteten.

5.1 Tidligere studier

Resultatene i denne studien viste at både Neers, Hawkins og Painful Arc tester hadde høyere sensitivitet enn spesifisitet ved sammenlikning av SIS og annen skulderpatologi, hvilket er i samsvar med Silva et al. (2008), Michener et al. (2009) og Alqunae et al. (2012) (Tabel 3). Ved Neers- og Hawkinstest viste Calis et al. (2000), Kelly et al. (2010) og Park et al (2005) også dette, men i motsetning viste Painful Arc høyere spesifisitet enn sensibilitet. Dette kan forklares ved standardiseringen av Painful Arc. I for eksempel Kelly et al. (2010) var deltagerne stående og utførelsen var i abduksjon. I denne studien sitter deltageren og

utførelsen er i scapulas plan/scaption. Hawkins test var i denne studien positiv på alle deltagerne med annen skulderpatologi. I de tidligere nevnte studiene var ikke dette tilfellet, hvilket kan forklares med flere deltagere i gruppen annen skulderpatologi. I motsetning til de andre testene har Jobe test høyere spesifisitet enn sensitivitet i denne studien. Dette er i samsvar med tidligere studier som Park et al. (2005) og Michener et al. (2009), men i motsetning til Silva et al. (2000) og Kelly et al. (2010). Årsaken til dette kan igjen kanskje skyldes antall deltagere eller utførelse av testen. Michener et al. (2009) og Park et al. (2005) brukte samme standardisering av testen. Likevel kan det være individuelle forskjeller i for eksempel hvor stor motstand terapeuten gir. Sammenliknet med de tidligere nevnte studiene er nøyaktigheten omtrent lik hvis man ser på alle testene samlet. Denne studien illustrerer at Jobe test har høyest nøyaktighet (0.82). Det har derimot ikke noen av de tidligere nevnte studiene. Dette kan igjen skyldes få antall deltagere og dermed kunstig gode tall. I motsetning kan det skyldes andre faktorer som for eksempel at vi har kort tid i mellom diagnostisering ved ultralyd og utførelsen av testene, noe som styrker studien. I tillegg skåres Jobe test forskjellig i studien til Kelly et al. (2010) hvor positiv test også representerte svakhet fremfor smerte. I Kelly et al. (2010), Silva et al. (2008) og Calis et al. (2000) har Hawkins test høyest nøyaktighet, og nøyaktigheten samsvarer helt (0.71) med denne studien ved Kelly et al. (2010) og Calis et al. (2000). Grunnen til dette kan være lik standardisering av testen. Kelly et al. (2010) brukte også ultralyd som referanseramme.

Tidligere studier og metaanalyser har vist varierende resultat når det gjelder diagnostisk nøyaktighet for Hawkins og Neer test (Calis et al. 2000, Park et al. 2005, Silva et al. 2008, Michener et al. 2009, Kelly et al. 2010, Alqunae et al 2012). Generelt gjelder det at denne studien samsvarer disse delvis når det gjelder Hawkins og Neer test. Painful Arc har vist bedre resultater i studiene til både Park et al. (2005) og Michener et al. (2009) med både høyere positiv LR og lavere negativ LR. Park et al. (2005) demonstrerte generelt bedre verdier bortsett fra på Jobe test. Noe av dette kan kanskje skyldes bruk av annen referanseramme. Calis et al. (2000) bekreftet diagnosen ved å sette en 10 ml lidocaine injeksjon i det subacromiale rommet. Videre har de patologiske tilstandene ved SIS variert fra studie til studie. Park et al. (2005) inndelte deltagerne i 3 grupper bestående av ingen rotatorcuff ruptur, partiel ruptur og total ruptur. Vi har i vår studie hatt deltagere med alle 3 stadiene samlet i SIS gruppen. Calis et al. (2000) valgte blant annet å ikke ta med total rotatorcuff ruptur. Dette kan forklare noe av forskjellen mellom studiene.

5.2 Konfoundere

Patologi

Hawkins test var falsk positiv på alle deltagerne med annen skulderpatologi og Neers test fikk 75 % positive. En forklaring kan være at deltagerne kan ha hatt en begynnende SIS som ikke var synlig ved ultralyd, da irritasjonen ikke var på et nivå for patologiske forandringer.

Gartsman (1995) så på rotatorcuffplager ved artroskopi og viste i sin studie at mange mennesker har smertefull impingement uten en inflammatorisk reaksjon. Dette kan også ha vært tilfelle ved de friske falske positive. På den andre siden har parietalartikulære skulderlesjoner vist seg å være symptomfrie hos noen pasienter (Milgrom et al. 1995). Dette kan reflektere den subjektive smerteskåren ved testene og eventuelt unøyaktig tolkning av smerte hos pasientene. I tillegg skåret deltagerne i vår studie negativt på Jobe test, hvilket tyder på at denne testen burde vært falsk positiv.

Tidsaspektet

Tiden mellom ultralydskanning og utførelsen av de kliniske testene varierte i denne studien fra en dag til to uker. Vi tilstrebet kortest mulig tid grunnet faren for patologiendringer, noe som er en styrke ved denne studien. I tidligere studier hvor tiden er representert varierer denne fra dager til måneder. I studien til Michener et al. (2009) er den for eksempel 2 måneder. Fremtidige studier, samt manuellterapeuter i klinisk praksis, burde minimalisere tiden mellom indeksen og referansetesten for å redusere faren for endringer av patologien ved SIS.

Blinding og standardisering

I denne studien var terapeutene som utførte testene blindet, hvilket er en styrke i denne studien. Blinding sikrer at registrering og fortolkning ikke påvirkes av subjektive antakelser (Lindbæk & Skovlund, 2002). Dette var hovedgrunnen til å også inndra friske i studien. Ingen av de tidligere studiene beskrevet har sammenliknet testene for SIS mot friske. I videre studier bør standardiseringen av utførelsen til de kliniske testene vurderes. I denne studien ble kun testprosedyren lest igjennom og hver test ble kun utført en gang for å bevare den økologiske validiteten ved at dette normalt gjøres i klinisk praksis (Svartdal, 2009). I noen tilfeller var deltagerne litt usikre på om det var kjent smerte ved testen eller ubehag grunnet strekk i ytterstilling. I disse tilfellene kunne det ha vært en fordel å for eksempel utføre testen to ganger. Videre kunne en gradering av scorene ha gitt et mer presist svar. På den andre siden er dette ikke normal klinisk praksis (Juel, 2007, Solberg, 2002, Magee, 2008, Moen,

2010). En begrensning i denne studien kan være tiden i mellom provokasjonstestene. I denne studien ble det valgt to minutters pause mellom hver test og tre minutter mellom hver av de to terapeutene for å hindre at den første testen påvirker svarene ved test nummer to. I klinisk praksis er prosedyren også å vente til eventuell smerte forvinner, men tiden vil derfor variere. For å unngå andre systematiske feil (Lund, 2002) altenerte terapeutene i å teste pasientene først og sist og deltagerne trakk testrekkefølgen, noe som er en styrke i studien. Hawkins test var positiv på alle med annen skulderpatologi og to friske. Likevel var testrekkefølgen på Hawkins test forkjellig, og det er derfor liten sannsynlighet for at den var falsk positiv grunnet testrekkefølgen.

Diagnostisering

Diagnostisering av skulderpatologi er en utfordring og kompleks prosess. Accident Compensation Corporations (2011) retningslinjer forslår at tester som Neers og Hawkins kan være nyttige som en del av undersøkelsen. Solberg (2002) hevder at skuldertestene bør ses i sammenheng med resten av undersøkelsen og at en test i seg selv har begrenset betydning. Dette bekreftes av Michener et al. (2009) som i sin studie konkluderte med at ved tre av fem positive SIS tester er sjansen for patologi stor og ved mindre enn tre av fem liten. Dette er i tråd med Park et al. (2005) sin studie som også konkluderte med at tre positive tester ga større sjanse for å indikere SIS. Som fysio – og manuellterapeut kan det være nyttig å klassifisere pasientene etter verdens helseorganisasjons (WHO) klassifikasjonssystem ”International Classification of Functioning” (ICF) (Üstün, 2002). En diagnostisk prosess kartlagt ved ICF inneholder en helhetlig vurdering med anamnese og et bredt undersøkelsesregister. I denne studien ble ikke alle potensielle historiske spørsmål, jobb eller fritidsaktiviteter undersøkt. Det etterlyses studier som tar for seg hele den diagnostiske undersøkelsen, da en test alene ikke er nok for å stille en diagnose.

En tidlig diagnostisering og kartlegging av pasienten er sentralt for optimalisering av tilstanden. Retningslinjene anbefaler gjenopptrening hos fysioterapeut ved subacromial bursitt og partiel ruptur av rotatorcuffen (Accident Compensation Corporation, 2011). Videre bør pasienter med totalruptur av rotatorcuffen eller omfattende patologiske skader i rotatorcuffen opereres. Det er spesielt viktig dersom pasienten er meget aktiv og ung (Accident Compensation Corporation, 2011).

Ultralyd

Valget om å bruke ultralyd som referanseramme for å diagnostisere SIS og annen skulderpatologi er skjønnsmessig og operatørvhengig. I denne studien kan det være en bias med henblikk på dette forsøkt redusert ved å kartlegge de diagnostiske kriteriene for positiv SIS, samt å velge en erfaren lege til å diagnostisere skulderpatologi med ultralyd.

Skuldresmerter kan begrense diagnostikken. Dette kan igjen begrense nøyaktigheten og utførelsen. Videre kan anatomiske variasjoner og andre patologiske forandringer kanskje influere på nøyaktigheten. I denne studien har dette blitt forsøkt redusert ved at deltagerne har blitt undersøkt med ultralyd i begge skuldre for sammenlikning.

Bruken av MR eller artroskopi kan muligvis være et bedre alternativ som referanseramme, da ultralyd er vurdert til å være mindre nøyaktig ved diagnostikk av rotatorcufflesjoner enn artroskopi eller åpen kirurgi. Ultralyd er likevel diagnostisk representativ sammenliknet med kirurgi (Naredo et al. 2002). Ultralyd har også fordelen ved færre kontraindikasjoner, ikke er inngrepene, rask og enkel å utføre, tillater dynamisk testing og er økonomisk gunstig (Naredo et al. 2002).

Smerte

Opplevelse av smerte er subjektiv og komplekst sammensatt. Derfor er det vanskelig å vite om eksterne faktorer spiller inn og eventuelt i hvor stor grad. I den utførte studien har jeg forsøkt å ta høyde for dette ved å registrere medisinførbruk og ved at deltagerne ikke skulle motta annen form for behandling mellom ultralydskanningen og utførelsen av de kliniske testene, da dette kunne påvirke smerten. Dette er også en del av den kliniske hverdagen, hvor vi registrer medisinførbruk og eventuell annen behandling.

Jensen et al. (2009) beskriver smerte som en multidimensjonal problemstilling som involverer flere biopsykososiale faktorer. Det subjektive aspekt ved smerte kan i flere tilfeller bli påvirket av personlig og økonomiske aspekter (Jensen et al. 2009). Det er blant deltagerne i denne studien ingen kjente verserende saker av denne karakter.

Smerteutbredelse og symptomer assosiert fra strukturer rundt skulderen, spesielt rundt det subacromiale rommet, kan være like (Juel, 2007). Noen pasienter kan ha en kombinasjon av flere smertegivende strukturer og dermed flere diagnoser som kan gjøre de vanskelige å diagnostisere.

Deltagere

Den diagnostiske nøyaktigheten varierer blant de tidligere beskrevne studiene (Calis et al. 2000, Park et al. 2005, Silva et al. 2008, Michener et al. 2009, Kelly et al. 2010, Alqunae et al 2012). Dette kan muligvis skyldes faktorer som alder, kjønn, jobb og fritidsaktiviteter. Det er i følge Harvey (2009) vist at prevalensen for SIS er forskjellig for aldersgrupper og kjønn, mens, Alqunae et al. (2012) påpekte at jobb og underforliggende faktorer kan påvirke den diagnostiske validiteten. I denne studien er det ut fra subanalysene små differanser når det gjelder kjønn, alder og aktuell side testet, og differansen vurderes derfor til å ikke ha klinisk relevans. Utfra resultatene er det ikke noe som indikerer at verdiene er bestemt av alder, kjønn, smertevarighet eller hvilken skulder som er testet. Derimot er det store begrensinger ved antall deltagere i gruppen annen skulderpatologi der det kun var fire deltagere, hvilket er en begrensning i denne studien. Grunnen til at positiv LR ikke kunne regnes ut ved for eksempel Jobe test ved sammenlikning av SIS og annen skulderpatologi kan skyldes få deltagere i gruppen annen skulderpatologi. Dette medførte at noen av verdiene ble kunstig gode eller dårlige. Grunnen til at antallet er lavere i gruppen annen skulderpatologi skyldes mest sannsynlig at 74% av pasienter med skuldersmerter har impingement (Östör et al. 2005). Dette stemmer veldig godt overens med vår studie der 76% av deltagerene fikk diagnosen impingement ved ultralyd.

Syv av 40 deltagere i vår studie ble ekskludert. Halvparten av disse ble ekskludert på bakgrunn av at de mottok annen fysikalsk behandling og to takket nei grunnet tidsmangel. Det kan ikke trekkes parallell mellom deltagerens skulderplager og årsak til eksklusjon. Det vil derfor sannsynligvis ikke ha stor betydning for validiteten i studien.

Konfidensintervallet er regnet ut på bakgrunn av formelen i tabell 4, som er basert på tilnærmelsen av normalfordelingen. Aalen (2006) viser til at antall deltagere bør være over fem. Tilnærmelsen er derfor ikke helt optimal, men konfidensintervallet sier likevel noe om hvor stor usikkerheten er i estimatene av sensitivitet og spesifisitet. Det kunne også vært mulig å beregne konfidensintervall på for eksempel LR, men dette er ikke gjort grunnet for få antall observasjoner (Aalen, 2006).

Utvelgelse av impingementtestene

En annen styrke i denne studien kan være utvelgelsen av de fire skuldertestene for impingement, da testene ble nøye utvalgt fra verdiene ved den kliniske validiteten og

reliabiliteten til tidligere studier (Calis et al. 2000, Park et al. 2005, Silva et al. 2008, Michener et al. 2009, Kelly et al. 2010, Alqunae et al. 2012). Alle de fire testene beskrives som impingementtester i litteraturen (Magee, 2008, Juel, 2007, Solberg, 2002, Hattam & Steatham, 2010, Bahr & Mæhlum, 2003, Gross et al. 1996, Bahr & Mæhlum, 2003). Samtidig er det flere andre studier, samt flere andre tester som kan anvendes i klinisk praksis som ”drop arm” test, ”lift off” test, utadrotasjon med motstand og infraspinatustest (Calis et al. 2000, Park et al. 2005, Michener et al. 2009, Kelly et al. 2010, Alqunae et al. 2012). Men grunnet utvelgelse på bakgrunn av tidligere studiers verdier ville indragelse av disse testene mest sannsynlig ikke skilt seg bedre ut i resultatet.

5.3 Ekstern validitet

Ekstern validitet innebærer muligheten for å generalisere resultatene til eller over relevante individer, situasjoner og tider (Lund, 2002). Generaliseringsproblemet vil ofte være et multidimensjonalt problem i form av om det er relevant å generalisere ut over individer, men også til situasjoner og tid (Lund, 2002). Resultatet må kunne generaliseres til en populasjon som overholder inklusjon – og eksklusjonskriteriene. I dette forskningsprosjektet minskes den eksterne validiteten grunnet få deltagere i gruppen annen skulderpatologi. De få pasientene i denne gruppen kan derfor gi liten generaliserbarhet i forhold til dataene som innsamles. I motsetning kan det være en styrke med få deltagere, da det er lettere å klarlegge og kontrollere mulige feilkilder (Bordens og Abbott, 2005). I impingementgruppen ($n = 13$) og gruppen med friske ($n = 16$) var antallet deltagere høyere, noe som styrker generaliseringsspørsmålet. I forhold til deltagerne vil det også alltid være et spørsmål om individhomogenitet, da utvelgelsen av deltagerne blir skjønnsmessige. Det vil derfor være risiko for at utvalget systematisk kan avvike fra populasjonen og at generaliserbarheten blir redusert (Holme & Solvang, 1996).

Det er viktig med god ekstern validitet dersom resultatene skal kunne generaliseres. En tilfredsstillende statistisk validitet, intern validitet og begrepsvaliditet er irrelevant hvis resultatene ikke kan generaliseres til en populasjon. Derved vil forskningsproblemet bli tilsvarende dårlig belyst og undersøkelsen som helhet bli svekket (Lund, 2002). I denne studien kunne den eksterne validiteten blitt større ved økt antall deltagere, men dette lot seg dessverre ikke gjøre grunnet studiens begrensede tidsramme. Til tross for lavere deltagertall enn ønsket kan resultatene likevel antyde at det er liten sannsynligheten for SIS ved negativ

test av Neers, Hawkins eller Painful Arc, mens Jobe test indikerer at sjansen for SIS er meget stor ved positiv test sammenliknet med annen skulderpatologi. Ved sammenlikning mot friske indikerte testene Neers, Jobe og Painful Arc at sjansen for SIS var stor ved positiv test, mens Hawkins test indikerte at sjansen for SIS var liten ved negativ test. Dette gjør at vi som manuellterapeuter i klinisk praksis kan mistenke utelukking av SIS ved negativ Hawkinstest og høyst sannsynlig mistenke SIS ved positiv Jobe test.

5.4 Videre forskning

Grunnet tidsfristen for levering av masteroppgave ble antall deltagere i studien kun 33. Terwee et al. (2007) anbefaler minimum 50 pasienter for å få tilstrekkelig sikkerhet og statistisk power i estimatene. På bakgrunn av dette vil vi fortsette rekrutering og testing frem til 1. desember 2012 for at forskningsprosjektet som helhet blir styrket. Målet vil være å skrive en artikkel for å publisere studien. Videre kunne det vært interessant å vurdere som Michener et al. (2009) og Park et al. (2005) om inklusjon av flere tester sammen kan forbedre den diagnostiske nøyaktigheten.

6.0 Konklusjon

Diagnostikk av SIS er en utfordring for klinikere og den diagnostiske nøyaktighet av de kliniske testene trenger å anses i kontekst av hele vurderingen av pasienten. Neers, Hawkins og Painful Arc test indikerte ved sammenlikning mot annen skulderpatologi liten sannsynligheten av SIS ved negativ test, mens Jobe test indikerer at sjansen for SIS er meget stor ved positiv test. Jobe test har også den høyeste nøyaktigheten på 82%. Ved sammenlikning mot friske deltagere indikerte testene Neers, Jobe og Painful Arc at sjansen for SIS var stor ved positiv test. I motsetning illustrerte Hawkins test at sjansen for SIS var liten hvis testen var negativ. I tillegg var Hawkins test den mest diagnostisk representative testen med lavest negative LR (0.09). Et positivt funn på en test må likevel brukes med varsomhet alt etter hvor valid testen er, og sannsynligvis er det mest fornuftig å bruke flere tester. I tillegg må testresultatene samsvare og supplere med funnene fra resten av undersøkelsen. Nøyaktig diagnostisering av SIS kan føre til redusert feilaktige henvendelser til kirurgi og fysioterapeut. Dette kan spare samfunnet for unødvendig tid og kostnader for både pasienten og samfunnet.

Referanseliste

Aalen OO: Statistiske metoder i medisin og helsefag, Gyldendal, 2006.

Accident Compensation Corporation (ACC): The diagnosis and management of soft tissue shoulder injuries and related disorders, Accessed February 25, 2011

Altamann DG, Deeks JJ: Diagnostic test 4: Likelihood Ratios, statistics notes, BMJ Volume 329, 2004

Alqunaee M, Galvin R, Fahey T: Diagnostic accuracy of clinical tests for subacromial impingement syndrome: a systematic review and meta-analysis, Arch Phys Med Rehabil, 2012 Feb;93(2):229-36

Awerbuch, M.S.: The clinical utility of ultrasonography for rotator cuff disease: shoulder impingement syndrome and subacromial bursitis. Med J Aust 2008;188(1):50-3

Bahr R, Mæhlum S: Idrettsskader. Oslo, Gazette bok, 2010

Bandy, WD, Rusche KR, Tekulve FY: Reliability and limb symmetry for five unilateral functional tests of the lower extremities. Isokinetics and Exercise Science, 1994.

Batterham, AM, George, KP: Reliability in evidence-based clinical practice: A primer for allied health professionals. Physical Therapy in Sport, 2003.

Benestad HB, Laake P: Forskningsmetode i medisin og biofag, Gyldendal, 2004

Brox JL, Sunde P, Schrøder CP, Engebretsen K, Skare Ø, Ekeberg OM, Juel NG: Atraumatiske skulderlidelser, tidsskrift for den norske legeförening, 2011

Beyer N, Magnusson P: Målemetoder i fysioterapi. København, Munksgaard, 2003.

Bigliani LU, Morrison DS, April DW: The morphology of the acromion and its relationship to rotator tears. Orthop Trans 10: 216, 1986

Bojsen-Møller F: Bevægeapparatets anatomibog, Munksgaard, 12 utgave, 2012

Brodal P: Sentralnervesystemet. Universitetsforlaget, 2007.

Calis M, Akgun K, Calis H, Tuzun F: Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Am Rheum Dis* 2000; 59: 44-7

Dalland O: Metode og oppgaveskriving for studenter, Gyldendal akademisk, 4. utg. 2007.

Dinnes J, Loveman E, McIntyre L, Waugh N: The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systemic review. *Health Technol Assess* 2003; 7: 96–112.

Ekeberg OM, Bautz-Holter E, Tveitå EK: Subacromial ultrasound guided or systemic steroid injection for rotator cuff disease: randomised double blind study. *BMJ* 2009; 338: a3112.

Gartsman GM: Arthroscopic treatment of rotator cuff disease. *J Shoulder Elbow Surg* 1995; 4: 228–41.

Gross J, Fetto J, Rosen E: *Muskuloskeletal examination*, 2nd Edition, Blackwell Publishing Company, 2006

Halvorsen K: *Å forske på samfunnet: En innføring i samfunnsvitenskapelig metode*. Cappelen akademisk forlag Oslo 2008.

Harvey H: *The diagnosis of subacromial impingement syndrome and associated pathology in the primary care setting*. Aukland: Aukland University of Technology; 2009

Hattam P, Smeatham A: *Special tests in musculoskeletal examination*. Churchill Livingstone 2010.

Holme IM, Solvang BK: *Metodevalg og metodebruk* 2, 3. utgave, TANO, Oslo, 1996.

Holtby R, Razmjou H: Validity of the supraspinatus test as a single clinical test in diagnosing patients with rotator cuff pathology. *J Orthop Sports Phys Ther* 2004; 34: 194–200.

Ihlebak C, Lærum E: Plager flest koster mest: muskel - og skjelettlidelser i Norge, Nasjonalt ryggnettverk, Forskningsenheten/formidlingsenheten, 2004
(http://www.formi.no/images/uploads/pdf/rapport_sept_04.pdf)

Itoi E, Minagawa H, Yamamoto N, Seki N, Abe H: Are pain location and physical examinations useful in locating a tear site of the rotator cuff? *Am J Sports Med* 2006; 34: 256–64.

Jamtvedt G, Hagen KB, Bjørndal A: Kunnskapsbasert fysioterapi. Metoder og arbeidsmåter. Gyldendal akademisk, Oslo 2003.

Jensen TS, Dahl JB, Lerche B: Smerter- baggrund, evidens og behandling, Danmark, FADL, 2009

Juel NG: Norsk fysikalsk medisin. 2.utgave. Fagbokforlaget, Bergen 2007.

Kelly SM, Brittle N, Allen GM: The value of physical tests for subacromial impingement syndrome: a study of diagnostic accuracy. *Clin Rehabil.* 2010 Feb; 24(2): 149-58

Kessel L, Watson M: The painful arc syndrome. Clinical classification as a guide to management. *J Bone Joint Surg Br.* 1977 May;59(2):166-72.

Kromer TO, de Bie RA, Bastiaenen CH: Effectiveness of individualized physiotherapy on pain and functioning compared to a standard exercise protocol in patients presenting with clinical signs of subacromial impingement syndrome. A randomized controlled trial, *BMC Musculoskelet Disord.* 2010 Jun 9;11:114.

Kruuse E.: Kvantitative forskningsmetoder – i psykologi og tilgrænsende fag. Danske psykologisk forlag, 6 utgave 2006.

Lindbæk M, Skovlund E: Kontrollerte kliniske forsøk - jakten på sann effekt av behandling
Tidsskr Nor Lægeforen 2002; 122:2631-5

Lund T, Kleven TA, Kvernbekk T, Christophersen KA: Innføring i forskningsmetodologi.
Akademika, Oslo 2002.

Lærum E, Brox JI, Storheim K, Espeland A, Haldorsen E, Munch-Ellingsen J, Nielsen LL,
Rossvoll I, Skouen JS, Stig LC, Werner EL: Nasjonale kliniske retningslinjer,
Korsryggsmarter-med og uten nerverotaffeksjon, Formi (Formidlingsenheten for muskel- og
skjelettlidelser), Sosial - og Helsedirektoratet, 2007:
http://www.formi.no/images/uploads/pdf/Formi_net.pdf

Løvås GG: Statistikk for univertiteter og høyskoler, Univeritetsforlaget, 2.utgave, 2004

Lu et al: Predicting Subcellular Localization of Proteins using Machine-Learned Classifiers,
Bioinformatics, Volume 20, Issue 4, March 2004, pp. 547 - 556.

Magee, DJ.: Orthopedic physical assessment, ed 5. Saunders Elsevier, 2008

Malhi AM, Khan, R: Correlation between clinical diagnosis and arthroscopic findings of the
shoulder. Postgrad Med J 2005; 960(81): 657-9

MacDonald PB, Clark P, Sutherland K: An analysis of the diagnostic accuracy of the
Hawkins and Neer subacromial impingement signs. J Shoulder Elbow Surg 2000; 9: 299-301.

Meislin RJ, Sperling JW, Stitik TP: Persistent shoulder pain: epidemiology, pathophysiology,
and diagnosis. American Journal Orthopedics (2005) 34:5-9.

Michener LA, Walsworth MK, Doukas WC, Murphy KP: Reliability and diagnostic accuracy
of 5 physical examination tests and combination of tests for subacromial impingement. Arch
Phys Med Rehabil. 2009 Nov; 90(11): 1898-903

Milgrom C, Schaffler M, Gilbert S, van Holsbeeck M: Rotator-cuff changes in asymptomatic adults: the effect of age, hand dominance and gender. *J Bone Joint Surg Br* 1995; 77-B: 296–98.

Moosmayer S, Heir S, Aaser P, Smith H: Ultralyddiagnostikk av skulderbuen - en metodebeskrivelse, *Tidsskrift for den Norske Legeforening*, nr. 2 2004; 124:177-80

Moen MH, de Vos RJ, Ellenbecker TS, Weir A: Clinical tests in shoulder examination: how to perform them. *Br J Sports Med* 2010;44(5):370-5

Naredo E, Aguado P, De Miguel E: Painful shoulder: comparison of physical examination and ultrasonographic findings. *Ann Rheum Dis* 2002;61: 132–36.

Neer CS: Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. *J. Bone Joint Surg.* 1972, [Am] 54, 41–50.

Park HB, Yokota A, Gill HS, El Rassi G, McFarland EG: Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndorme. *J Bone Joint Surg Am* 2005;87(7):1446-55

Portney LG, Watkins MP: *Foundations of clinical research: applications to practice.* (vols. 2nd ed.) Upper Saddle River, N.J.: Prentice Hall, 2000.

Reed Group: Rotator cuff tear. In: *The medical disability advisor: Workplace for guidelines for disability duration.* 2009; Westminster: Reed Group.

Ringdal K: *Enhet og mangfold, Samfunnsvitenskapelig forskning og kvantitativ metode*, 2. utgave. Kunnskapsforlaget, 2009.

Ruyter KW, Førde R, Solbakk JH: Medisinsk etikk –en problembasert tilnærming, Gyldendal norsk forlag, 2000.

Silva L, Andreu JL, Munoz P, Pastrana M, Millan I, Sanz J, Barbadillo C, Fernandez-Castro M: Accuracy of physical examination in subacromial impingement syndrome. *Rheumatology* 2008; 47(5): 679-83

Solberg AS: Klinisk undersøkelse av nakke-skulder. Høyskoleforlaget, Kristiansand 2002.

Streiner DL, Norman GR: Health measurement scales: a practical guide to their development and use. (vols. 4th ed.) Oxford: Oxford University Press, 2008.

Svartdal F: Psykologiens forskningsmetoder, En introduksjon, 3. Utgave, Bergen: Fagbokforlaget 2009

Svendsen IF: Arbeidstilsynet, 2011 (<http://www.hmsmagasinet.no/id/51067.0>)

Straus SE, Richardson WS, Glasziou P, Haynes RB: Evidence-based medicine. How to practise and teach EBM. Elsevier Churchill Livingstone 2005.

Torstensen TA, Østerås H, Harms-Ringdahl K: Dose –responseeffekten ved bruk av medisinsk treningsterapi ved langvarig subacromialt smertesyndrom, *Fysioterapeuten* nr. 10, 2005

Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC: Quality criteria were proposed for measurement properties of health status questionnaires. *VU University Medical Center*, 2007 Jan;60(1):34-42.

Üstün TB: International Classification of Functioning, World Health Organization, Geneva, 2002

Valadie AL, Jobe CM, Pink MM, Ekman EF, Jobe FW: Anatomy of provocative tests for impingement syndrome of the shoulder. *J Shoulder Elbow Surg* 2000; 9(1): 36-46

Van Der Windt DAWM, Koes BW, Boeke AJP, Devillé W, De Jong BA, Bouter LM:
Shoulder disorders in general practice: prognostic indicators of outcome. *British Journal of General Practice* 1996; 46: 519-23

Östör AJK, Richards CA, Prevost AT, Speed CA, Hazleman BL: Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology* 2005; 44: 800-5

Vedlegg

Vedlegg 1: Anonymt skåringskjema

SKÅRINGSSKJEMA TESTER

- Test-rekkefølge anonymiseres!
- Terapeut-rekkefølge anonymiseres!
- 2 min mellom hver test, 3 min mellom hver terapeut

Pasientnr:

Test	Rekkefølge	Terapeut #1 positiv = 1 negativ = 0	Terapeut #2 positiv = 1 negativ = 0
<i>Jobes test</i>			
<i>Neers test</i>			
<i>Hawkins test</i>			
<i>Painful arc</i>			
<i>Sumscore</i>			

Vedlegg 2: Forespørsel om deltagelse i studien/samtykke

Forespørsel om deltagelse i forskningsprosjektet

”Gyldighet av skuldertester”

Bakgrunn og hensikt

Mellom 30-40 prosent av befolkningen opplever skuldersmerter hvert år. Vanligst er plager knyttet opp mot musklene rundt skulderen og avklemminger lokalt. Nedsatt plassforhold av slimpose og/eller muskulatur i skulderen (subacromialt impingment) er den vanligste skulderlidelsen i befolkningen, og representerer 44-65% av alle registrerte skulderplager i kliniske praksis.

Dette er et spørsmål til deg om å delta i en studie for å undersøke gyldigheten av tester som vi bruker for å diagnostisere smertetilstander i skulder. Vi ønsker å undersøke om testene våre er nøyaktige nok.

Hva innebærer studien?

Dersom du takker ja til å delta i studien, vil du bli kalt inn til selve testprosedyren 2 ganger med en ukes mellomrom. Det vil da bli gjennomført 4 enkle kliniske tester av to ulike terapeuter første gang, og av kun en terapeut den ene andre gangen. Testene som utføres brukes for å se om du har det gitte skulderproblemet eller ikke, og er helt ufarlige. Det vil ta ca. 20 minutter hver gang.

Mulige fordeler og ulemper

Fordeler:

Det vil bli gitt informasjon, råd og behandlingsforslag til deg som deltar.

Ulemper:

Du må komme til testing to ganger.

Hva skjer med prøvene og informasjonen om deg?

Prøvene tatt av deg og informasjonen som registreres om deg skal kun brukes slik som beskrevet i hensikten med studien. Alle opplysningene og prøvene vil bli behandlet uten navn og fødselsnummer eller andre direkte gjenkjennende opplysninger. En kode knytter deg til dine opplysninger og prøver gjennom en navneliste.

Det er kun autorisert personell knyttet til prosjektet som har adgang til navnelisten og som kan finne tilbake til deg.

Det vil ikke være mulig å identifisere deg i resultatene av studien når disse publiseres.

Frivillig deltagelse

Det er frivillig å delta i studien. Du kan når som helst og uten å oppgi noen grunn trekke ditt samtykke til å delta i studien. Dette vil ikke få konsekvenser for din videre behandling.

Dersom du ønsker å delta, undertegner du samtykkeerklæringen på siste side. Om du nå sier ja til å delta, kan du senere trekke tilbake ditt samtykke uten at det påvirker din øvrige

behandling. Dersom du senere ønsker å trekke deg eller har spørsmål til studien, kan du kontakte

Ytterligere informasjon om studien finnes i kapittel A – utdypende forklaring av hva studien innebærer.

Ytterligere informasjon om biobank, personvern og forsikring finnes i kapittel B – Personvern, biobank, økonomi og forsikring.

Samtykkeerklæring følger etter kapittel B

Kapittel A- utdypende forklaring av hva studien innebærer

Kriterer for deltagelse:

Det vil bli rekrutert pasienter i 2 forskjellige utvalg. Et utvalg der ultralydundersøkelse har vist nedsatt plassforhold i skulderen (positiv impingement) og et utvalg der skulderplagene skyldes andre forhold.

For å være med i gruppen med nedsatt plassforhold i skulderen gjelder følgende kriterier:

Diagnostisert med nedsatt plassforhold i skulderen ved ultralyd

Skuldersmerter med varighet mer enn én måned

Skuldersmertene må være pasientens aktuelle hovedproblem

Må kunne gi et informert samtykke

Mellom 18 og 60 år

Du kan ikke inngå i gruppen med redusert plassforhold i skulderen ved:

Andre kjente plager i skulder

Tidligere skade mot skulder

Tidligere skulderoperert

Betennelsesdannende leddlidelser

Smerter ved bevegelse av nakken

Smerter i nakke med utstrålende plager ut til fingrene

Mistenke om alvorlig sykdom

For å være med i gruppen for andre skulderplager, gjelder følgende kriterier:

Diagnostisert med andre skulderplager ved hjelp av ultralyd

Skuldersmerter med varighet mer enn én måned

Skuldersmertene må være pasientens aktuelle hovedproblem

Må kunne gi et informert samtykke

Mellom 18 og 60 år

Du kan ikke inngå i gruppen med andre skulderplager ved:

Andre kjente plager i skulder

Tidligere skade mot skulder

Tidligere skulderoperert

Betennelsesdannende leddlidelser

Smerter ved bevegelse av nakken

Smerter i nakke med utstrålende plager ut til fingrene

Mistenke om alvorlig sykdom

Bakgrunnsinformasjon om studien:

Vi møter i vår hverdag mange pasienter med skulderproblemer. For å skille de ulike skulderproblemene fra hverandre bruker vi ulike tester for å kunne stille riktig diagnose. Det er derfor svært sentralt at testene vi bruker er pålitelige og diagnostisk nøyaktige. Hensikten med denne studien er nettopp å undersøke kvaliteten på fire tester brukt for å diagnostisere nedsatt plassforhold av slimpose og/eller muskulatur i skulderen.

Tidsskjema:

Rekrutteringen av pasienter til mastergradstudiet vil skje på

Datainnsamlingen vil foregå på Det vil bli gjort målinger på testene i skulderen, og det vil registreres om den gitte testen fremprovoserer smerte i skulderen eller ikke.

Datainnsamlingene vil foregå vår/høst 2012 og bearbeidelse høst 2012. Tidsfristen for ferdig prosjekt er 1. november 2012. Studiet vil bli gjennomført som en test- retest analyse. De kliniske testene som vil gjennomføres er Neers test, Hawkins-Kennedy test, Jobe/empty can-test og painful arc.

Mulige fordeler:

Du vil delta i et prosjekt som kan bedre kvaliteten på tester som brukes ved diagnostisering av skulderproblemer. Du vil også få informasjon og råd vedrørende dine skulderplager.

Mulige bivirkninger:

Det er ingen bivirkninger da det ikke gis medikamenter eller tas blodprøver etc.

Mulige ubehag:

Testene kan gi lett forbigående ubehag i skulder. Det legges derfor opp til pauser mellom hver test slik at eventuelle ubehag er borte før utførelse av hver test.

Pasientens/studiedeltakerens ansvar:

Eneste krav til deg som deltager i studiet er at du møter opp til avtalt tid 2 ganger med en ukes mellomrom. Ved ønske om eventuell videre behandling avtales dette på klinikken.

Kompensasjon/utgifter:

Eventuell kompensasjon til og dekning av utgifter for deltakere gis ikke. Informasjon, råd og treningsprogram med øvelser er gratis. Eventuell videre behandling vil følge vanlig takstsystem.

Kapittel B - Personvern, biobank, økonomi og forsikring

Personvern

Opplysninger som registreres om deg er navn, kjønn, alder, medikamentbruk, diagnose og testens resultat. Alle som får innsyn i studien har taushetsplikt og ingen data vil bli misbrukt. Hver deltager vil bli nummerert og testerne får kun oppgitt nummeret og ingen annen informasjon.

..... er ansvarlig for rekruttering av pasienter.

..... er ansvarlig for innsamling av data og rekruttering av pasienter.

Utlevering av materiale og opplysninger til andre

Ingen persondata vil i ettertid kunne bli misbrukt. Ved publikasjon av studien vil alle data være anonymisert.

Rett til innsyn og sletting av opplysninger om deg og sletting av prøver

Hvis du sier ja til å delta i studien, har du rett til å få innsyn i hvilke opplysninger som er registrert om deg. Du har videre rett til å få korrigert eventuelle feil i de opplysningene vi har registrert. Dersom du trekker deg fra studien, kan du kreve å få slettet innsamlede opplysninger, med mindre opplysningene allerede er inngått i analyser eller brukt i vitenskapelige publikasjoner.

Økonomi

Det er ingen sponsor som finansierer studiet. Det er derfor ikke mulig å gi kompensasjon for eventuelle utgifter deltageren måtte ha.

Forsikring

Forsikring på vanlig måte gjennom pasientskadeloven.

Informasjon om utfallet av studien

Alle deltakerne i masterstudiet har rett til å få informasjon om utfallet/resultatet av studiet.

Samtykke til deltakelse i studien

Jeg er villig til å delta i studien

(Signert av prosjektdeltaker, sted/dato)

Jeg bekrefter å ha gitt informasjon om studien

(Signert, rolle i studien, sted/dato)

Vedlegg 3: Godkjenning fra REK



Region: REK sør-øst	Saksbehandler: Øyvind Grønlie Olsen	Telefon: 22845522	Vår dato: 05.03.2012	Vår referanse: 2012/57/REK sør-øst D
			Deres dato: 17.01.2012	Deres referanse:

Vår referanse må oppgis ved alle henvendelser

Att: Liv Inger Strand

2012/57 D Reliabilitet og validitet av skuldertester

Vi viser til søknad om forhåndsgodkjenning av ovennevnte forskningsprosjekt. Søknaden ble behandlet av Regional komité for medisinsk og helsefaglig forskningsetikk i møtet 09.02.2012.

Prosjektleder: Liv Inger Strand
Forskningsansvarlig: Universitetet i Bergen

Prosjektomtale

Hensikten med studien er tredelt:

- 1. Undersøke den diagnostiske validiteten ved fire kliniske tester som brukes for å stille en impingement-diagnose*
- 2. Undersøke intertester-reliabiliteten ved de samme testene*
- 3. Undersøke test-retest-reliabiliteten ved de samme testene.*

Skuldersmerter generelt er et omfattende problem med konsekvenser for individet og samfunnet. En stor andel av smertene skyldes avklemming av strukturer i øvre deler av skulderleddet, kalt impingement syndrom. For å stille diagnosen brukes i dag blant annet MR, ultralyd og røntgen i tillegg til en rekke kliniske tester. De første tre testene regnes som ganske sikre tester for å påvise skulder impingement. Det er ønskelig å kunne påvise tilstanden ved hjelp av enkle diagnostiske tester i vanlig klinisk praksis. Dette fordrer at testene er diagnostisk valide og at de har høy reliabilitet.

Testene som skal utføres er Hawkins test, Neers test, painful arc og Jobes test. Testene blir scorede med utgangspunkt i dikotomien positiv eller negativ test. For Hawkins, Neers og painful arch, betyr dette at testen scores som positiv hvis den primært gir smerter i skulderområdet.

Totalt deltar 50 pasienter med skulderplager. Halve gruppen har fått indikasjon på skulderimpingement ved hjelp av ultralydundersøkelse mens den andre halvparten har fått indikasjon på at skulderplagene ikke skyldes impingement.

Vurdering

Komiteen anser prosjektet som en prospektiv sammenligning av etablerte metoder. Prosjektets formål er etter komiteens mening mer enn en ren evaluering av praksis. En effektivisering av behandlingkjeden kan føre til at pasienter blir gitt riktig diagnose på raskere tid – noe som kan føre til at riktig behandling til vedkommende pasient gis på et tidligere tidspunkt. Forskningsprosjektet er etter dette fremleggelsespliktig, jf. helseforskningsloven § 10.

Besøksadresse: Gullhaug torg 4A, Nydalén, 0484 Oslo	Telefon: 22845511 E-post: post@helseforskning.etikkom.no	All post og e-post som inngår i saksbehandlingen, bes adressert til REK sør-øst og ikke til enkelte personer	Kindly address all mail and e-mails to the Regional Ethics Committee, REK sør-øst, not to individual staff
	Web:		

Informasjonssikkerhet

Komiteen kan ikke se at det er oppgitt databehandlingsansvarlig for studien. Det forutsettes at Universitetet i Bergen er databehandlingsansvarlig for prosjektet.

Informasjonsskriv

I informasjonsskrivet benyttes det til dels vanskelige medisinske ord og uttrykk. I den grad det er mulig skal innholdet i informasjonsskrivet presenteres på en lettfattelig måte. Medisinske ord og uttrykk som må erstattes med alminnelige ord og uttrykk.

Det er et uttalt krav at presentasjonen av forskningsprosjektets innhold skal være nøytralt. Det innsendte informasjonsskrivet inneholder formuleringer som ikke tilfredsstillende dette kravet. Før rekruttering til forskningsprosjektet gjennomføres forutsetter komiteen at informasjonsskrivet bearbejdes slik at opplysninger om prosjektets innhold presenteres i nøytral form. Revidert informasjonsskriv ettersendes komiteen til orientering.

Vedtak

Komiteen har vurdert søknaden og godkjenner prosjektet med hjemmel i helseforskningsloven § 9 jf. § 33. Godkjenningen er gitt under forutsetning at de vilkår som er stilt innarbejdes i protokoll og informasjonsskriv.

Godkjenningen av prosjektet gjelder til 31.12.2014. Av dokumentasjonshensyn skal opplysningene likevel bevares inntil 31.12.2019. Opplysningene skal deretter slettes eller anonymiseres, senest innen 1.6.2020.

Opplysningene skal lagres avidentifisert, det vil si adskilt i en nøkkel- og en opplysningsfil.

Forskningsprosjektets data skal oppbevares forsvarlig, se personopplysningsforskriften kapittel 2, og Helsejektivitatets veileder for «Personvern og informasjonssikkerhet i forskningsprosjekter innenfor helse- og omsorgssektoren».

Prosjektet skal sende sluttmelding til REK sør-øst D på fastsatt skjema senest 1.6.2015.

Tillatelsen er gitt under forutsetning av at prosjektet gjennomføres slik det er beskrevet i søknaden og protokollen, og de bestemmelser som følger av helseforskningsloven med forskrifter.

Dersom det skal gjøres vesentlige endringer i prosjektet i forhold til de opplysninger som er gitt i søknaden, må prosjektleder sende endringsmelding til REK.

Komiteens vedtak kan påklages til Den nasjonale forskningsetiske komité for medisin og helsefag, jfr. helseforskningsloven § 10, 3 ledd og forvaltningsloven § 28. En eventuell klage sendes til REK Sørøst D. Klagefristen er tre uker fra mottak av dette brevet, jfr. forvaltningsloven § 29.

Vennligst oppgi vårt referansenummer i korrespondansen.

Med vennlig hilsen,

Stein A. Evensen (sign.)
dr. med.
leder

Øyvind Grønlie Olsen
saksbehandler

Vedlegg 4: Baselinedata og testscore

Gruppe/Diagnose	Deltager	Kjønn	Alder	Sykdoms varighet i mnd	Testside	J	N	H	P
Gruppe 1/Impingement	1	K	49	12	H	1	0	1	1
Gruppe 1/Impingement	2	M	55	1	V	1	1	1	1
Gruppe 1/Impingement	3	M	60	9	V	1	1	1	0
Gruppe 1/Impingement	13	K	28	9	V	1	1	1	1
Gruppe 1/Impingement	15	K	46	360	V	1	0	1	1
Gruppe 1/Impingement	17	K	53	12	H	1	1	1	0
Gruppe 1/Impingement	19	M	44	6	V	1	1	1	0
Gruppe 1/Impingement	20	M	60	18	V	1	1	1	1
Gruppe 1/Impingement	21	M	57	12	V	0	1	1	1
Gruppe 1/Impingement	22	M	60	12	V	1	1	1	1
Gruppe 1/Impingement	24	K	33	48	V	0	1	1	1
Gruppe 1/Impingement	25	M	18	1	H	1	0	0	0
Gruppe 1/impingement	33	K	42	12	V	0	1	1	1
Gruppe 2/Annen skulderpatologi	12	K	33	24	H	0	1	1	0
Gruppe 2/Annen skulderpatologi	23	M	42	9	H	0	0	1	0
Gruppe 2/Annen skulderpatologi	26	M	45	12	H	1	1	1	1
Gruppe 2/Annen skulderpatologi	28	K	55	19	H	0	1	1	1
Gruppe 3/friske	4	K	31	-	H	0	1	1	0

Gruppe 3/friske	5	M	35	-	V	0	0	1	0
Gruppe 3/friske	6	K	30	-	H	1	1	1	1
Gruppe 3/friske	7	M	30	-	V	0	1	1	1
Gruppe 3/friske	8	K	21	-	H	0	0	0	0
Gruppe 3/friske	9	M	60	-	H	0	0	0	0
Gruppe 3/friske	10	K	22	-	H	0	0	0	0
Gruppe 3/friske	11	K	23	-	V	0	0	0	0
Gruppe 3/friske	14	K	59	-	H	0	0	0	0
Gruppe 3/friske	16	K	30	-	H	0	0	0	0
Gruppe 3/friske	18	K	23	-	V	0	0	0	0
Gruppe 3/friske	27	K	46	-	H	0	0	0	0
Gruppe 3/friske	29	M	60	-	H	0	0	1	0
Gruppe 3/friske	30	M	55	-	H	0	0	0	0
Gruppe 3/friske	31	K	55	-	V	0	0	0	0
Gruppe 3/friske	32	K	60	-	H	0	0	0	0

K = Kvinne
M = Mann

H = Høyre
V = Venstre

J = Jobe
N = Neers
0 = Negativ test

H = Hawkins
P = Painful Arc
1 = Positiv test