

# Skulderfunksjon og sammenheng mellom måleinstrumenter etter kompliserte skulderbrudd

En tverrsnittstudie seks måneder postoperativt

**Kari Krokan**

Kandidatnummer: 25  
HELSEF4503  
Masterstudium i tverrfaglig helseforskning  
Antall studiepoeng: 120

Institutt for helse og samfunn - Avdeling for tverrfaglig helsevitenskap  
Universitetet i Oslo



# Sammendrag

**Bakgrunn:** Proximal humerus fraktur (PHF) oppstår hyppigst blant eldre og insidensen øker. Basert på manglende kunnskap om postoperativ funksjon og mangelfull data om psykometriske egenskaper ved måleinstrumentene Disability of the arm, shoulder and hand (DASH), Constant -Murley Score (CMS) og Oxford Shoulder Score (OSS), var det ønskelig å oppnå bedre kunnskap om nevnte instrumenter.

**Hensikt:** Hensikten var å øke kunnskapen om postoperativ funksjon, undersøke eventuelle tak- eller gulveffekter og undersøke sammenheng mellom måleinstrumentene og undersøke sammenheng mellom bakgrunnsvariabler og lav CMS skår hos pasienter med tre og fire-parts PHF seks måneder postoperativt.

**Metode:** Masteroppgaven benyttet sekundærdata fra en randomisert kontrollert studie som omhandler operativ behandling av PHF. Masteroppgaven er en tverrsnittstudie og utvalget bestod av 73 personer med tre eller fire-parts PHF operert med nagle eller plating. Måletidspunktet var seks måneder postoperativt. Pearsons korrelasjon, lineær og logistisk regresjon ble benyttet.

**Resultater:** Resultatene viste gjennomsnittlig god skulderfunksjon målt med DASH og OSS og dårlig skulderfunksjon målt med CMS. Blant menn var det 23% som nådde best mulig skår på OSS, og blant utvalget 22% som ikke oppfylte kriteriene for å utføre CMS styrketest. Analysene viste sterk sammenheng mellom CMS og DASH ( $r = -0,72$ ,  $p < 0,01$ ), CMS og OSS ( $r = 0,73$ ,  $p < 0,01$ ) og DASH og OSS ( $r = -0,84$ ,  $p < 0,01$ ). Personer med ASA 3 hadde høyere odds for lav CMS skår i ujustert estimer (OR= 8,00; 95% KI 1,4 – 45,8) og justerte estimer (OR= 10,70; 95% KI 1,4 – 80,9) sammenlignet med de som hadde ASA 1 eller ASA 2, men det var imidlertid stor usikkerhet i estimatene.

**Konklusjon:** Til tross for sterk sammenheng mellom måleinstrumentene var det lav enighet mellom pasient- og klinikerrapporterte måleinstrumenter. Resultatene kan indikere takeffekt hos menn ved bruk av OSS, og høy andel manglende data ved styrkeområdet ved CMS kan indikere lav enighet for instrumentet. Resultatene kan indikere at personer som er vurdert til ASA 3 har høyere odds for lav CMS skår enn de med lavere ASA skår seks måneder postoperativt.

## Abstract

**Introduction:** Proximal humerus fracture (PHF) occurs most frequently among the elderly and the incidence is increasing. Based on a lack of knowledge about postoperative function and insufficient data on psychometric properties of the measuring instruments Constant - Murley Score (CMS), Disability of the arm, shoulder and hand (DASH) and Oxford Shoulder Score (OSS), the aim was to achieve better knowledge of these instruments.

**Purpose:** The purpose was to increase knowledge of postoperative function, examine any ceiling- or floor effects, examine the relationship between the instruments and examine the association between background variables and low CMS score in patients with three- and four-part PHF six months postoperatively.

**Method:** This master thesis used secondary data from a randomized controlled study which examined operative treatments of PHF. This master thesis is a cross-sectional study. The sample consisted of 73 people with three- or four-part PHF surgically treated with nail or plate. The follow up time was six months postoperatively. Pearsons correlation, linear and logistic regression were used.

**Results:** On average the results showed good shoulder function measured with DASH and OSS and poor shoulder function measured with CMS. Among men, 23% achieved the best possible score on the OSS, and among the sample 22% did not meet the criteria for performing CMS strength test. The results showed a strong correlation between CMS and DASH ( $r = -0,72$ ,  $p < 0,01$ ), CMS and OSS ( $r = 0,73$ ,  $p < 0,01$ ) and DASH and OSS ( $r = -0,84$ ,  $p < 0,01$ ). People with ASA 3 had higher odds of a low CMS score compared to those who had ASA 1 or ASA 2 in unadjusted model (OR= 8,00; 95%KI 1,4 – 45,8) and adjusted model (OR= 10,70; 95%KI 1,4 – 80,9), however, there was great uncertainty in the estimates.

**Conclusion:** Despite the strong correlation between the instruments, there was low agreement between patient- and clinician-reported measures. The results may indicate a ceiling effect for men when using OSS, and a high proportion of missing data in the strength domain of CMS may indicate a low suitability of the instrument. The results may indicate that people assessed as ASA 3 have higher odds for a low CMS score six months postoperatively compared to people with lower ASA scores.

## **Forord**

Proessen med å skrive masteroppgaven har inspirert meg både faglig og personlig. Det har vært en krevende øvelse å kombinere studier og full jobb, men det har samtidig vært svært lærerikt og givende. Utdanningen og masteroppgaven har gitt meg ny kunnskap og nye perspektiver, som jeg håper på å kunne bruke og bygge videre på senere.

Jeg vil takke min hovedveileder Nina Østerås for utrolig god veiledning, hjelp og støtte underveis. Jeg ønsker å takke Annette Wikerøy for samarbeidet under datainnsamlingen og tillatelse til å benytte datamaterialet fra hennes prosjekt. Jeg må også rette en stor takk til min arbeidsgiver, Akershus Universitetssykehus, for denne muligheten og mine nærmeste for deres optimisme og støtte.

Lørenskog, 06.11.23  
Kari Krokan

# Begrepsavklaring

<b>ASA-klassifisering</b>	Operativ risikoskår fra 1-5 (American Society of Anesthesiologists)
<b>Clavicula</b>	Kragebeinet
<b>Elevasjon</b>	Løfte armen over skulderhøyde
<b>Fleksjon, ekstensjon og abduksjon</b>	Bøy, strekk og utoverføring i skulderleddet
<b>Humerus</b>	Overarmsbeinet
<b>Insidens</b>	Nye tilfeller av skade
<b>Komorbiditet</b>	Flere sykdommer samtidig
<b>Prevalens</b>	Forekomst av skade
<b>Reliabilitet</b>	Hvor nøyaktig målet er
<b>Responsivitet</b>	I hvilken grad et måleinstrument kan oppdage endring over tid
<b>Scapula</b>	Skulderbladet
<b>Validitet</b>	Hvilken grad en måler det som er hensikten å måle

## Forkortelser

<b>ADL</b>	Dagligdagse aktiviteter (Activities of daily living)
<b>CMS</b>	Constant Murley Score (Måleinstrument)
<b>CROM</b>	Klinikerrapportert måleinstrument (Clinician-reported outcome measure)
<b>DASH</b>	Disability of the arm, shoulder and hand (måleinstrument)
<b>ICF</b>	Internasjonal klassifisering av funksjon, funksjonshemning og helse (International Classification of functioning, disability and health)
<b>KI</b>	Konfidensintervall, et intervall hvor den sanne verdien ligger innenfor
<b>MCID</b>	Minste kliniske viktig forskjell som er meningsfull for pasienten (Minimal Clinically Important Difference)
<b>MDC</b>	Minste oppdagbare endring (Minimal Detectable Change)
<b>OSS</b>	Oxford Shoulder Score (måleinstrument)
<b>PROM</b>	Pasientrapportert måleinstrument (Patient-reported outcome measure)
<b>P-verdi</b>	Signifikansnivå
<b>RCT</b>	Randomisert kontrollert studie (randomised controlled trial)
<b>ROM</b>	Bevegelsesutslag (Range of motion)
<b>SD</b>	Standardavvik, gjennomsnittlig avstand fra gjennomsnittet
<b>VAS</b>	Pasientrapportert skala for gradering av smerte (Visual Analog Scale,)

# Figuroversikt

Nummer	Beskrivelse	Sidetall
1	Anatomi av proksimale humerus inndelt i farger	16
2	Knokler og leddforbindelser i skulderledd og skulderbue sett forfra	17
3	Neer klassifisering illustrert med CT bilder	19
4	Metall goniometer	27
5	Utførelse av måling av fleksjon i skulderen	28
6	Utførelse av styrkedomene Constant Murley Score	28
7	Fjærvekt Berkley Precision Digital Scale	29
8	Flytskjema viser endeligutvalg i masteroppgaven	37
9	Histogram med aldersfordelingen i utvalget	43
10	Histogram med aldersfordeling gruppert på kjønn	44
11	Boksplott med aldersfordeling gruppert på kjønn	44
12	Histogram av skåringsfordeling for CMS (n=73)	46
13	Histogram gruppert på kjønn med skåringsfordeling CMS (n=73)	46
14	Stolpediagram for kategorisert CMS gruppert på kjønn	47
15	Stolpediagram for dikotomisert utført vs. ikke utført styrketest, gruppert på CMS skår $\geq 56$ og $< 56$	47
16	Histogram av skåringsfordeling for DASH (n=73)	48
17	Stolpediagram for kategorisert DASH skår fordelt på kjønn	48
18	Histogram av skåringsfordeling for OSS (n=73)	49
19	Stolpediagram for dikotomisert OSS skår gruppert på kjønn	51
20	Manglende data ved styrkedomenet av CMS	51
21	Spredningsdiagram mellom variablene CMS og DASH (n=73)	52
22	Spredningsdiagram mellom variablene CMS og OSS (n=73)	52
23	Spredningsdiagram mellom variablene DASH og OSS (n=73)	53
24	Spredningsdiagram mellom uavhengig variabel Leverage verdi og avhengig variabel predikert sannsynlighet (n=73)	56
25	Spredningsdiagram mellom uavhengig variabel Cook's D og avhengig variabel predikert sannsynlighet (n=73)	56

## Tabelloversikt

Nummer	Beskrivelse	Sidetall
1	Normative gjennomsnittsverdier CMS	30
2	Normative gjennomsnittsverdier DASH	31
3	Normative gjennomsnittsverdier OSS	33
4	Sammenheng mellom måleinstrumentene i tidligere litteratur	34
5	Resultater fra litteratursøk, sammenheng mellom bakgrunnsvariabler og lav CMS	35
6	Inklusjon og eksklusjonskriterier i hovedstudien	37
7	Demografiske og bruddrelaterte karakteristika i utvalget	45
8	Deskriptiv statistikk av kontinuerlig funksjonsdata	49
9	Deskriptiv statistikk av kategorisk funksjonsdata	50
10	Tak- eller gulveffekter (eller manglende skårer)	51
11	Korrelasjonstabell mellom data fra måleinstrumentene ved bruk av parametrisk Pearsons korrelasjon	53
12	Lineær regresjon mellom data fra måleinstrumentene CMS, DASH og OSS	54
13	Logistisk regresjonsmodell for å estimere oddsen for lav CMS skår med uavhengige bakgrunnsvariabler	55



## **Datamateriale og tilknytning i forskningsprosjekt**

I masteroppgaven er det benyttet datamaterialet fra en randomisert kontrollert studie (RCT), «*Plate fixation versus intramedullary nailing of 3 and 4 part proximal humerus fractures. A prospective, double blinded randomized controlled trial*», som omhandler skulderbrudd.

Studien foregikk på Akershus Universitetssykehus (Ahus) og datainnsamlingen ble fullført juni 2023. Utvalget bestod av tre og fire-parts skulderbrudd. Hensikten med RCTen var å sammenligne to ulike operasjonsmetoder i lys av komplikasjonsrate, radiologiske funn, helseøkonomi og skulderfunksjon. Utvalget ble randomiserte til innvendig nagle eller plating. Primærutfallsmålet i RCTen var Disability of the arm, shoulder and hand (DASH).

Radiologiske funn (osteosyntesesvikt og avvaskulær nekrose), Oxford Shoulder Score (OSS), Constant -Murley Score (CMS), EuroQol-5D (EQ-5D), helseøkonomi og komplikasjonsrate var sekundærutfallsmål. Masterkandidaten samlet inn data seks måneder og to år postoperativt med pasient og klinikerrapporterte målemetoder for arm eller skulderfunksjon, helserelatert livskvalitet og helseøkonomi. Prosjektansvarlig for RCTen samlet inn data ved baseline, seks uker, tre måneder og ett år postoperativt. RCTen var dobbelblindet ved seks måneder og to år postoperativt. Prosjektansvarlig: PhD kandidat Annette Wikerøy.

# Innholdsfortegnelse

<b>Sammendrag</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>Forord</b>	<b>3</b>
<b>Begrepsavklaringer</b>	<b>4</b>
<b>Forkortelser</b>	<b>5</b>
<b>Figuroversikt</b>	<b>6</b>
<b>Tabelloversikt</b>	<b>7</b>
<b>Datamaterialet og tilknytning til forskningsprosjekt</b>	<b>8</b>
<b>1.0 Innledning</b>	<b>12</b>
1.1 Skulderbrudd	12
1.2 Behandling og målemetoder for evaluering av behandlingseffekt	12
1.3 Bakgrunn for masteroppgaven	13
<b>2.0 Teori</b>	<b>16</b>
2.1 Skulderens anatomi	16
2.1.1 Skjelett i overarmen	16
2.1.2 Skulderleddet og skulderbuen	17
2.1.3 Bevegelser i skulderen	17
2.1.4 Muskler i skulderen	18
2.2 Skulderbrudd	18
2.2.1 Klassifisering av skulderbrudd	18
2.2.2 Insidens av bruddtyper	19
2.3 Hvordan oppstår skulderbrudd blant voksne og eldre	20
2.4 Risikofaktorer for skulderbrudd	20
2.4.1 Osteoporose	20
2.4.2 Kjønn	21
2.4.3 Alder	21
2.5 Behandling av skulderbrudd	22
2.5.1 Komplikasjoner etter skulderbrudd	23
2.6 Måleinstrumenter	23
2.6.1 Psykometriske egenskaper	24
2.7 Måleinstrumenter som benyttes ved PHF	26
2.7.1 Måleinstrumentene benyttet i masteroppgaven	26
2.7.1.1 Constant Murley Score (CMS)	26
2.7.1.3 Disabilities of Arm, Shoulder and hand Questionnaire (DASH)	30

2.7.1.5	Oxford Shoulder Score (OSS)	31
2.8	Tidligere litteratur	33
2.8.1	Sammenheng mellom måleinstrumentene	33
2.8.2	Sammenheng mellom bakgrunnsvariabler og lav CMS skår	34
2.9	Teoretisk forankring	35
<b>3.0</b>	<b>Metode</b>	<b>36</b>
3.1	Datamaterialet	36
3.2	Studiedesign for masteroppgaven	36
3.3	Utvalget i masteroppgaven	36
3.4	Datainnsamling	38
3.4.1	Måleinstrumenter	38
3.4.1.1	Constant Murley skår (CMS)	38
3.4.1.2	Disabilities of Arm, Shoulder and hand Questionnaire (DASH) og Oxford Shoulder Score (OSS)	39
3.4.2	Demografiske variabler	39
3.5	Statistisk analyse	39
3.5.1	Deskriptive analyser	39
3.5.2	Sammenheng mellom data fra måleinstrumentene	40
3.5.3	Sammenheng mellom lav CMS skår og uavhengige bakgrunnsvariabler	41
3.6	Forskningsetikk	42
<b>4.0</b>	<b>Resultater</b>	<b>43</b>
4.1	Deskriptiv statistikk av demografiske variabler	43
4.2	Deskriptiv statistikk av data fra måleinstrumentene	46
4.2.1	Postoperativ funksjon målt med CMS	46
4.2.2	Postoperativ funksjon målt med DASH	48
4.2.3	Postoperativ funksjon målt med OSS	49
4.3	Tak- eller gulveffekter (eller manglende skårer) for de tre måleinstrumentene	50
4.4	Sammenheng mellom data fra måleinstrumentene	52
4.5	Sammenheng mellom lav CMS skår og uavhengige bakgrunnsvariabler	54
<b>5.0</b>	<b>Diskusjon</b>	<b>57</b>
5.1	Metodediskusjon	57
5.1.1	Studiedesign	57
5.1.2	Utvalg og data	58
5.1.3	Målemetoder	60
5.1.4	Generaliserbarhet	62
5.1.5	Statistiske analyser	64
5.1.6	Styrker og svakheter	66
5.2	Resultatdiskusjon	67

5.2.1	Deskriptiv statistikk av data fra måleinstrumentene _____	67
5.2.2	Tak- eller gulveffekter (eller manglende skårer) for de tre måleinstrumentene _____	72
5.2.3	Sammenheng mellom data fra måleinstrumentene _____	73
5.2.4	Sammenheng mellom lav CMS skår og uavhengige bakgrunnsvariabler _____	76
5.3	Kliniske implikasjoner og videre forskning .....	77
<b>6.0</b>	<b>Konklusjon</b> _____	<b>80</b>
<b>7.0</b>	<b>Referanseliste</b> _____	<b>81</b>
<b>8.0</b>	<b>Vedlegg</b> _____	<b>94</b>
	Vedlegg 1. CMS spørreskjema _____	94
	Vedlegg 2. DASH spørreskjema _____	95
	Vedlegg 3. OSS spørreskjema _____	100
	Vedlegg 4. Forutsetninger for Pearsons korrelasjon CMS og DASH _____	102
	Vedlegg 5. Forutsetninger for Pearsons korrelasjon CMS og OSS _____	102
	Vedlegg 6. Forutsetninger for Pearsons korrelasjon DASH og OSS _____	103
	Vedlegg 7. Normalfordelingstest av residualer CMS og DASH _____	103
	Vedlegg 8. Spredningsdiagram mellom funksjonsdata for DASH og CMS kun kvinner _____	104
	Vedlegg 9. Spredningsdiagram mellom funksjonsdata for DASH og OSS kun kvinner _____	104
	Vedlegg 10. Spredningsdiagram mellom funksjonsdata for CMS og OSS kun kvinner _____	105
	Vedlegg 11. Korrelasjonstabell mellom funksjonsdata kun kvinner _____	105
	Vedlegg 12. Lineær regresjonstabell mellom CMS, DASH og OSS kun kvinner _____	106
	Vedlegg 13. Logistisk regresjonstabell uten røyk _____	106
	Vedlegg 14. Godkjenning av endringsmelding fra REK _____	107

## **1.0 Innledning**

### **1.1 Skulderbrudd**

Det finnes ikke kunnskap om hvor mange som rammes av skulderbrudd i Norge, men cirka 500 personer i Akershus Universitetssykehus (Ahus) sitt opptaksområde rammes av skulderbrudd årlig (hentet fra QlikLis-applikasjonen på Ahus). I en kohortstudie utført i Oslo fant Enger et al. (2018) en insidensrate på 94 per 100 000 person- år i perioden mai 2013 til april 2014. Når skulderbruddet oppstår i den øvre enden av overarmsbenet, kalles det proximal humerus fraktur (PHF) (Mostafa, Imonugo & Varacallo, 2023). Konsekvensene på kort og lang sikt for personer med PHF og for samfunnet er betydelige (Palvanen, Kannus, Niemi & Parkkari, 2006). Personer med PHF opplever begrensninger i fysisk funksjon, aktivitet og deltakelse, og PHF utgjør en stor og gradvis økende økonomisk belastning for samfunnet (Handoll, Elliott, Thillemann, Aluko & Brorson, 2022).

Insidensen av PHF øker med alderen, og kvinner rammes to til tre ganger oftere enn menn (Launonen et al., 2015; Sumrein et al., 2017). Majoriteten som rammes er over 60 år og har benskjørhet (Handoll et al., 2022), og bruddet oppstår oftest som et resultat av fall fra egen høyde (Court-Brown, Garg & McQueen, 2001; Launonen et al., 2015). De fleste PHF er minimalt forskyvnet, men andelen med større forskyvninger og mer komplekse bruddmønstre er rapportert fra 7-64% i litteraturen (Bahrs et al., 2013; Bergdahl, Ekholm, Wennergren, Nilsson & Möller, 2016; Launonen et al., 2015). Andelen med komplekse bruddmønstre er over 50% hos pasienter over 60 år (Court-Brown et al., 2001).

PHF står for cirka 6% av alle brudd blant voksne (Court-Brown & Caesar, 2006; Handoll et al., 2022). Denne bruddtypen er den tredje vanligste blant eldre, og insidensen forventes å øke på grunn av aldring i befolkningen (Bahrs et al., 2014; Launonen et al., 2015; van de Water, Davidson, Shields, Evans & Taylor, 2016). Det har blitt anslått en tredobling i løpet av de neste 30 årene (Launonen et al., 2015).

### **1.2 Behandling og målemetoder for evaluering av behandlingseffekt**

Det finnes ulike typer PHF i forhold til grad av bruddforskyvning, og det brukes ofte klassifiseringssystemer for å bedre forståelsen av bruddmønstre og behandlingen (Marongiu

et al., 2020). For minimalt forskjøvet PHF er ikke-operativ (konservativ) behandling det vanligste behandlingsvalget med immobilisering i fatle og oppfølging av øvelser hos fysioterapeut (Handoll et al., 2022). Komplekse PHF blir i økende grad behandlet med operasjon, men det er ingen konsensus når det gjelder operasjonsmetoder. Oppfølging av øvelser hos fysioterapeut er også rutinemessig behandling etter operasjon (Handoll et al., 2022).

For å kunne vurdere behandlingseffekt i klinisk praksis og i forskning, måles symptomer og funksjon med standardiserte måleinstrumenter underveis i rehabiliteringen (Kyte et al., 2015). Måleinstrumentene kan deles opp i klinikerrapporterte (Clinician-reported outcome measurements (CROMs)) eller pasientrapporterte (Patient-reported outcome measurements (PROMs)) måleinstrumenter (Nowak et al., 2019).

### **1.3 Bakgrunn for masteroppgaven**

Det finnes ingen gullstandard når det gjelder bruk av instrumenter for å måle effekt av behandling etter PHF (Nowak et al., 2019). Det benyttes ofte en kombinasjon av CROMs og PROMs (Bahrs et al., 2010; Plath et al., 2019; Ye et al., 2013). Resultatet er at det brukes mange ulike måleinstrumenter, men ingen er diagnosespesifikke for PHF (Maurer et al., 2021; Nowak et al., 2019; Richard et al., 2020). I tidligere forskningsstudier er det i gjennomsnitt brukt 2,2 instrumenter for å måle skulderfunksjon (Richard et al., 2020). Til tross for at det er publisert mange effektstudier som omhandler behandling av PHF, er behandlingen kontroversiell. Denne kontroversen kan delvis tilskrives den store variasjonen av målemetoder som er benyttet, fordi det vanskeliggjør sammenstilling av studier i systematiske oversikter og begrenser mulighetene til å «poole» data i meta-analyser (Nowak, Davis, Mamdani, Beaton & Schemitsch, 2021; Richard et al., 2020).

Temaet for denne masteroppgaven har utspring i egen erfaring fra datainnsamlingen i RCTen. For å måle symptomer og skulderfunksjon ble det benyttet to måleinstrumenter for PROMs, i tillegg til et instrument som kombinerer PROMs med CROMs av bevegelighet og styrke i skulderen. Studiedeltakerne uttrykte stadig fortvilelse over mengden av spørsmål, og mente de hadde besvart samme spørsmål gjentatte ganger. De tre instrumentene måler nesten de samme domenene (Constant & Murley, 1987; Dawson, Fitzpatrick & Carr, 1996; Hudak, Amadio &

Bombardier, 1996). Datainnsamlingen i RCTen bød også på utfordringer knyttet til manglende data, da flere av deltakerne ikke oppfylte kriteriene til å utføre styrketesten, der kriteriene er smertefrihet og minimum elevasjon av arm til skulderhøyde. Dette resulterte i at deltakere med smerter eller med svært redusert bevegelighet i skulderen, ble vurdert til å ha dårlig styrke uten at dette var målt hos denne gruppen. Hos de som oppfylte kriteriene, var det som kliniker vanskelig å følge de standardiserte retningslinjene for utførelse av styrketesten (Constant et al., 2008). Dette var fordi det var vanskelig for studiedeltakerne å opprettholde fullt spenn i armen og unngå kompensering med aktiv bruk av tyngdeoverføring av kroppen. I tidligere forskning er det diskrepans når det gjelder psykometriske egenskaper for de tre måleinstrumentene (Nowak et al., 2019; Vrotsou et al., 2018). Andre har undersøkt sammenheng mellom måleinstrumenter hos pasienter med PHF postoperativt, og det er rapportert moderat til høy sammenheng (Maurer et al., 2021; van de Water, Shields, Davidson, Evans & Taylor, 2014; Ye et al., 2013), men basert på litteratursøk har ingen undersøkt sammenheng mellom Constant-Murley Score (CMS), The Disabilities of the arm, shoulder and hand (DASH) og Oxford Shoulder Score (OSS) seks måneder postoperativt hos tre og fire-parts PHF operert med innvendig nagle eller plating. I tillegg finnes det ingen kunnskap om måleinstrumentene benyttet på pasienter med tre og fire-parts PHF seks måneder etter plating eller nagle og hvilken funksjon det forventes hos denne populasjonen seks måneder postoperativt, da tidligere randomiserte kontrollerte studier som har sammenlignet behandlingsmetodene har utvalg med to-parts PHF (Zhu, Lu, Shen, Zhang & Jiang, 2011), to og tre-parts PHF (Gracitelli et al., 2016) og to-fire parts PHF (Plath et al., 2019).

I lys av usikre psykometriske egenskaper, manglende homogenitet når det gjelder valg av måleinstrumenter, etterlyst kunnskap om måleinstrumentene i litteraturen og erfaringsmessig stor pasientbyrde, var det ønskelig å bedre kunnskapen om måleinstrumentene benyttet for evaluering av tre og fire-parts PHF. I tillegg er det ønskelig å bedre kunnskap om seks måneder postoperativ funksjon hos pasienter med tre og fire-parts PHF operert med plating eller nagle. Hensikten med denne masteroppgaven var derfor å øke kunnskapen om postoperativ funksjon hos pasienter med tre og fire-parts PHF, og å sammenligne resultatene fra måleinstrumentene CMS, DASH og OSS, for å kunne vurdere om ett måleinstrument er bedre egnet enn andre. Dette vil kunne bidra til å redusere bruken av multiple instrumenter for å måle funksjon hos denne pasientgruppen, slik at kliniker - og pasientbyrden kan reduseres. Dersom klinikere og forskere kan enes om hvilke måleinstrumenter som skal benyttes, kan

dette gi bedre muligheter til å sammenligne resultater i forskningsstudier. På bakgrunn av erfaringsmessig større pasient og klinikerbyrde ved CROMs delen av CMS og manglende data på styrkedomenet, er det ønskelig med ytterligere kunnskap om dette måleinstrumentet for å bedre kunne vurdere instrumentets egnethet for pasienter operert for tre og fire-parts PHF.

I denne masteroppgaven vil følgende problemstillinger bli besvart:

1. Hvilken skulderfunksjon har pasienter med tre og fire-parts PHF operert med plating eller nagle seks måneder postoperativ målt med CMS, DASH og OSS?
2. I hvilken grad sees tak- eller gulveffekter (eller manglende skårer) for de tre måleinstrumentene hos pasienter med tre og fire-parts proksimal humerus fraktur seks måneder postoperativt?
3. I hvilken grad er det sammenheng mellom skulderfunksjon målt med CMS, DASH og OSS hos pasienter med tre og fire-parts proksimal humerus fraktur seks måneder postoperativt?
4. I hvilken grad er det sammenheng mellom bakgrunnsvariabler og lav skår (<56) på skulderfunksjon målt med CMS hos pasienter med tre og fire-parts proksimal humerus fraktur seks måneder postoperativt?



## 2.0 Teori

I dette kapittelet presenteres teori som belyser problemstillingene i masteroppgaven.

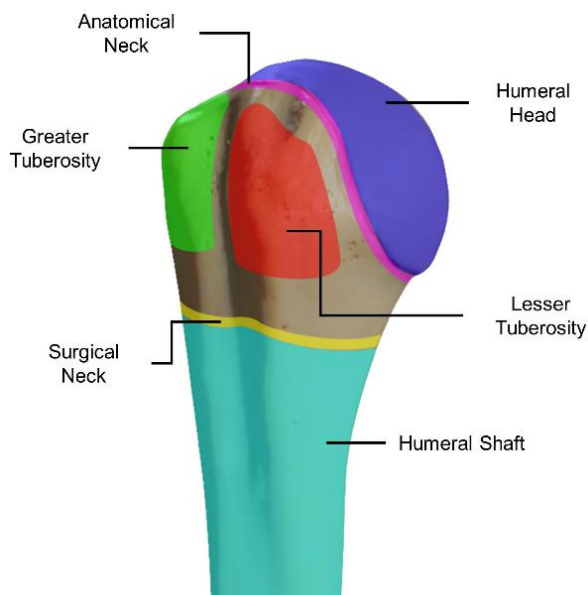
### 2.1 Skulderens anatomi

Skulderen er et sammensatt ledd, bestående av skulderbuen og skulderleddet, hvor skulderbevegelse skjer i samarbeid mellom disse to (Dahl & Rinvik, 2010, s. 407).

#### 2.1.1 Skjelett i overarmen

Overarmen består av en knokkel, humerus (Dahl & Rinvik, 2010, s. 373). Knokkelen er en lang rørknokkel, den proksimale (øvre) delen har et halvkuleformet leddhode, caput humeri. Den proksimale enden består også av collum anatomicum (den anatomiske halsen). Tuberkulum majus (den store overarmsknuten) ligger på lateralsiden (yttersiden) av proksimale humerus. Tuberkulum minus (den lille overarmsknuten) ligger distalt for (nedenfor) skulderleddet på fremsiden av humerus. Området distalt for tuberkulum minus og majus kalles collum chirurgicum (den kirurgiske halsen) (Dahl & Rinvik, 2010, s. 373-375) (Figur 1).

**Figur 1:** Anatomi av proksimale humerus inndelt i farger



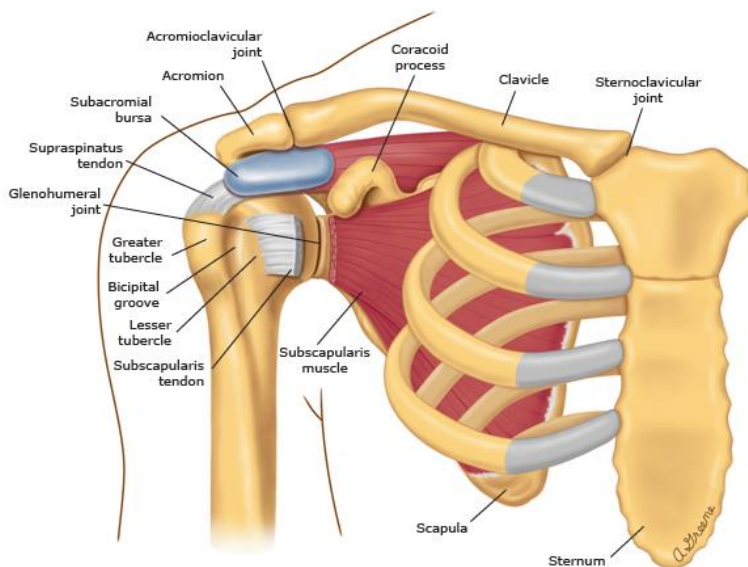
*Caput humeri (blå), Tuberkulum majus (grønn), Tuberkulum minus (rød), collum anatomicum (rosa) og collum chirurgicum (gul). Hentet fra Castro-Franco et al. (2020), Trends in the Characterization of the Proximal Humerus in Biomechanical Studies: A Review. Brukes med tillatelse fra Creative Commons CC BY License*

### 2.1.2 Skulderleddet og skulderbuen

Skulderleddet, glenohumeralleddet (GH-leddet), er et kuleledd mellom *cavitas glenoidalis* (leddskålen) på *scapula* (skulderbladet) og *caput humeri* (Dahl & Rinvik, 2010, s. 375).

Skulderbuen består av *scapula*, *clavicula* (kragebenet) og *sternum* (brystbenet) som sammen danner en bue på hver skulder (Dahl & Rinvik, 2010, s. 369-370). Leddforbindelsene i skulderbuen, *acromioclavicularleddet* (AC-leddet), *sternoclavicularleddet* (SC-leddet) og *scapulathoracalleddet* (ST-leddet) bidrar til bevegelse i skulderen i samarbeid med GH-leddet (Terry & Chopp, 2000) (Figur 2).

**Figur 2:** Knokler og leddforbindelser i skulderledd og skulderbue sett forfra



*Hentet fra Bassett (2023), Proximal humeral fractures in adults. Brukes med tillatelse fra Bassett R. Proximal humeral fractures in adults. In: UpToDate, Post TW (Ed), UpToDate, Waltham, MA. Accessed on 10/02/23. Copyright © 2023 UpToDate, Inc. and its affiliates and/or licensors. All rights reserved*

### 2.1.3 Bevegelser i skulderen

Skulderleddet er kroppens mest bevegelige ledd (Dahl & Rinvik, 2010, s. 376). Leddet har tre bevegelsesbaner: fleksjon (frem overføring) - og ekstensjon (bak overføring), abduksjon (ut overføring) og adduksjon (inn overføring) og utover- og innover rotasjon. En kombinasjon av alle disse bevegelsene skaper sirkumduksjon. Normalt bevegelsesutslag (ROM; *range of motion*) i skulderleddet er fleksjon 150-180 grader, ekstensjon fra 40-60 grader, abduksjon fra

150-180 grader, utover rotasjon fra 60-90 grader og innover rotasjon fra 50-70 grader (Bakhsh & Nicandri, 2018).

#### **2.1.4 Muskler i skulderen**

Brystet og ryggens ekstremitetsmuskler går fra trunkus (overkroppen) til skulderbuen og de proksimale delene av humerus (Dahl & Rinvik, 2010, s. 391). Musklene bidrar til bevegelse av skulderleddet og- eller skulderbuen (Dahl & Rinvik, 2010, s. 393-398). Skuldermusklene går fra skulderbuen til proximale humerus og beveger skulderleddet (Dahl & Rinvik, 2010, s. 391). Skuldermusklene supraspinatus, infraspinatus, subscapularis og teres minor utgjør rotatormansjetten, og samlet har de nevnte musklene stor betydning for stabiliteten i skulderleddet (Dahl & Rinvik, 2010, s. 400). Rotatormansjetten fester seg til områdene som omfatter PHF (se punkt 2.2).

## **2.2 Skulderbrudd**

PHF inkluderer brudd i områdene collum anatomicum, collum chirurgicum, tuberkulum majus og tuberkulum minus (Mostafa et al., 2023). Ved mistanke om PHF, benyttes billeddiagnostikk med røntgen i ulike plan, og suppleres med Computertomografi (CT) med tredimensjonale bilder for ytterligere informasjon hvis det er nødvendig (Robinson, Athwal, Sanchez-Sotelo & Rispoli, 2008).

### **2.2.1 Klassifisering av skulderbrudd**

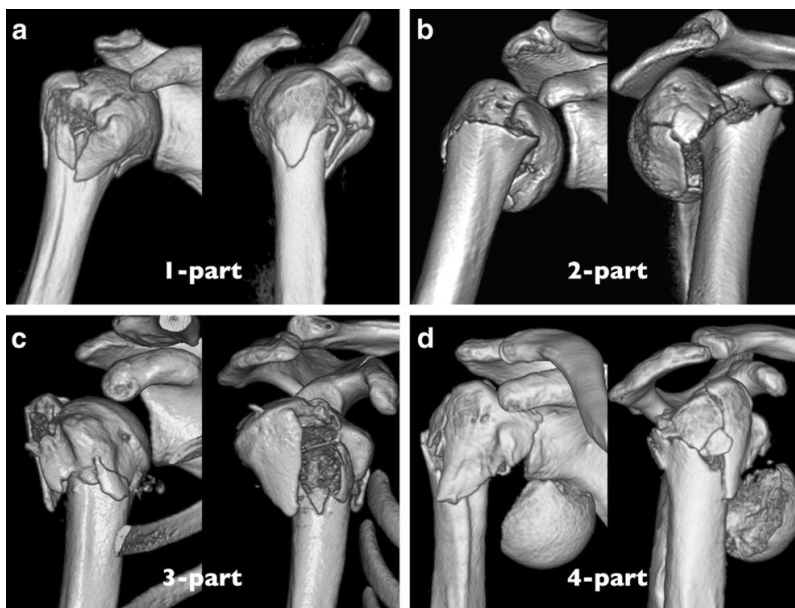
For å bedre forståelsen og behandlingen av PHF er ulike bruddklassifiseringssystemer rutinemessig benyttet i klinisk praksis (Marongiu et al., 2020).

Neer og Arbeitsgemeinschaft für Osteosynthesefragen/Orthopaedic Trauma Association (AO/OAT) klassifisering er mest brukt for klassifisering av PHF (Marongiu et al., 2020). PHF kan grovt deles inn i forskjøvet og ikke- forskjøvet brudd. I masteroppgavens datamateriale er Neer klassifiseringen benyttet. Neer klassifiseringen har en høyere klinisk og vitenskapelig aksept enn AO/OTA-klassifiseringen (Bahrs et al., 2013), og er mest benyttet (Handoll et al., 2022; Nowak et al., 2017).

I 1970 presenterte Neer sitt klassifiseringssystem (Neer, 1970). Basert på røntgenbilder av høy kvalitet beskrev han minimalt forskjøvet frakturer som mindre enn 45 grader vinkling mellom hovedsegmentene eller mindre enn en cm forskyvning av segmentene. Forskjøvet

frakturer beskrev han som mer enn en cm forskyvning og- eller mer enn 45 grader vinkelforskyvning. Proximale humerus ble delt inn i fire hovedsegmenter, en, to, tre eller fire-part. De potensielle segmentene som er involvert er tuberkulum major, tuberkulum minor, caput humeri og humerusskaftet. En-parts er minimalt forskjøvet. Ved to, tre eller fire-parts PHF er det forskyvning (mer enn en cm) eller vinkling (mer enn 45 grader) mellom noen eller alle segmentene (Neer, 1970). I 2002 ble valguspåvirkede fire parts frakturer inkludert i en revidert versjon av Neer klassifiseringen (Neer, 2002). Det beskrives 16 bruddkategorier i klassifikasjonen (Neer, 2002) (Figur 3).

**Figur 3:** Neer klassifisering illustrert med CT bilder



*a:* en-parts fraktur. Caput humeri holder seg samlet med de andre segmentene. *B:* to-parts fraktur. Caput humeri og tuberkulum major er rotert i forhold til humerusskaftet. *C:* tre-parts fraktur. Caput humeri er forskjøvet fra humerusskaftet og tuberkulum majus. *D:* fire-parts fraktur. Caput humeri er helt forskjøvet fra de tre andre segmentene. Hentet fra Matsumura et al. (2020), Reproducibility of the modified Neer classification defining displacement with respect to the humeral head fragment for proximal humeral fractures, brukes med tillatelse fra Creative Commons CC BY License

### 2.2.2 Insidens av bruddtyper

En-parts PHF er vanligst blant yngre pasienter, og reduseres med alderen, i motsetning til alle andre bruddtyper i samme segment (Bergdahl et al., 2016). Andel en-parts PHF er rapportert mellom 13 -51% i litteraturen (Bahrs et al., 2013; Bahrs et al., 2014; Court-Brown et al., 2001; Launonen et al., 2015). Andelen forskjøvet PHF er ulikt rapportert i litteraturen. To-parts frakturer er rapportert fra 29-64%, tre-parts frakturer er rapportert 19-44 % og tre og

fire-parts frakturer 7-11 % (Bahrs et al., 2013; Bergdahl et al., 2016; Launonen et al., 2015). Tre og fire-parts var hyppigere rapportert hos kvinner (28%) enn hos menn (18%) i Launonen et al. (2015). I en retrospektiv registerstudie med 815 personer fra Tyskland, ble det rapportert 86% to-fire partsfrakturer (Bahrs et al., 2014). Totalt 49% var tre-fire partsfrakturer. Andelen tre og fire partsfrakturer økte ikke i perioden 1997-2011 (Bahrs et al., 2014).

## **2.3 Hvordan oppstår skulderbrudd blant voksne og eldre**

Majoriteten av PHF oppstår etter fall fra egen høyde (lavenergitraume) (Court-Brown et al., 2001). Fall fra egen høyde omtales ofte som lavenergitraume, og brudd som oppstår etter lavenergitraume (eller mindre) kalles fragile brudd (Kanis et al., 2001). PHF er rapportert i 59-97 % av tilfellene etter lavenergitraume (Bahrs et al., 2013; Bahrs et al., 2014; Bergdahl et al., 2016; Launonen et al., 2015; Palvanen et al., 2000). Andel PHF oppstod hos 2-10% etter fall fra større høyder og 4-6% i trafikkulykker (Bergdahl et al., 2016; Launonen et al., 2015). Skademekanisme er relatert til alder. Høyenergitraume hos yngre pasienter og lavenergitraume hos eldre pasienter (Launonen et al., 2015). Kun 20% av tre og fire-partsfrakturer oppstod hos yngre menn etter høyenergitraumer og majoriteten av tre og fire-partsfrakturer oppstod hos eldre kvinner etter lavenergitraumer (Bahrs et al., 2014). Palvanen et al. (2000) fant at 54% hadde falt utendørs. I Sverige fant de en økt insidens av PHF i vintermånedene (desember-februar), som trolig kan forklares ved glatt føre utendørs (Bergdahl et al., 2016). Bahrs et al. (2013) gjorde tilsvarende funn i en registerstudie fra Tyskland. Insidensen var økt mellom november og mars også i Finland (Launonen et al., 2015).

## **2.4 Risikofaktorer for skulderbrudd**

### **2.4.1 Osteoporose**

Osteoporose er en systemisk skjelettsykdom som gir redusert benmasse og bentetthet, og derfor påfølgende økt risiko for brudd (Peck et al., 1993; van Oostwaard, 2018). Det er sjelden symptomer på osteoporose før det eventuelt oppstår et lavenergibrudd (Cosman et al., 2014). Kvinner rammes av osteoporose tre ganger hyppigere enn menn (Kanis et al., 2021). Diagnosen kan stilles klinisk ved påvist lavenergibrudd, eller ved bruk av en radiologisk undersøkelse, dual-energy x-ray absorptiometri (DEXA) (Peck et al., 1993). Primær osteoporose er mest utbredt og oppstår som følge av overgangsalder og aldringsprosesser

(Peck et al., 1993). Risikofaktorer for utvikling av osteoporose er blant annet østrogenmangel i overgangsalder hos kvinner, høy alder, lavt inntak av kalsium og vitamin D i kosten, lav BMI, alkoholmisbruk, røyking, lavt aktivitetsnivå og immobilisering, arv, komorbiditet og generelt dårlig helse (Cosman et al., 2014). Osteoporose er en sterk risikofaktor for brudd, og 30% av alle postmenopausale kvinner i Europa og USA har osteoporose (Bahrs et al., 2014). Gjennom hele livet vil 50% av kvinner over 50 år gjennomgå brudd (Bahrs et al., 2014). Osteoporose er en faktor som påvirker insidensen av PHF (Bahrs et al., 2013; Sumrein et al., 2017). PHF oppstod ved lavenergitraume hos 86 % over 50 år, som kan bety at PHF er relatert til osteoporose (Bergdahl et al., 2016).

#### **2.4.2 Kjønn**

I studier er det rapportert høyest forekomst av PHF blant kvinner. Det er rapportert fra 67-90% (Bahrs et al., 2014; Bergdahl et al., 2016; Clement, Duckworth, McQueen & Court-Brown, 2014; Fraser et al., 2020; Gracitelli et al., 2016; Olerud, Ahrengart, Ponzer, Saving & Tidermark, 2011; Plath et al., 2019; Sumrein et al., 2017). Det er rapportert en økende andel tre og fire-parts PHF blant kvinner over 60 år (Bahrs et al., 2013; Bahrs et al., 2014).

#### **2.4.3 Alder**

Sumrein et al. (2017) fant at gjennomsnittsalder på skadetidspunkt var 62 år for menn og 72 år for kvinner. Insidensen økte med alderen, størst økning etter fylte 50 år. Insidensen var 105.6 per 100 000 person-år 50-59 år, 181.9 per 100 000 person-år 60-69 år og 380.5 per 100 000 person-år ved fylte 70 år eller eldre (Sumrein et al., 2017). Kim et al. (2020) rapportert signifikant høyere sannsynlighet for forskjøvet PHF hos kvinner og alder over 70 år (uavhengig av kjønn). Sannsynligheten for forskjøvet PHF økte med 1,1 for hvert økte leveår, og 3,8 høyere sannsynlighet for kvinner sammenlignet med menn. Bergdahl et al. (2016) fant at gjennomsnittsalderen for kvinner på skadetidspunkt var 70,5 år og 61.8 år for menn.

#### **2.4.4 Andre risikofaktorer**

Falltendens, skrøpeligheit og medisinske tilstander (epilepsi, depresjon og diabetes) ble identifisert som risikofaktorer (Chu et al., 2004). Bahrs et al. (2013) og Bahrs et al. (2014) fant signifikant sammenheng mellom høy komorbiditet og falltendens når det gjaldt PHF.

Thorin, Wihlborg, Åkesson og Gerdhem (2016) fant økt risiko for PHF hos eldre kvinner som røyket eller hadde røyket tidligere. Andre kjente risikofaktorer for fragile brudd generelt er lys hud, kognitiv svikt, lav Body Mass Index (BMI), røyking og alkoholmisbruk (Friedman & Mendelson, 2014).

Økende insidensrate over 50 år, for alle kjønn, viser påvirkningen av aldersrelaterte risikofaktorer som osteoporose og større risiko for fall i den eldre befolkningen (Bergdahl et al., 2016).

## **2.5 Behandling av skulderbrudd**

Behandling av PHF er kontroversiell og blant de mest omdiskuterte av alle bruddbehandlinger (Bergdahl et al., 2016), og det er rapportert diskrepans mellom klinisk praksis og nåværende kunnskap om behandling (Sumrein et al., 2017).

Behandlingsvalg kan vurderes ut ifra bruddtype, pasientens aktivitetsnivå, generelle helsetilstand, og effekt og begrensninger ved behandlingsmetodene som er tilgjengelig (Neer, 2002). American Society of Anesthesiologists Physical Status scoring system (ASA skår) benyttes ofte for klassifisere pasientens generelle helsestatus (basert på grad av komorbiditet) før eventuell operasjon (Bjorgul, Novicoff & Saleh, 2010) og har vist seg og kunne måle risiko for intraoperative og postoperative ikke-ortopediske komplikasjoner (Perka, Arnold & Buttgerit, 2000).

Konservativ (ikke-operativ) behandling, er det vanligste behandlingsvalget for en-part PHF, men velges også noen ganger hos eldre med to-fire parts PHF (Handoll et al., 2022).

Konservativ behandling innebærer vanligvis at armen immobiliseres i fatle i en uke eller lengre, deretter oppstart av fysioterapi for øvelser.

To-fire parts PHF behandles oftere med operasjon (Handoll et al., 2022). Det finnes ulike operasjonsmetoder, men de vanligste er innvendig fiksering med plating eller innvendig nagling. Det benyttes også totalproteser (erstatning av både caput humeri og cavitas glenoidalis), ofte reverserte (omvendt) proteser i bruddbehandling (Handoll et al., 2022).

Sumrein et al. (2017) fant at antall PHF operasjoner i 2012 var 75% høyere enn 2001.

Operasjonsraten økte mest i populasjonen over 50 år. Det er en økning i antall operasjoner av to-fire parts PHF uten kjent årsak, da det ikke er noe evidens som støtter dette (Beks et al., 2018; Launonen et al., 2015; Rangan et al., 2015; Sumrein et al., 2017). En randomisert

kontrollert studie fra Norge av Fraser et al. (2020) fant at reversert skulderprotese ga bedre funksjon enn plating to år postoperativt hos 104 tre og fire-parts PHF over 65 år. En Cochrane-oversikt av Handoll et al. (2022) fant ingen forskjell i funksjon mellom operasjon eller konservativ behandling av to-fire parts PHF blant eldre.

### **2.5.1 Komplikasjoner etter skulderbrudd**

Det er en økt risiko for død blant eldre som pådrar seg PHF, spesielt den første måneden, men risikoen vedvarer minst ett år og er størst blant menn (Bergdahl, Wennergren, Ekelund & Möller, 2020). Clement, Duckworth, et al. (2014) rapporterte at 10 % (61 av 637) døde i løpet av det første året, og de som ikke var hjemmeboende, hadde størst risiko.

Rehabiliteringen tar flere måneder, og dårlig skulderfunksjon og smerte er vanlig komplikasjoner uavhengig av behandlingsmetode (Hodgson, Mawson, Saxton & Stanley, 2007). Mange, uavhengig av bruddtype, rapporterer noe eller stor funksjonsnedsettelse etter to år (Hodgson et al., 2007). Clement, Duckworth, et al. (2014) fant at mer enn en fjerdedel blant opererte og konservativt behandlede PHF hadde dårlig funksjon målt med CMS etter ett år uavhengig av alder. Blant mange eldre vil PHF begrense muligheten til å utføre dagligdagse aktiviteter.

Komplikasjoner rapportert postoperativt er infeksjon, nerveskade, avaskulær nekrose (dødt benvev som følge av tapt blodforsyning), osteosyntesevikt, manglende eller ingen bruddtilheling, stivhet og redusert funksjon i skulderen, kronisk regionalt smertesyndrom og smerter (Rangan et al., 2015). Noen av komplikasjonene vil kreve reoperasjon.

Komplikasjoner rapportert ved konservativ behandling er avaskulær nekrose, manglende eller ingen bruddtilheling, stivhet, redusert funksjon i skulderen og smerte (Rangan et al., 2015). Også ved konservativ behandling kan noen komplikasjoner være operasjonskrevende.

Det er funnet høyere komplikasjonsrate og mer alvorlige komplikasjoner hos opererte PHF sammenlignet med de som er behandlet konservativt (Beks et al., 2018; Handoll et al., 2022).

## **2.6 Måleinstrumenter**

PROMs er standardiserte instrumenter som måler pasientens opplevelse av egen helse, funksjon og- eller symptomer. PROMs kan også måle psykososiale faktorer (Jayakumar et al.,



2018). CROMs er standardiserte instrumenter hvor kliniker måler pasientens funksjon eller helsetilstand, som for eksempel radiologiske undersøkelser, bevegelsesutslag og styrkemål. PROMs og CROMs måler ofte ulike domener ved et eller flere fenomener.

Måleinstrumentene kan være generiske, diagnose eller regionspesifikke (van de Water, Shields & Taylor, 2011).

### **2.6.1 Psykometriske egenskaper**

Datamaterialets kvalitet er avhengig av målemetodene som benyttes i innsamlingen (Pallant, 2020, s. 3). I utvelgelsen av målemetoder er det viktig å kartlegge at metoden måler den påståtte behandlingseffekten og er egnet for den tiltenkte pasientgruppen for å unngå systematiske måleskjevheter (U.S. Department of Health, 2006). Det samme gjelder i klinisk praksis, for å måle pasientens funksjon i rehabiliteringen må metoden måle det som er hensikten og være egnet for den tiltenkte pasientgruppen (L. Mokkink, Terwee & de Vet, 2021). Dette vurderes ut ifra psykometriske egenskaper som validitet, reliabilitet og responsivitet som også avhenger av hvilken kontekst og hvilken populasjon som måles, og kan derfor variere mellom populasjoner (Nowak et al., 2019; van de Water et al., 2011). Hvis målemetodene ikke har tilstrekkelig psykometriske egenskaper er det stor risiko for skjevhet i resultater som kan føre til feil konklusjoner (L. B. Mokkink et al., 2010a).

#### **2.6.1.1 Validitet, reliabilitet og responsivitet**

Validitet innebærer i hvilken grad metoden måler det som er hensikten å måle (L. B. Mokkink et al., 2010b; Pallant, 2020, s. 7). Validitet deles inn i kategoriene innholdsvaliditet, kriterievaliditet, og konstruktvaliditet. Innholdsvaliditet er hvilken grad metoden måler alle dimensjoner av det fenomenet en ønsker å måle. Kriterievaliditet er hvilken grad metoden samsvarer med en sann verdi eller «gullstandard». Konstruktvaliditet er hvilken grad metoden samsvarer med andre relaterte (konvergent validitet) og ikke-relaterte domener (diskriminant validitet) (L. B. Mokkink et al., 2010b; Pallant, 2020). Det finnes ingen «gullstandard» når det gjelder målemetode for funksjonsevaluering av PHF (Richard et al., 2020).

Reliabilitet er hvilken grad metoden måler nøyaktig (L. B. Mokkink et al., 2010b; Pallant, 2020). Test-retest reliabilitet evaluerer sammenhengen på målinger gjort i samme kontekst, men ved ulike tidspunkter. Intern konsistens evaluerer evnen til å skape de samme resultatene ved å bruke ulike utvalg til å måle et fenomen på samme tidspunkt. Interrater reliabilitet

evaluerer målenøyaktighet mellom ulike personer i samme kontekst (L. B. Mokkink et al., 2010b).

Responsivitet er målemetodens evne til å oppdage endring av fenomenet som måles over tid (de Vet et al., 2006; L. B. Mokkink et al., 2010b).

### **2.6.1.2 Tolkbarhet**

Tolkbarhet er ikke ansett som en psykometrisk egenskap, men viktig karakteristikk av en målemetode og innebærer tolkning av endringen i skår (L. B. Mokkink et al., 2010b). Statistisk signifikant bedring etter behandling i studier betyr ikke nødvendigvis at endringen har en klinisk verdi eller er meningsfull for pasienten (Copay, Subach, Glassman, Polly & Schuler, 2007; Dabija & Jain, 2019) og det er derfor viktig å gjøre hensiktsmessige vurderinger om endringen eller forskjellen er av klinisk betydning (Rodrigues, 2020).

En målestokk for klinisk relevans av endring i funksjon er måleinstrumentets minste viktige endring (MIC; *Minimal important change*) (de Vet et al., 2006). Endring i skår over MIC verdien er klinisk viktig endring. Måleinstrumentets minste målbare endring (MDC; *Minimal detectable change*) er grad av endring nødvendig for at endringen er sann, hvor det er tatt høyde for standard målefeil (SEM; *Standard error of mean*) ved instrumentet (de Vet et al., 2006). Minste kliniske viktige forskjell (MCID; *Minimal clinical important difference*) er terskelen på minste forskjell i skår som er meningsfull eller viktig for pasienten (Copay et al., 2007; Dabija & Jain, 2019). Identifisering av MCID er spesielt viktig for tolkning av resultater hvor måleinstrumentene er PROMs (Copay et al., 2007). Det er vist at MCID verdier ved samme måleinstrument varierer mellom ulike skulderlidelser, derfor er det viktig at den spesifikke diagnosen tas i betraktning når man rapporterer og tolker MCID (Dabija & Jain, 2019). En absolutt MCID er kanskje ikke aktuell for en heterogen populasjon med varierende diagnoser (Dabija & Jain, 2019).

MIC, MCID og MDC er alle viktige målestokker for tolkning av endring eller forskjell i skår på et måleinstrument (de Vet et al., 2006). Tolkningen kan være kontekstavhengig da det kan være variasjon mellom tilstander, populasjoner og behandlinger (Rodrigues, 2020).

Tak- eller gulveffekt ansees å være tilstede hvis mer enn 15% av utvalget oppnår henholdsvis lavest eller høyest mulig skår på et måleinstrument (McHorney & Tarlov, 1995).

### **2.6.1.3 Pasientbyrde ved måleinstrumenter**

De fleste forskere opplever utfordringer med å balansere behovet for å samle inn tilstrekkelig data til å besvare problemstillingen, og samtidig unngå stor pasientbyrde (Polit & Beck, 2017, s. 175). I mange studier samles det inn mer data enn nødvendig (Polit & Beck, 2017, s. 175). Eventuell pasientbyrde i forskning er en helhetlig sammensetning av flere faktorer, blant annet gjennomføring av ulike målemetoder (Bodart, Byrom, Crescioni, Eremenco & Flood, 2019). Det er vist signifikant sammenheng mellom antall spørsmål på PROMs og grad av manglende data (Rolstad, Adler & Rydén, 2011). Manglende data kan skape utfordringer i analysene, påvirke tolkbarheten og verdien av funnene fra studien (Bodart et al., 2019).

## **2.7 Måleinstrumenter som benyttes ved PHF**

I en systematisk oversikt av Richard et al. (2020) fant forfatterne at det ble brukt 22 ulike måleinstrumenter for pasienter med PHF. CMS og DASH var mest benyttet, OSS, American Shoulder and Elbow Surgeons (ASES), Neer score, QuickDASH, og Singel Assesment Numeric Evaluation (SANE) ble også ofte rapportert (Richard et al., 2020). I en systematisk oversikt av van de Water et al. (2011) fant forfatterne 17 ulike måleinstrumenter, de mest brukte var CMS, DASH og Neer, etterfulgt av OSS, UCLA (University of California, Los Angeles Shoulder Score) og ASES. En nyere systematisk oversikt av Padua, de Girolamo, Grassi og Cucchi (2021) fant bruk av 11 ulike PROMs, hvorav DASH og ASES var hyppigst brukt. CMS ble ikke inkludert da det ikke ble klassifisert som et PROMs.

### **2.7.1 Måleinstrumentene benyttet i masteroppgaven**

Instrumentene benyttet for å måle skulderfunksjon var CMS, DASH og OSS.

#### **2.7.1.1 Constant Murley Score (CMS)**

CMS ble presentert i 1987 som et skulderspesifikt instrument for å måle fysisk funksjon i skulder (Constant & Murley, 1987). CMS er delvis PROMs og delvis CROMs, men klassifisert ofte som CROMs i litteraturen (Padua et al., 2021). CMS måler fire ulike domener relatert til skulderfunksjon. De to PROMs domenene måler smerte og ADL (ADL; *activities of daily living*) og de to CROMs domenene måler ROM og styrke. I PROMs domenene er maksimale skår 35 (ca. 33%) og i CROMs domenene er maksimale skår 65 (ca. 67%), som

resulterer i en maksimal skår på 100 som er best mulig målte skulderfunksjon (Constant & Murley, 1987).

Smertedomenet kan skåres opp til 15 poeng, og er målt ved PROMs gradering på en VAS (VAS; *Visual analog scale*) skala av de mest uttalte smertene ved normal aktivitet i løpet av et døgn og ett spørsmål om grad av smerte på en Likert-skala, ingen, milde, moderate eller sterke ved normale aktiviteter (Constant et al., 2008). ADL-domenet er PROMs og kan skåres opp til 20 poeng. Uforstyrret søvn gir to poeng, sporadisk forstyrret søvn gir ett poeng og forstyrret søvn hver natt gir null poeng. Jobb og fritidsaktiviteter er tildelt åtte poeng, og skåres fra en-fire. Ett tilleggselement målt i ADL-domene er evne til å bruke armen uten smerte fra under midjen (null poeng) til over hodet (10 poeng).

I Constant et al. (2008) retningslinje for CROMs delen, gir ROM-domenet maksimalt 40 poeng og er delt likt mellom fleksjon, abduksjon, funksjonell utover rotasjon og funksjonell innover rotasjon i skulderen. Bevegelsene må være smertefrie og aktive. Fleksjon og abduksjon måles med metall goniometer 0-180 grader (Figur 4). Referansepunkt for goniometeret er armens akse og processus spinosus i brystryggen (Figur 5). Når det gjelder funksjonell innover og utover rotasjon gis to poeng for hver bevegelse oppnådd, maksimalt 10 poeng. Funksjonell utover rotasjon måles ved oppnåelse av hånd bak hode med albue fremover, hånd bak hode med albue bakover, hånd på toppen av hode med albue fremover, hånd på hode med albue bakover og full elevasjon av arm. Innover rotasjon måles på samme måte ved bruk tommelen som peker mot de anatomiske landemerkene; bak setet, iliosakralledet, midjen, den tolvte brystgyrvirvelen og mellom skulderbladene (Constant et al., 2008).

**Figur 4:** Metall goniometer (bilde tatt av Kari Krokan, 2023)



**Figur 5:** Utførelse av måling av fleksjon i skulderen (bilde tatt av Kari Krokan, 2023)



Styrkedomenet er CROMs og gir maksimalt 25 poeng (Constant et al., 2008). Utførelsen av styrkemålingen ble ikke spesifisert i 1987, men i 2008 publiserte Constant et al. (2008) en modifisert retningslinje for bruk av instrumentet hvor utførelsen av styrkemålingen ble forsøkt spesifisert. Styrkemålingen gjøres i 90 grader abduksjon i scapulas plan, underarm er pronert med håndgrep på fjærvekt eller Isobex med armen i fullt spenn (Figur 6). Pasienter som ikke oppnår 90 grader skulderabduksjon oppnår null poeng (Constant et al., 2008).

**Figur 6:** Utførelse av styrkedomene CMS (bilde tatt av Kari Krokan, 2023)



I masteroppgaven ble fjærvekt Berkley Precision Digital Scale – 35 LB/16KG, modell BT35DS benyttet (Figur 7). Styrkemålingen ble utført i 5 sekunder og 5 repetisjoner, og resultatene fra målingen oppgis i kilo og deretter omregning til poeng.

**Figur 7:** Fjærvekt Berkley Precision Digital Scale (bilde tatt av Kari Krokan, 2023)



CMS er det mest brukte instrumentet for å måle effekt av behandling etter PHF (Nowak et al., 2021; Richard et al., 2020; van de Water et al., 2011). I lys av ICF (ICF; *International Classification of functioning, disability and health*) måler CMS kroppsfunksjoner og strukturer (80%), og begrensninger i deltakelse (20%) (Nowak et al., 2021). CMS skår 86-100 ansees som svært god skår, 71-85 ansees som god skår, 56-70 ansees som moderat skår og under 56 ansees som dårlig skår (Bahrs et al., 2010).

### 2.7.1.2 Psykometriske egenskaper ved CMS

Det er diskrepans mellom resultatene i systematiske oversikter som omhandler psykometriske egenskaper ved CMS benyttet hos PHF. Vrotsou et al. (2018) rapporterte moderat reliabilitet, lav validitet og akseptabel responsivitet. Nowak et al. (2019) rapporterte god reliabilitet, akseptabel validitet og lav responsivitet. van de Water et al. (2011) fant ikke tilstrekkelig psykometrisk data i inkluderte studier. Bruk av CMS har vist høy pasient og klinikerbyrde (Nowak et al., 2019). CMS er ikke oversatt eller validert til norsk. MCID ved PHF er fra 5,4 (distribusjonsbasert) til 11,6 (ankerbasert) (Nowak et al., 2019; van de Water et al., 2014), men i intervensjonsstudier som omhandler PHF rapporteres oftest MCID 10 som er utregnet for pasienter med rupturer i rotatormansjetten (Kukkonen, Kauko, Vahlberg, Joukainen & Äärimaa, 2013). Se normative gjennomsnittsverdier ved CMS i tabell 1.

**Tabell 1:** Normative gjennomsnittsverdier CMS (Constant, 1986)

Kjønn/alder	Gjennomsnitt	SD
<b>Kvinner</b>		
41-50	80	3,8*
51-60	73	2,8*
61-70	70	4*
71-80	69	3,9
<b>Menn</b>		
41-50	92	3,6*
51-60	90	3,1*
61-70	83	4,2*
71-80	75	3,6

\*Signifikant forskjell mellom kjønn  
SD, standardavvik

### 2.7.1.3 Disabilities of Arm, Shoulder and hand Questionnaire (DASH)

DASH er et regionspesifikt PROM og ble presentert i 1996 (Hudak et al., 1996). Hensikten er å måle fysisk funksjon og symptomer den siste uken hos pasienter med muskelskjelettskade i arm, skulder og hånd. DASH inneholder 30 elementer som skåres med en Likert skala fra en-fem rettet mot symptomer og vansker med å utføre ulike fysiske og sosiale aktiviteter og åtte tilleggselementer relatert til arbeid, musikk og idrett. Symptomene smerte, svakhet, stivhet og prikking i armen blir målt. Fysiske funksjoner som dagligdagse aktiviteter, husarbeid, innkjøp, fritidsaktiviteter, personlig hygiene og påkledning, spising, seksuell aktivitet, søvn og eventuelt arbeid, musikk og idrett blir målt (Hudak et al., 1996). DASH gir en sumskår fra 0-100, hvor 0 er beste skår (van de Water et al., 2014). Domene relatert til dagligdagse aktiviteter inneholder 21 elementer, domene relatert til symptomer inneholder seks elementer og domene relatert til sosial funksjon inneholder tre elementer (van de Water et al., 2014). DASH er validert og oversatt til norsk (Finsen, 2008). DASH skår kan kategoriseres til svært god (0-24 poeng), god (25-49 poeng), moderat (50-74) og dårlig (75-100) funksjon (Mattyasovszky et al., 2011). DASH er et av de mest brukte instrumentene til å måle effekt av behandling etter PHF (Nowak et al., 2019; Padua et al., 2021; Richard et al., 2020; van de Water et al., 2011). I lys av ICF måler DASH kroppsfunksjoner og strukturer (23%), aktivitetsbegrensninger (57%) og deltakelsesbegrensning (20%) (Nowak et al., 2021).

#### 2.7.1.4 Psykometriske egenskaper ved DASH

Når det gjelder psykometriske egenskaper ved DASH for pasienter med PHF er det diskrepans mellom resultatene. Det er rapportert moderat til sterk konstruktvaliditet og god responsivitet (Slobogean, Noonan & O'Brien, 2010; van de Water et al., 2014), mens en annen studie rapporterte god reliabilitet, moderat validitet og akseptabel responsivitet (Nowak et al., 2019). Bruk av DASH har vist lav pasient og klinikerbyrde (Nowak et al., 2019). MCID er rapportert til å være 10,83 hos pasienter med generell muskel og skjelettlidelse i overekstremiteter (Franchignoni et al., 2014). Normativ gjennomsnittlig DASH skår hos den generelle befolkningen i Norge vises i tabell 2 (Aasheim & Finsen, 2014).

**Tabell 2:** Normative gjennomsnittsverdier DASH (Aasheim & Finsen, 2014).

Kjønn/Alder	Gjennomsnitt	SD	95% KI
<b>Kvinner</b>			
40-49	11	16	12-15
50-59	12	16	9-16
60-69	18	18	15-22
70-70	22	21	17--26
<b>Menn</b>			
40-49	9	16	5-12
50-59	12	14	7-13
60-69	11	16	8-14
70-79	13	15	9-16

SD, standardavvik  
KI, konfidensintervall

#### 2.7.1.5 Oxford Shoulder Score (OSS)

Oxford Shoulder Score (OSS) er et skulderspesifikt pasientrapportert måleinstrument og ble presentert i 1996 (Dawson et al., 1996). Hensikten er å måle postoperativt funksjon i skulder de siste fire ukene, med unntak av stabiliserende skulderoperasjoner (Dawson et al., 1996). OSS inneholder 12 elementer som skåres med en Likert skala fra 0-4 rettet mot symptomer og vansker med å utføre ulike dagligdagse aktiviteter, der fire representerer ingen vansker eller symptomer (Dawson, Rogers, Fitzpatrick & Carr, 2009). OSS måler smerte, selvstendighet i personlig ADL (bruk av kniv og gaffel, påkledning, handling, hårstell, personlig hygiene),



søvnforstyrrelser og eventuelle arbeidsbegrensninger. Tidligere skåret OSS fra 12-60, hvor 60 representerte dårligste skår, men denne skåring ble endret i 2009. Modifisert OSS skårer fra 0-48, hvor 48 er best mulig oppnådde skår (Dawson et al., 2009). Det er ikke publisert en kategorisert inndeling av OSS skårer. OSS er mye brukt for å måle effekt av operasjon etter PHF (van de Water et al., 2014).

#### **2.7.1.6 Psykometriske egenskaper OSS**

OSS er oversatt og validert på norsk (Fjalestad, Iversen, Hole, Smedsrud & Madsen, 2014). Bruk av OSS har vist adekvat pasient og klinikerbyrde (Nowak et al., 2019) og har vist lav andel manglende data (Dawson et al., 2009). I lys av ICF måler OSS kroppsfunksjoner og strukturer (25%), aktivitetsbegrensninger (67%) og deltakelsesbegrensning (8%) (Nowak et al., 2021).

Når det gjelder psykometriske egenskaper har OSS har vist god reliabilitet, tilfredsstillende validitet og responsivitet hos pasienter med PHF (Nowak et al., 2019). Uavhengig av diagnose eller operasjonsmetode forventes lavere skår enn 48 hos eldre pasienter (Dawson et al., 2009). MCID ved PHF er 11,4 ved ankerbasert og 5,1 ved distribisjonsbasert utregning (Nowak et al., 2019; van de Water et al., 2014). Det finnes ikke normative gjennomsnittsverdier for den modifiserte OSS skåren. Se normative gjennomsnittsverdier ved bruk av den ikke-modifiserte skåringen i tabell 3.

**Tabell 3:** Normative gjennomsnittsverdier OSS ved bruk av den ikke-modifiserte skåringen av OSS fra 12-60, hvor 60 er dårligste skår (Clement & Court-Brown, 2014)

Kjønn/alder	Gjennomsnitt	SD
<b>Kvinner</b>		
41-50	14,6	2,9
51-60	18,3	3,7
61-70	20,3	3,8*
71-80	20,4	4,9
<b>Menn</b>		
41-50	13,6	2,2
52-60	16,9	4,1
61-70	17,8	4,1*
71-80	17,5	5,0

SD, standardavvik

\*Signifikant forskjell mellom kjønn

## 2.8 Tidligere litteratur

### 2.8.1 Sammenheng mellom måleinstrumentene

Tidligere studier har undersøkt sammenheng mellom måleinstrumentene hos pasienter med PHF. Det er rapportert god til sterk sammenheng mellom de ulike måleinstrumentene i litteraturen (Beastall, Fielding, Christie & Johnstone, 2012; Maurer et al., 2021; van de Water et al., 2014; Ye et al., 2013) (Tabell 4). Det ble ikke funnet studier som har undersøkt sammenheng mellom CMS, DASH og OSS seks måneder postoperativt hos tre og fire-parts PHF operert med nagle eller plating.

**Tabell 4:** Sammenheng mellom måleinstrumentene i tidligere litteratur

Forfatter (år)	Studie-design	n	Alder Kvinner %	Brudd type	Behandling	Måletids punkt	Måleinstrument	Resultater
Maurer et al. (2021)	Tverrsnitt	76	x 50%	2, 3 og 4 parts	Plate og skuer	5 år	CMS og DASH CMS og OSS DASH og OSS	$r_s = -0,62^*\Delta$ $r_s = 0,64^*\Delta$ $r_s = 0,88^*\Delta$
van de Water et al. (2014)	Prospektiv kohort	20	68 80%	2 og 3 parts	Operasjon/ konservativ (80%)	6uker  12uker	CMS og DASH CMS og OSS DASH og OSS CMS og DASH CMS og OSS DASH og OSS	$r_p = -0,62^{*1}$ $r_p = 0,53^{*1}$ $r_p = 0,80^{*1}$ $r_s = -0,86^*\Delta$ $r_s = 0,79^*\Delta$ $r_s = 0,85^*\Delta$
Ye et al. (2013)	Retro- spektiv kohort	89	67 70%	3 og 4 parts	Plating	13-55 måneder	CMS og DASH	$r_s = -0,85^*\Delta$
Beastall et al. (2012)	Tverrsnitt	44	58 34%	x	Plating	1 år	OSS og CMS	$r_s = -0,85\Delta$

\*  $p < 0,05$  $r_s\Delta$  Spearman rho koeffisient $r_p^{*1}$  Pearsons koeffisient

Alder er oppgitt i gjennomsnitt

x, ikke oppgitt

### 2.8.2 Sammenheng mellom bakgrunnsvariabler og lav CMS skår

Tidligere studier har undersøkt sammenheng mellom bakgrunnsvariabler og lav CMS skår (reduisert skulderfunksjon) hos pasienter med PHF. En studie viste sammenheng mellom høy alder og kvinner og lav CMS skår (Südkamp, Audigé, Lambert, Hertel & Konrad, 2011), to studier viste sammenheng mellom høy alder og lav CMS skår (Clement, Duckworth, et al., 2014; Miquel, Elisa, Fernando, Alba & Torrens, 2021). Det var også diskrepans mellom resultatene når det gjelder sammenheng mellom kjønn og lav CMS skår (Clement, Duckworth, et al., 2014; Miquel et al., 2021; Südkamp et al., 2011) (Tabell 5).

**Tabell 5:** Resultater fra litteratursøk, sammenheng mellom bakgrunnsvariabler og lav CMS

Forfatter (år)	Studie- design	n	Alder Kvinner%	Brudd- type	Behandling	Måletids -punkt	Bakgrunns- variabler	p-verdi
Südkamp et al. (2011)	Retro- spektiv kohort	463	Menn 57 Kvinne 66 72%	3-4 parts	Plate og skuer	1 år	Alder>80 Kvinner	p<0,01 p<0,01
Clement, Duckworth, et al. (2014)	Retro- spektiv kohort	637		1-4 parts	Operasjon/ konservativ (90%)	1 år	Kvinner Alder	p=0,09 p<0,01
Miquel et al. (2021)	Tverrsnitt	126	72 år 79%	1-4 parts	Operasjon/ Konservativ (83%)	1 år	Alder Kvinner BMI	p<0,05 p=0,23 p=0,54

Alder er oppgitt i gjennomsnitt

BMI, body mass index

n, utvalgsstørrelsen

## 2.9 Teoretisk forankring

Masteroppgaven er forankret i et realistisk vitenskapsteoretisk perspektiv. I perspektivet menes det at virkeligheten eksisterer i en spesifikk og entydig form i verden uavhengig av vår erkjennelse av den (Justesen, 2012, s. 15). Objektiv kunnskap er sentralt i realismen, derav kvantitative metoder (Justesen, 2012, s. 29). Bruk av tall og statistikk er vurdert som nøkkelen for å oppnå relevant kunnskap innenfor realismen (Wifstad, 2018, s. 20). Datamaterialet som samles inn skal i stor grad kunne måles som størrelse, form eller antall (Wifstad, 2018, s. 60). Kvantitative undersøkelser gjøres på gruppenivå, og da kan kontekstuelle forhold som pasientens egen mening og erfaring utebli (Bondevik & Engebretsen, 2020, s. 63).

For å redusere skillet mellom kropp og sinn i det realistiske perspektivet, ble den biopsykososiale modellen utviklet (Wifstad, 2018, s. 76-78). Den biopsykososiale modellen bygger på samspillet mellom biologiske, psykologiske og sosiale faktorer for å forstå helse og sykdom (Wade & Halligan, 2017). Masteroppgaven har datamaterialet som omhandler skulderfunksjon. For å samle data om funksjon og funksjonshemning, eller ikke-dødelige helseutfall, utviklet Verdens Helseorganisasjon Internasjonal klassifisering av funksjon, funksjonshemning og helse (ICF) (Ustün, Chatterji, Bickenbach, Kostanjsek & Schneider, 2003). ICF bygger på den biopsykososiale modellen og er et verktøy som kan bidra til å forstå menneskelig funksjon og funksjonshemning. ICF modellen beskriver samspillet mellom helsetilstand (sykdom eller lidelse), kroppsstrukturer og funksjoner, aktiviteter, deltakelse, miljøfaktorer og personlige faktorer (Ustün et al., 2003).

## **3.0 Metode**

I masteroppgaven er det benyttet kvantitative analysemetoder på et tallmateriale for å besvare problemstillingene.

### **3.1 Datamaterialet**

I masteroppgaven er det benyttet datamaterialet fra en randomisert kontrollert studie (RCT) som omhandler PHF. RCTens formål var å sammenligne to ulike operasjonsmetoder hos tre og fire-parts PHF, og hovedutfallsmålet var skulderfunksjon målt med DASH.

Datainnsamlingen ble fullført juni 2023 (n=79). Masterstudenten samlet inn data seks måneder og to år postoperativt ved bruk av CMS, DASH og OSS.

### **3.2 Studiedesign for masteroppgaven**

Masteroppgavens design er en observasjonell tverrsnittstudie, med formål om å undersøke funksjonsnivå målt med ulike måleinstrumenter, undersøke måleinstrumentenes eventuelle tak- eller gulveffekter, undersøke sammenheng mellom måleinstrumenter og undersøke sammenheng mellom bakgrunnsvariabler og lav CMS skår blant pasienter med tre og fire-parts PHF seks måneder postoperativt.

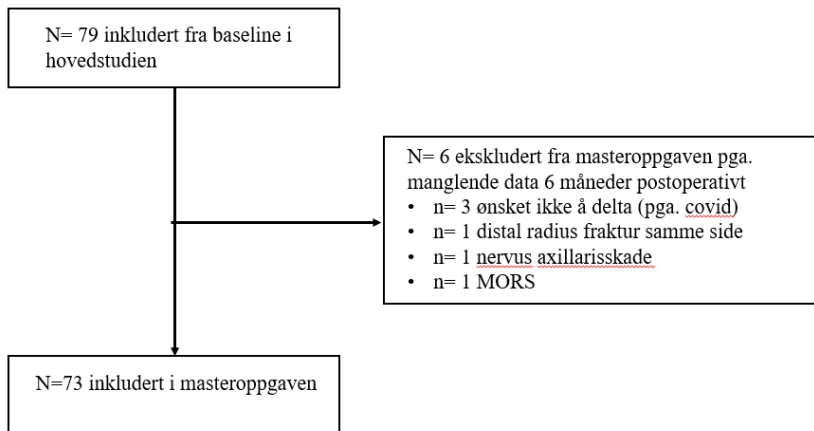
### **3.3 Utvalget i masteroppgaven**

Alle tre og fire-parts PHF i perioden oktober 2016 til august 2021 som oppfylte inklusjonskriteriene (Tabell 6) ble forespeilet å delta i studien. Utvalget bestod av 79 pasienter med operasjonstrengende tre eller fire-parts PHF ved baseline. Pasientene ble randomisert til to ulike intervensjonsgrupper, operasjon med innvendig nagle eller plating. Masterkandidaten og pasientene var blindet for operasjonsmetode, og i masteroppgaven er deltakere i begge gruppene inkludert i analysene. På grunn av manglende data for seks personer ved seks måneder postoperativ oppfølging består utvalget i masteroppgaven av 73 pasienter (Figur 8).

**Tabell 6:** Inklusjon og eksklusjonskriterier i hovedstudien

Inklusjonskriterier	Eksklusjonskriterier
<ul style="list-style-type: none"><li>- Pasienter &gt; 18 år</li><li>- Alvorlig forskyvning, definert som feilstilling på minst 45 grader vinkelavvik i valgus eller 30 grader i varus i ekte frontal projeksjon, avhengig av om bruddet er påvirket eller ikke</li><li>- Brudd med mer enn 50 % forskyvning av hodet mot kirurgisk nakke</li><li>- Tuberculum minor eller major må fraktureres i et 3 eller 4 parts brudd</li></ul>	<ul style="list-style-type: none"><li>- Frakturer eldre enn 3 uker</li><li>- Ipsilateral skade som kan påvirke gjenopprettings- og skåringssystemene</li><li>- Manglende evne til å beskytte osteosyntesen, for eksempel ved bruk av krykker</li><li>- Systemisk sykdom som kan påvirke tilhelingsprosessen eller skåringssystemer</li><li>- Pasienter med for liten humerusdiameter til å bruke skruer</li><li>- Pågående infeksjonsprosess rundt snittstedet for osteosyntesen</li><li>- Caput humeri bare et tynt skall eller splitt</li><li>- Manglende evne til å lese og forstå norsk</li><li>- Patologisk fraktur</li><li>- Nevrovaskulær skade</li><li>- Åpen fraktur</li><li>- Ikke- etterlevelse</li><li>- Medfødt anomali</li><li>- Bruddluksasjon</li><li>- Rusmisbruk</li><li>- Pasienter som ikke er bosatt i nedslagsfeltet til Ahus</li><li>- Pasienter som ikke er i stand til å forplikte seg og- eller forstå skriftlig samtykke</li></ul>

**Figur 8:** Flytskjema viser frafall etter inklusjon i hovedstudien, og endelig utvalg i masteroppgaven



### 3.4 Datainnsamling

PROMs og CROMs data om skulderfunksjon er samlet inn av masterkandidaten seks måneder postoperativt med måleinstrumentene CMS, DASH og OSS i perioden april 2017 til desember 2021. All datainnsamling ble gjort på papir, CMS ble i tillegg registrert elektronisk i Research Electronic Data Capture (REDCap).

#### 3.4.1 Måleinstrumenter

Siden tallmaterialet i denne i masteroppgaven er sekundærdata fra en RCT, var måleinstrumentene forhåndsbestemt. I masteroppgaven benyttes det data fra måleinstrumentene CMS, DASH og OSS. De er blant de mest brukte måleinstrumentene på skulderfunksjon etter PHF, men samtidig etterspørres det mer kunnskap om hvor godt de egner seg for å måle skulderfunksjon på denne gruppen (Richard et al., 2020). Målemetode for smerte og skulderfunksjon var to PROMs, DASH og OSS, og ett kombinert PROMs og CROMs, CMS. Se punkt 2.7.1.1-2.7.1.5 i teorikapittelet for nærmere beskrivelse av måleinstrumentene.

##### 3.4.1.1 Constant Murley skår (CMS)

På grunn av manglende oversettelse av CMS ble spørsmål i PROMs delen, smerte og ADL- domenet, oversatt av masterkandidaten muntlig for pasientene. I CROMs delen beholdt deltakerne t-skjorte på for å ivareta blinding. CROMs delen ble utført i tråd med Constant et al. (2008) retningslinjer. CMS ble utført på et behandlingsrom uten andre enn pasient og

masterkandidaten til stede. Ved styrketesten var fjærvekten festet til en behandlingsbenk, og pasient ble instruert i hvordan styrketesten skulle utføres både muntlig og praktisk. Styrketesten ble ikke utført hos de som ikke oppfylte kriteriene. Det tok ca. 15-20 minutter å gjennomføre hele måleinstrumentet. Se vedlegg 1 for CMS spørreskjema hvor også utregning av resultatene fra styrketesten av spesifisert.

#### **3.4.1.2 Disabilities of Arm, Shoulder and hand Questionnaire (DASH) og Oxford Shoulder Score (OSS)**

Deltakerne fikk muntlig og skriftlig informasjon på hvordan DASH og OSS skulle besvares, og spørreskjemaet ble besvart av deltakerne selvstendig eller med hjelp fra pårørende på venterommet. Masterkandidaten var tilgjengelig for spørsmål angående spørreskjemaet. Spørreskjemaene tar 3-5 minutter å besvare. Se spørreskjemaene i vedlegg 2 og 3.

#### **3.4.2 Demografiske variabler**

Tilgjengelige demografiske variabler i denne masteroppgaven var alder, kjønn, BMI, røykestatus, arbeidsstatus, ASA skår, bruddside og dominant arm. ASA skår ble satt av anestesipersonell preoperativt, resterende bakgrunnsvariabler ble samlet inn av prosjektleder ved baseline.

### **3.5 Statistisk analyse**

Statistiske analyser er utført i programvaren IBM Statistical Package for the Social Sciences Version 29.0.1.0 (SPSS). I oppgaven er det utført univariate, bivariate og multivariate analyser for å besvare problemstillingene.

#### **3.5.1 Deskriptive analyser**

##### **3.5.1.1 Deskriptiv analyse av demografisk data**

Deskriptive analyser ble brukt for å beskrive og syntetisere data i masteroppgaven (Polit & Beck, 2017, s. 356). Sentraltendens, spredning og normalfordeling ved kontinuerlige variabler ble testet ved gjennomsnitt og median, standardavvik og maks/minimum skår, histogram, normalfordelingsplott og boksplokk. Histogram og boksplokk er presentert i oppgaven.



Kategoriske variabler er testet i krysstabeller og kji-kvadrat, og presentert i stolpediagram. Sentraltendens for normalfordelte data ble beskrevet med gjennomsnitt og standardavvik ved normalfordelte kontinuerlige variabler og som frekvens og prosent ved kategoriske variabler. Resultatene presenteres i tabell. På bakgrunn av at det var langt færre menn enn kvinner i utvalget, er deskriptive analyser også utført gruppert på kjønn. Forskjell mellom grupper ble testet med kji-kvadrat test og ficher exact test ved kategoriske variabler og ble testet med to-utvalgs t-test for kontinuerlige variabler.

### **3.5.1.2 Deskriptiv analyse av data fra måleinstrumentene**

Tilsvarende analyser og beskrivelser som nevnt i punkt 3.5.1.1 ble gjort for måleinstrumentene CMS, DASH og OSS og resultatene presenteres i histogrammer og tabell. For ytterligere deskriptive analyser ble kontinuerlige data fra CMS og DASH kategorisert, da det finnes etablerte kategoriseringer for dårlig, moderat, god eller svært god funksjon. Grafisk tolkning av histogram ble benyttet for å undersøke andel med høyest eller lavest mulig skår (eller manglende skårer) for de tre måleinstrumentene. Eventuelle tak- eller gulveffekter (eller manglende data) for de tre måleinstrumentene ble presentert i stolpediagram og tabell.

### **3.5.2 Sammenheng mellom data fra måleinstrumentene**

Ved undersøkelse av korrelasjon mellom kontinuerlige variabler CMS, DASH og OSS er det gjort parametriske analyser med Pearsons korrelasjon ( $r_p$ ) (Pallant, 2020, s. 128).

Korrelasjonskoeffisientenes styrke ble tolket ut i fra Hinkle, Jurs og Wiersma (2003) retningslinjer,  $r_p = 0,00-0,30$  viser liten sammenheng,  $r_p = 0,30-0,50$  viser lav sammenheng og  $r_p = 0,50-0,70$  viser moderat sammenheng,  $r_p = 0,70-0,90$  viser høy sammenheng og  $r_p = 0,90-1,00$  viser veldig høy sammenheng. Forutsetninger for Pearsons korrelasjon er testet ved linearitet i spredningsdiagram ved grafisk tolkning, og normalfordeling ved gjennomsnitt, standardavvik, median, skjevhet, histogram, pp-plot og boks-plott (Vedlegg 4, 5 og 6).

Forutsetningene for Pearsons korrelasjon var ikke oppfylt på grunn av fire utliggere i variabelen DASH og en utligger i variabelen CMS. Analysene ble derfor utført med både Spearmans rho korrelasjon og Pearsons korrelasjon med og uten utliggere. Siden disse analysene ga sammenlignbare resultater, ble det valgt å presentere resultatene fra Pearsons korrelasjonsanalyse. Utliggerne påvirket ikke resultatet av betydning, og analysene er derfor

gjort på hele utvalget (inkludert uteliggere). Resultatene fra Pearsons korrelasjon blir presentert med korrelasjonskoeffisienter ( $r_p$ ) og p-verdi i tabell.

For å undersøke sammenheng mellom de kontinuerlige variablene CMS, DASH og OSS ble lineær regresjon benyttet. Forutsetninger for enkel lineær regresjon ble testet med grad av linearitet i spredningsdiagram ved grafisk tolkning, uavhengige målinger, (uavhengige residualer), homoskedastisitet i spredningsdiagram ved grafisk tolkning, normalfordeling av residualer i histogram, spredningsdiagram og normalfordelingsplott, og residualer mellom 3 og -3 i spredningsdiagram. Forutsetningene var ikke oppfylt for CMS og DASH på grunn av residualer med verdi over 3 i spredningsdiagrammet og usikker homoskedastisitet (Vedlegg 7). Analysene ble utført med og uten utliggere, som viste at uteliggere ikke påvirket resultatene av betydning. Resultatene fra lineær regresjon blir presentert med regresjonskoeffisienter (B), 95 % konfidensintervall og p-verdi i tabell. Signifikansnivået er  $p < 0,05$ , og 95% KI er vurdert i forhold til variabelens måleskala.

Fordi det var en veldig liten andel menn i utvalget (18%), ble korrelasjons- og lineære regresjonsanalyser også utført kun for kvinner, da normative gjennomsnittsverdier for de ulike måleinstrumentene kan være påvirket av kjønn (Clement & Court-Brown, 2014; Constant et al., 2008; Aasheim & Finsen, 2014).

### **3.5.3 Sammenheng mellom lav CMS skår og uavhengige bakgrunnsvariabler**

For å estimere sannsynligheten for lav CMS skår (avhengig dikotom kategorisk variabel) basert på verdien til en uavhengig variabel ble logistisk regresjon benyttet (justert). Analysene ble først utført med de uavhengige variablene separat mot den avhengige variabelen (ujustert). De uavhengige variablene kjønn, alder, røykestatus, arbeidsstatus, ASA skår og BMI ble inkludert i regresjonsmodellen. Analysene ble utført med og uten variabelen røyking på grunn av lavt antall som røyker. Forutsetninger for logistisk regresjon ble testet med grad av multikollinearitet (variansinflasjonsfaktor  $< 10$  og tolleranseverdi  $> 0,1$ ), Pearsons korrelasjon, normalfordeling ved gjennomsnitt, standardavvik, median, skjevhet, histogram, pp-plot og boks-plott mellom kontinuerlige uavhengige variabler. For kategoriske uavhengige variabler ble Spearman's rho benyttet.

Det er anbefalt minst 10 hendelser per uavhengig variabel (Vittinghoff & McCulloch, 2007), så det ble valgt seks uavhengige variabler med kjent sammenheng fra tidligere forskning. Regresjonsmodellen ble testet med Homes and Lemeshow Test ( $> 0,05$ ) og Goodness of fit

test. Ytterligere forutsetninger ved analysen ble testet med Cook's distance (Cook's D) for å måle hvor mye estimatene i modellen endrer seg hvis observasjoner med høy Cook's D verdi fjernes fra datasettet. Leverage mål ble også benyttet for å indentifisere høye Leverage verdier som også kan påvirke estimatene i modellen. Ved visuell tolkning av spredningsdiagram for Leverage value ble det indentifisert fem høye leverage verdier, men dette ble vurdert som akseptabelt (Figur 24). I spredningsdiagrammet for Cook's D ble det indentifisert to høye verdier (Figur 25) og analysene ble utført uten disse uteliggerne. Det ble undersøkt om dataene for de aktuelle observasjonene var plottet riktig, noe det viste seg å være. Det kunne vært aktuelt å gå videre med å prøve transformasjon eller en robust modell, men det ble ansett å være utenfor rammene for en masteroppgave. Resultatene fra regresjonsmodellen presenteres derfor med justerte estimer både med og uten uteliggere (n=73 vs. n=71). Resultatene blir presentert med justert og ujustert odds ratio (OR), p-verdi og 95 % konfidensintervall i tabell. Signifikansnivået er  $p < 0,05$ , og 95% KI er vurdert i forhold til variabelens måleskala.

### **3.6 Forskningsetikk**

Masteroppgaven benyttet sekundærdata fra en RCT som er godkjent fra Regionale komiteer for medisinsk og helsefaglig forskningsetikk (REK 2016/626) og databeskyttelsesansvarlig på Ahus. RCTen blir gjennomført i samsvar med etiske prinsipper fra Helsinkideklarasjonen og ICH/Good Clinical Practice. Datalagring blir håndtert i tråd med nasjonale lover om personvern. Deltakerne er nummerert og aidentifisert. Deltakerne har gitt muntlig og skriftlig frivillig informert samtykke ved inklusjon i RCTen. Hensikt med masteroppgaven er meldt inn og godkjent hos REK med endringsmelding 14.12.22 (Vedlegg 14).

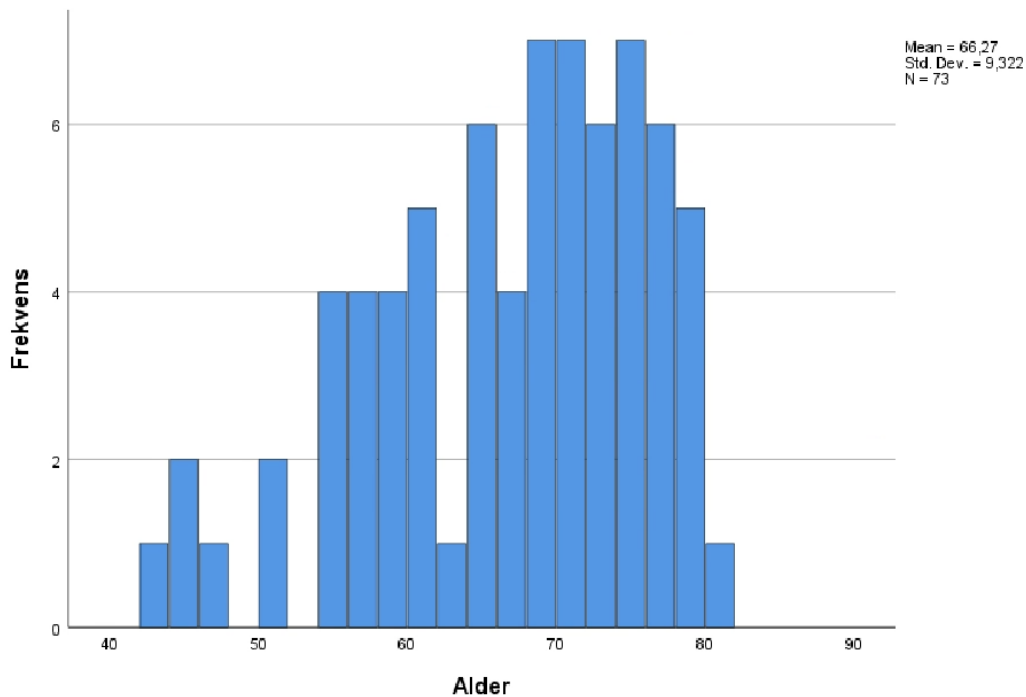
## 4.0 Resultater

I dette kapitlet presenteres først resultater fra deskriptiv analyse av demografisk data, deskriptiv analyse av funksjonsdata, deretter korrelasjons og regresjonsanalyser og til slutt beskrives resultatene fra logisk regresjonsanalyse.

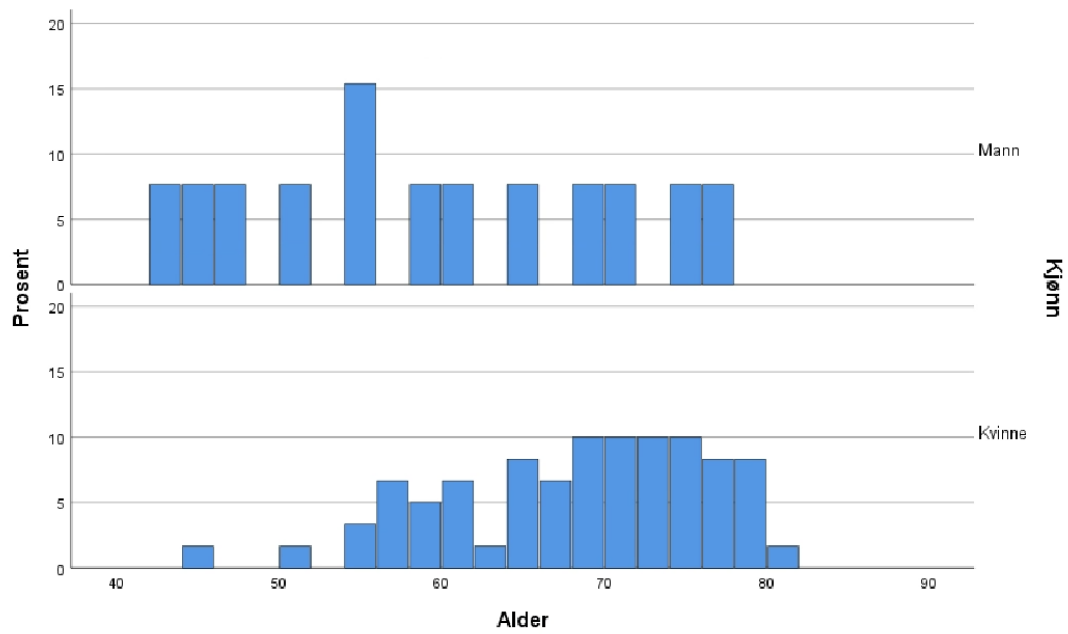
### 4.1 Deskriptiv statistikk av demografiske variabler

Masteroppgaven inkluderte 73 deltakere (Tabell 7). Gjennomsnittsalderen for utvalget var 66 år (SD 9,3) (Figur 9) og var signifikant høyere for kvinner sammenlignet med menn (67,9 år vs. 58,9 år,  $p < 0,01$ ) (Figur 10 og 11). Utvalget hadde en skjev kjønnsfordeling, henholdsvis 82% kvinner og 18% menn. Gjennomsnittlig BMI var 28. Andelen pensjonister/uføre var 67% i hele utvalget, blant kvinnene var andelen 75% og blant menn var andelen 31%. Blant hele utvalget var andelen røykere 8%. Det var signifikant færre kvinner enn menn som var i arbeid, men flere som var pensjonister/uføre ( $p = 0,04$ ). Preoperativt ble ASA skår 1 rapportert hos 37%, ASA skår 2 ble rapportert hos 49% og ASA skår 3 ble rapportert hos 14%. Andelen med fraktur i dominant arm var 47%.

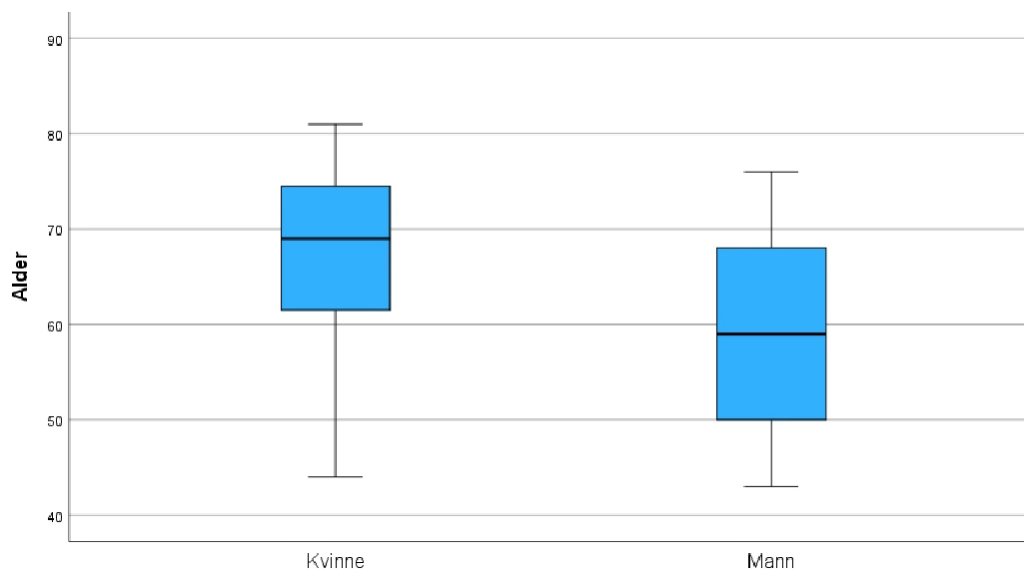
**Figur 9:** Histogram med aldersfordelingen i utvalget (n=73)



**Figur 10:** Histogram med aldersfordeling gruppert på kjønn



**Figur 11:** Boksplott med aldersfordeling gruppert på kjønn



**Tabell 7:** Demografiske og bruddrelaterte karakteristika i hele utvalget og gruppert på kjønn

Demografiske variabler	Hele utvalget (n=73)	Kvinner (n= 60)	Menn (n=13)
Alder ved operasjon (år), gj.snitt ( $\pm$ SD)	66,3 (9,3)	67,9 (8,1)	58,9(11,4)**
Kjønn, n (%)		60 (82,2)	13 (17,8)
BMI gjennomsnitt ( $\pm$ SD)	28,3 (4,8)	27,8 (4,9)	30,3 (3,8)
Røyker, n (%)	6 (8,2)	5 (8,3)	1 (7,7)
Arbeid			
Pensjonist/ufør n (%)	49 (67,1)	45 (75,0)	4 (30,8)*
I arbeid n (%)	24 (32,9)	15 (25)	9 (69,2)*
ASA skår ved operasjon			
ASA 1, n (%)	27 (37,0)	19 (31,7)	8 (61,5)
ASA 2, n (%)	36 (49,3)	32 (53,3)	4 (30,8)
ASA 3, n (%)	10 (13,7)	9 (14,5)	1 (7,7)
Brudd i høyre arm, n (%)	34 (46,6)	26 (76,5)	8 (23,5)
Brudd i dominant arm, n (%)	34 (46,6)		

Body Mass Index (BMI) måler forholdet mellom høyde og vekt for å estimere andel kroppsfett.  
The American Society of Anesthesiologist (ASA) er et system for gradering av risiko i forbindelse med operasjon, basert på faktorer som blant annet komorbiditet og allmenntilstand. Skårer fra 1-5, hvor 1 indikerer lav risiko og 5 indikerer betydelig risiko  
Standardavvik (SD)

\* Signifikant forskjell mellom kvinner og menn  $p < 0,05$

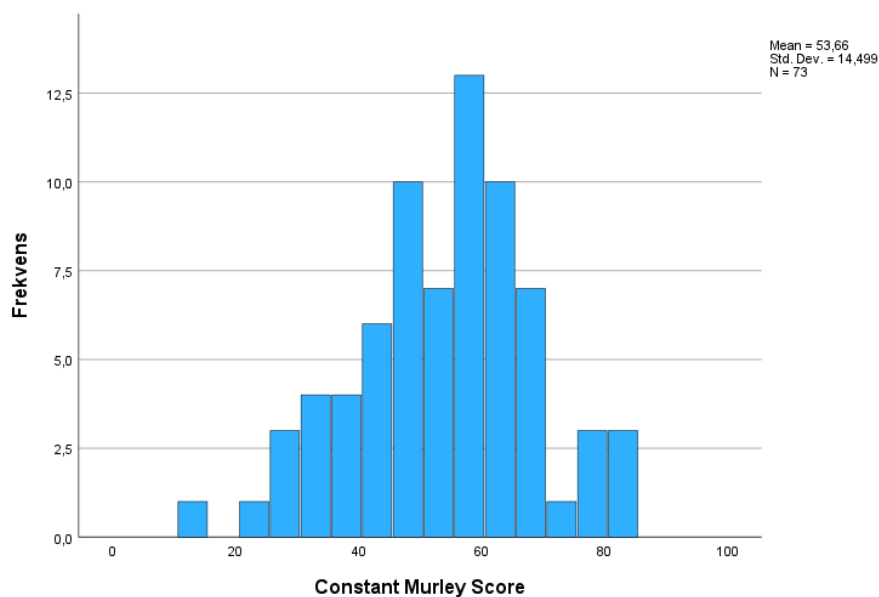
\*\* Signifikant forskjell mellom kvinner og menn  $p < 0,01$

## 4.2 Deskriptiv statistikk av data fra måleinstrumentene

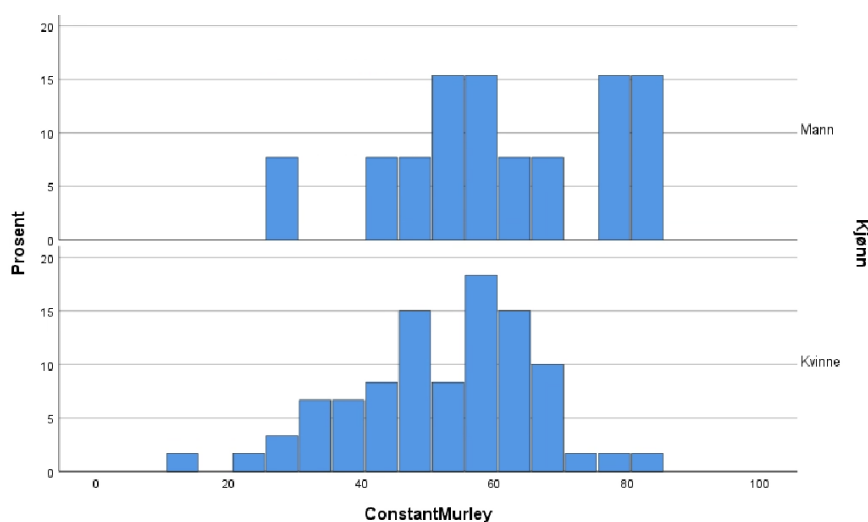
### 4.2.1 Postoperativ funksjon målt med CMS

Gjennomsnittlig CMS skår var 53,7 (Figur 12). Kvinner hadde signifikant lavere skår sammenlignet med menn (52,0 vs. 61,2,  $p=0,04$ ) (Figur 13).

**Figur 12:** Histogram av skåringsfordeling for CMS (n=73)



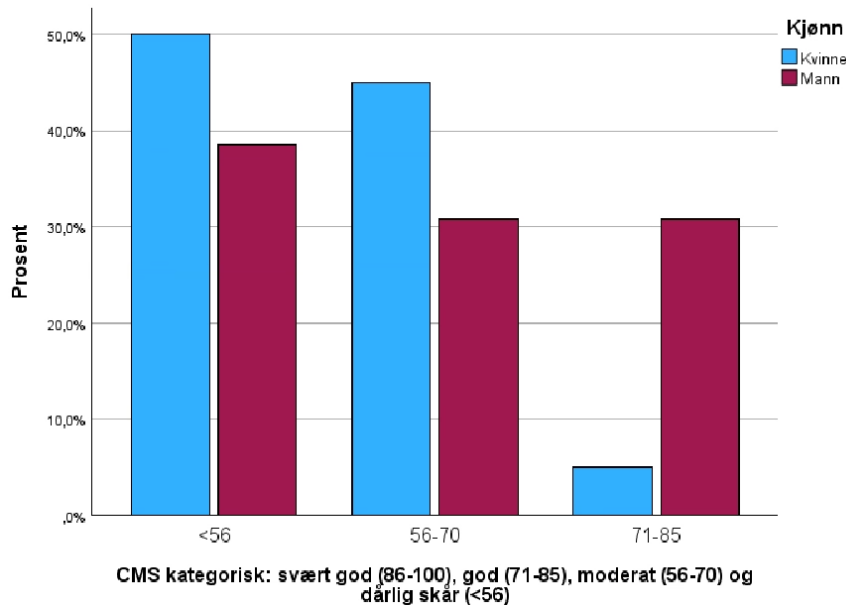
**Figur 13:** Histogram gruppert på kjønn med skåringsfordeling CMS (n=73)



Blant hele utvalget hadde 48% det som ansees som dårlig funksjon målt med CMS (<56 poeng) (Figur 14), og det var større andel kvinner enn menn med dårlig funksjon (50% vs. 39%,  $p<0,05$ ). Det var 43% som hadde moderat funksjon (CMS skår 56-70), og det var signifikant forskjell mellom kvinner og menn (45% vs. 31%,  $p<0,05$ ). Kun 10% hadde god

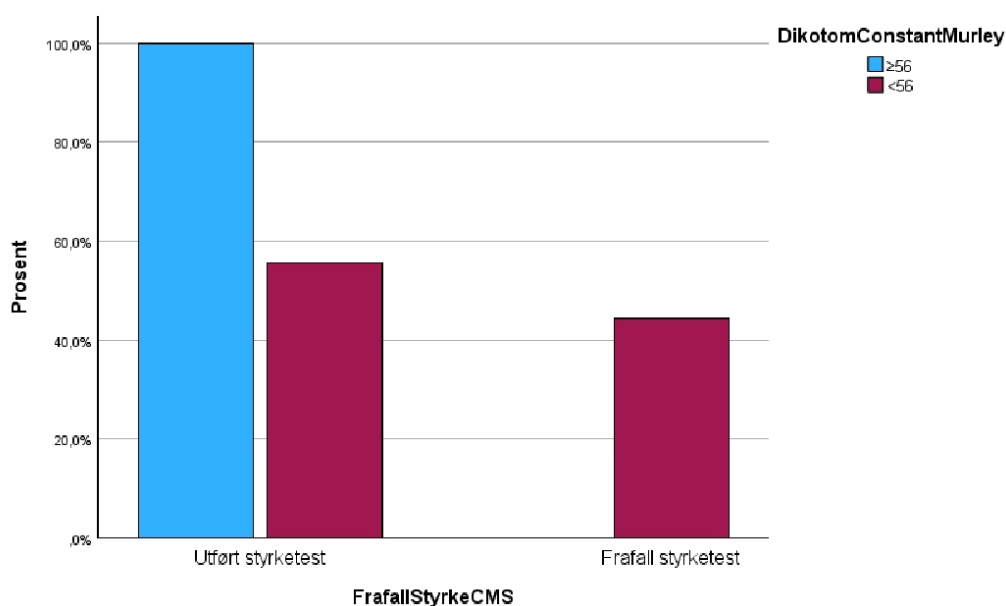
funksjon (CMS skår 71-85), og det var signifikant forskjell mellom kvinner og menn (5% vs. 31%,  $p < 0,05$ ). Det var ingen som hadde svært god funksjon (CMS skår 86-100).

**Figur 14:** Stolpediagram for kategorisert CMS gruppert på kjønn



Det var bare 56% blant de med CMS skår <56 som oppfylte kriteriene for styrketesten, mens 100% av de med CMS  $\geq 56$  oppfylte kriteriene (Figur 15). Det var signifikant forskjell mellom gruppene med CMS skår  $\geq 56$  vs. <56 når det gjaldt å oppfylle kriteriet for styrketesten (44% vs. 0%,  $p < 0,01$ ).

**Figur 15:** Stolpediagram for dikotomisert utført vs. ikke utført styrketest, gruppert på CMS skår  $\geq 56$  og <56

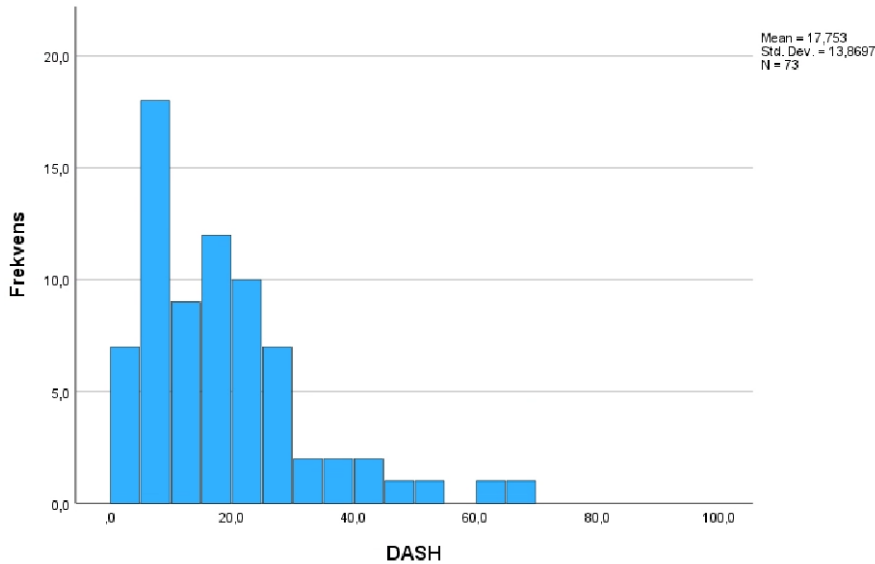




## 4.2.2 Postoperativ funksjon målt med DASH

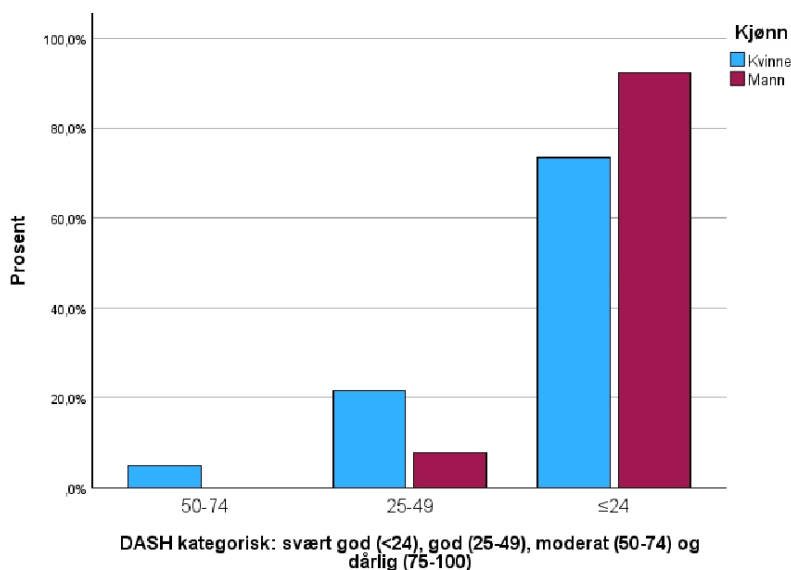
Gjennomsnittlig DASH skår var 17,8. Det var forskjell, men ikke signifikant, mellom kvinner og menn (19,2 vs. 10,9,  $p=0,05$ ) (Figur 16).

**Figur 16:** Histogram av skåringsfordeling for DASH (n=73)



Kategorisert DASH skår, viste at 77% hadde svært god funksjon (skår <24), 19% hadde god funksjon (skår 25-49), og 4% hadde moderat funksjon (skår 50-74) (Tabell 9). Fordelt på kjønn hadde 73% kvinnene og 92 % av mennene svært god funksjon, 22% kvinner og 8 % menn hadde god funksjon og 5% kvinner hadde moderat funksjon (Figur 17). Det var ikke signifikant forskjell mellom kjønn og kategorisert DASH skår.

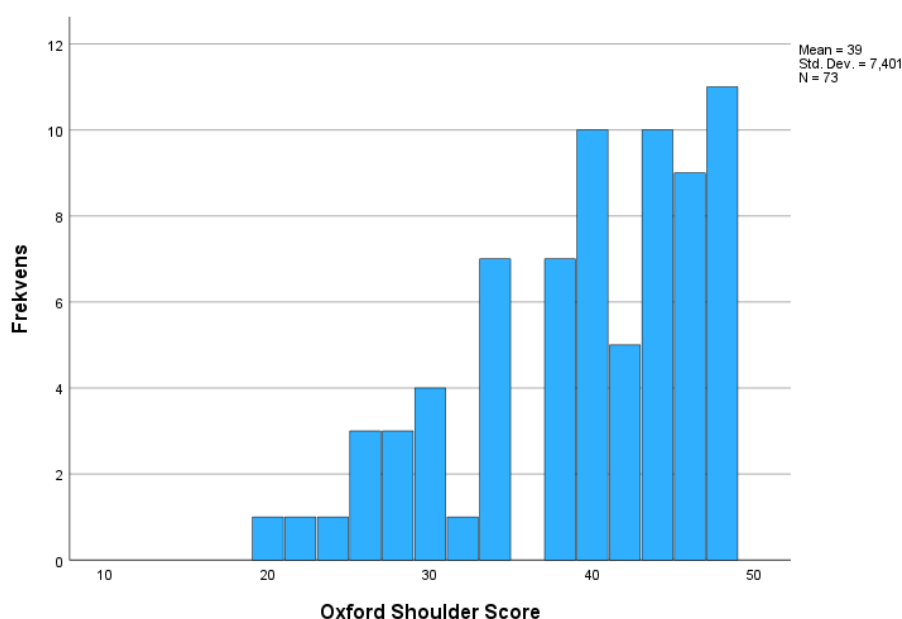
**Figur 17:** Stolpediagram for kategorisert DASH skår fordelt på kjønn



### 4.2.3 Postoperativ funksjon målt med OSS

OSS skår var gjennomsnittlig 39 for hele utvalget. Kvinner hadde en lavere skår enn menn, men forskjellen var ikke signifikant (38,2 vs. 42,6,  $p=0,05$ ) (Figur 18). Det er ingen kjent kategorisering av OSS skår.

**Figur 18:** Histogram av skåringsfordeling for OSS (n=73)



Tabell 8 presenterer deskriptiv analyse av data fra måleinstrumenter for hele utvalget og fordelt mellom kjønn. Tabell 9 viser kategorisering av kontinuerlige data fra måleinstrumentene, og resultater fordelt mellom kjønn.

**Tabell 8:** Deskriptiv statistikk av kontinuerlig funksjonsdata

Måleinstrument	Hele utvalget (n=73)	Kvinner (n= 60)	Menn (n=13)
CMS skår, gjennomsnitt ( $\pm$ SD)	53,7 (14,5)	52 (13,7)	61,2 (16,3)*
DASH skår, gjennomsnitt ( $\pm$ SD)	17,8 (13,9)	19,2 (14,5)	10,9 (8,6)
OSS skår, gjennomsnitt ( $\pm$ SD)	39,0 (7,4)	38,2 (7,4)	42,6 (6,5)

Constant Murley skår (CMS) måler skulderfunksjon, 2 er dårligst skår og 100 er beste skår  
The disabilities of arm, shoulder and hand (DASH) måler generell armfunksjon, 100 er dårligst skår og 0 er beste skår

Oxford shoulder skår (OSS) måler postoperativ funksjon i skulderen, og skårer fra 0-48, hvor 48 er best mulig skår

Standardavvik (SD)

\* Signifikant forskjell mellom kvinner og menn  $p<0,05$

**Tabell 9:** Deskriptiv statistikk av kategorisk funksjonsdata

Måleinstrument	Hele utvalget (n=73)	Kvinner (n=60)	Menn (n=13)
Kategorisert DASH, n (%)			
Svært god	56 (76,7)	44 (73,3)	12 (92,3)
God	14 (19,2)	13 (21,7)	1 (7,7)
Moderat	3 (4,1)	3 (5,0)	0
Dårlig	0	0	0
Kategorisert CMS, n (%)			
Svært god	0	0	0
God	7 (9,6)	3 (5,0)	4 (30,8)*
Moderat	31 (42,5)	27 (45,0)	4 (30,8)*
Dårlig	35 (47,9)	30 (50,0)	5 (38,5)*
CMS skår <56, n (%)	36 (49,3)	31 (51,7)	5 (38,5)
Maksimal OSS skår (48), n (%)	8 (11,0)	5 (8,3)	3 (23,1)

DASH skår: svært god (<24), god (25-49), moderat (50-74) og dårlig (75-100)

CMS skår: svært god (86-100), god (71-85), moderat (56-70) og dårlig (<56)

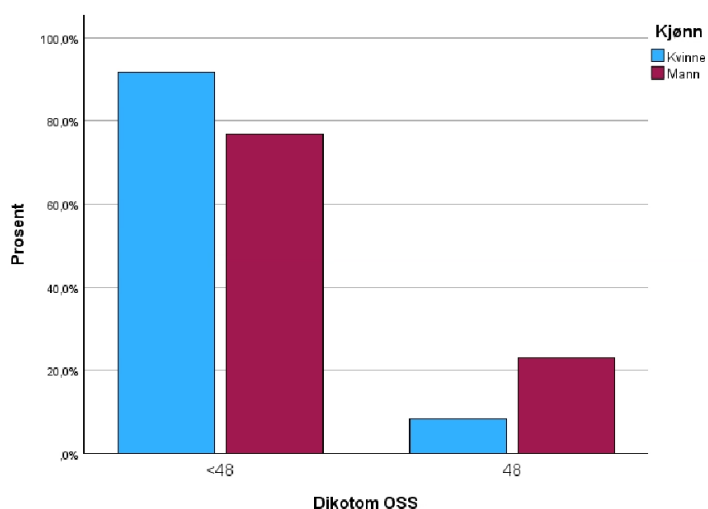
OSS, 48 er best mulig oppnådde skår

\* Signifikant forskjell mellom gruppene  $p < 0,05$

### 4.3 Tak- eller gulveffekter (eller manglende skårer) for de tre måleinstrumentene

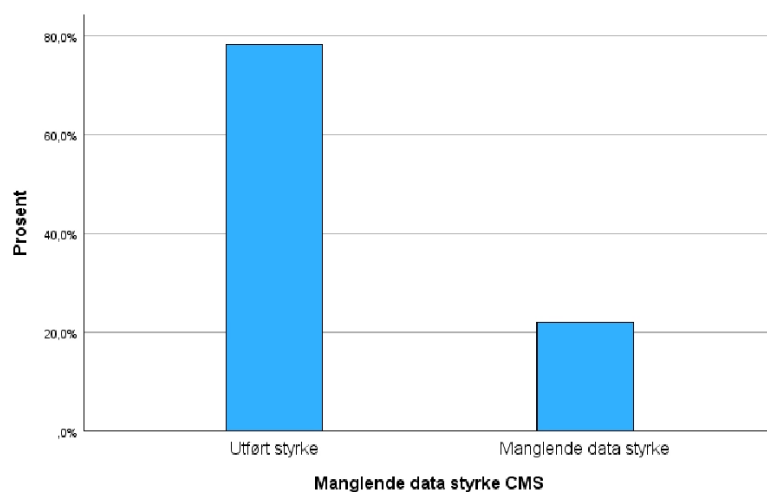
Ved deskriptive analyser ble det ikke identifisert tak- eller gulveffekt ved CMS (Figur 12) eller DASH (Figur 16). Blant hele utvalget skåret 11% best mulig OSS skår (skår 48) (Tabell 10). Det var ikke signifikant forskjell mellom kvinner og menn (8% vs. 23%,  $p=0,15$ ) (Figur 19). Resultatene kan indikere takeffekt ved instrumentet for menn.

**Figur 19:** Stolpediagram for dikotomisert OSS skår (<48 vs. 48) gruppert på kjønn. 48 er beste oppnådde skår



Ytterligere deskriptive analyser viste at 22% ikke oppfylte kriteriene til å utføre styrketesten i CMS (Figur 20), og resultatene viste en svært liten forskjell mellom kvinner og menn (22% vs. 23%,  $p=0,91$ ) (Tabell 10). Resultatene kan indikere gulveffekt for styrkedomenet i CMS.

**Figur 20:** Manglende data ved styrkedomenet av CMS



**Tabell 10:** Tak- eller gulveffekter (eller manglende skårer)

Måleinstrumenter	Hele utvalget (n=73)	Kvinner (n=60)	Menn (n=13)
Maksimal OSS skår (48), n (%)	8 (11,0)	5 (8,3)	3 (23,1)
Manglende oppfyllelse av kriteriene for CMS styrketest, n (%)	16 (22)	13 (22)	3 (23)

CMS, Constant Murley score  
OSS, Oxford shoulder score

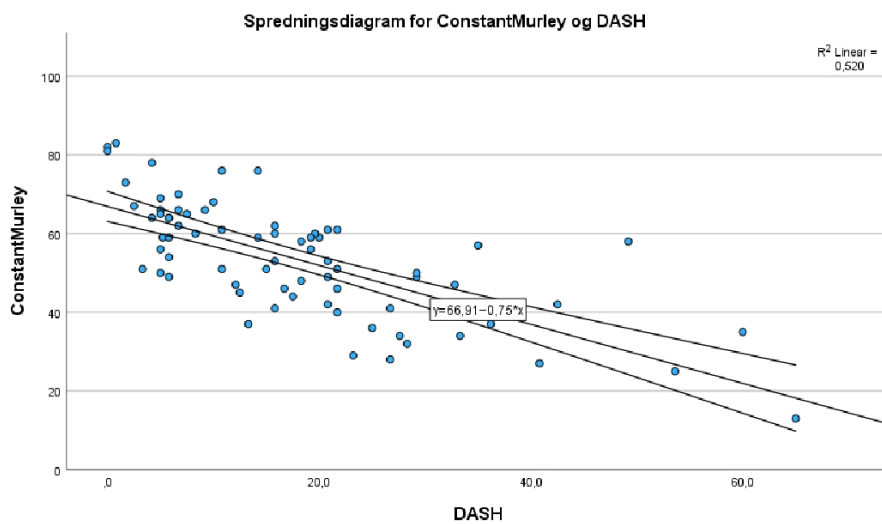
#### 4.4 Sammenheng mellom data fra måleinstrumentene

Figur 21, 22 og 23 viser spredningsdiagram mellom to og to av måleinstrumentene.

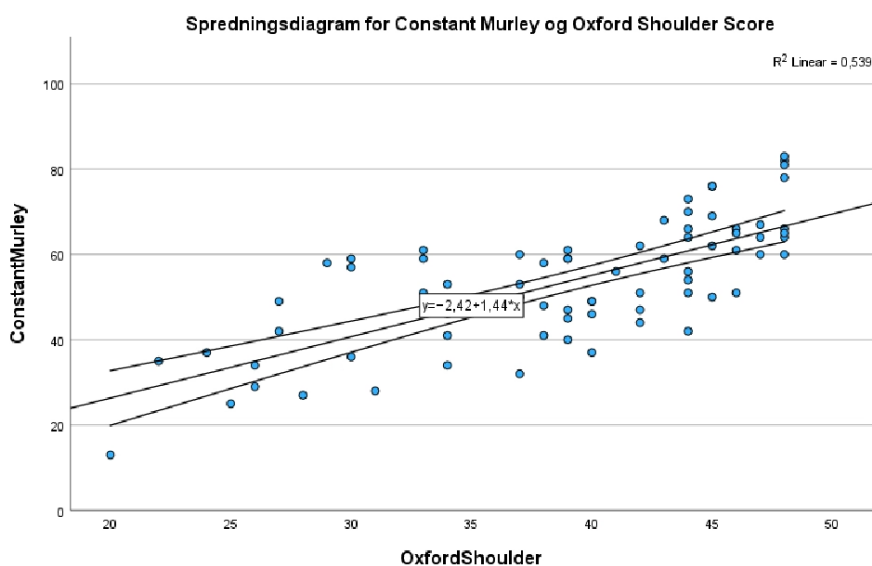
Korrelasjonsanalyser viste høy korrelasjon mellom måleinstrumentene (Tabell 11).

Resultatene viser en sterk negativ korrelasjon mellom DASH og CMS ( $r_p = -0,72$ ; 95% KI -0,82 til -0,59) og mellom DASH og OSS ( $r_p = -0,84$ ; 95% KI -0,89 til -0,75). Det er en sterk positiv korrelasjon mellom CMS og OSS ( $r_p = 0,73$ ; 95% KI 0,61-0,83). Sterk og signifikant ( $p < 0,01$ ) korrelasjon mellom alle måleinstrumentene og 95% KI viser relativt liten usikkerhet i resultatene.

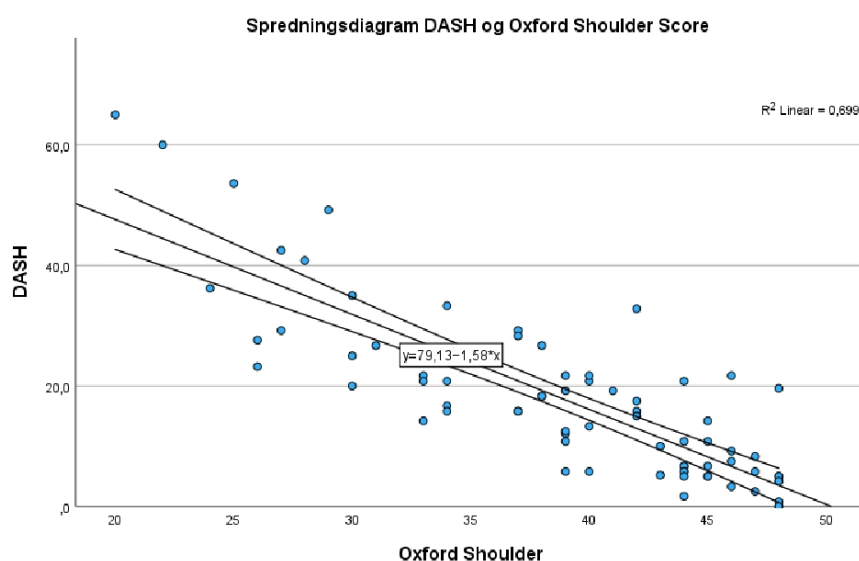
**Figur 21:** Spredningsdiagram mellom data fra måleinstrumentene CMS og DASH (n=73)



**Figur 22:** Spredningsdiagram mellom data fra måleinstrumentene CMS og OSS (n=73)



**Figur 23:** Spredningsdiagram mellom data fra måleinstrumentene DASH og OSS (n=73)



**Tabell 11:** Korrelasjonstabell mellom data fra måleinstrumentene ved bruk av parametriske Pearsons korrelasjon (n=73)

	CMS	DASH	OSS
CMS	1		
DASH	-0,72**	1	
OSS	0,73**	-0,84**	1

Pearsons korrelasjonskoeffisient ( $r_p$ ) <0,30 vurderes som veldig svak, 0,30-0,50 vurderes som svak og 0,50-0,70 vurderes som moderat, 0,70-0,90 vurderes som høy og 0,90-1,00 vurderes som veldig høy. DASH, disabilities of arm, shoulder and hand; OSS, Oxford Shoulder Score; CMS, Constant Murley Score.

\*\* Signifikant  $p < 0,01$

Resultatene fra lineær regresjon viste at hver gang CMS øker med 1 enhet (ett poeng), synker DASH med 0,8 enheter (poeng) ( $B_1 = -0,75$ ; 95% KI: -0,92 til -0,58) (Tabell 12). Når CMS øker med 1 enhet (ett poeng), øker OSS med 1,44 enheter (poeng) ( $B = 1,44$ ; 95% KI: 1,12 til 1,75). Hvis DASH synker med 1 enhet (ett poeng), øker OSS med 1,6 enheter (poeng) ( $B = -1,6$ ; 95% KI: -1,81 til -1,32). Det er liten usikkerhet i estimatene på grunn av relativt smale konfidensintervall.

**Tabell 12:** Lineær regresjon mellom data fra måleinstrumentene CMS, DASH og OSS (n=73)

Uavhengig variabel	Avhengig variabel	B	P	95% KI
DASH	CMS	-0,75	<0,01	-0,92, -0,58
OSS	CMS	1,44	<0,01	1,12, 1,75
OSS	DASH	-1,56	<0,01	-1,81, -1,32

DASH, disabilities of arm, shoulder and hand; OSS, Oxford Shoulder Score; CMS, Constant Murley Score.

B, regresjonskoeffisient

KI, konfidensintervall

P, p-verdi

#### 4.5 Sammenheng mellom lav CMS skår og uavhengige bakgrunnsvariabler

Resultatene fra ujusterte regresjonsanalyser viste at det var 8 ganger større odds ratio (OR) for lav CMS skår for en person med ASA 3 sammenlignet med en person med ASA 1 (OR= 8,0; 95% KI 1,4 – 45,8;  $p < 0,05$ ) hos  $n=73$  (Tabell 13). Resultatene viste også høyere odds for lav CMS skår for kvinner (OR= 1,71; 95% KI 0,50-5,83;  $p=0,39$ ) og for pensjonist/uføre (OR= 2,06; 95% KI 0,75-5,56;  $p=0,16$ ) hos  $n=73$ , men ingen av resultatene var statistisk signifikante. Justert regresjonsanalyse viste at det er 10,7 ganger større odds for lav CMS skår for en person med ASA 3 sammenlignet med en person med ASA 1 hos  $n=73$  (OR= 10,70; 95% KI 1,4 - 80,9;  $p < 0,05$ ) og 30,6 ganger større odds for lav CMS skår for en person med ASA 3 sammenlignet med en person med ASA 1 hos  $n=71$  (OR= 30,63; 95% KI 2,4 -384,3;  $p < 0,05$ ) (Tabell 13).

Figur 24 viser spredningsdiagram mellom uavhengig variabel Leverage verdi og avhengig variabel predikert sannsynlighet og figur 25 viser spredningsdiagram mellom uavhengig variabel Cook's D og avhengig variabel predikert sannsynlighet. Ved analyser uten de to uteliggerne (Figur 25) økte tre av koeffisient-estimatene i modellen med mer enn 15%. OR for pensjonist/ufør økte fra 1,99 til 2,84 (30% økning), OR ASA 2 økte fra 2,87 til 3,78 (31,5% økning) og OR ASA 3 fra 10,70 til 30,63 (186% økning). Svært stort konfidensintervall og manglende oppfyllelse av forutsetningene indikerer stor usikkerhet i estimatene. Det var høyere odds for lav CMS skår ved å være pensjonist/ufør (OR=1,99; 95% KI 0,45-8,90;  $p=0,37$ ), men ingen betydelig forskjell i odds når det gjaldt de andre uavhengige variablene (alder, kjønn, BMI og røyking) i justerte analyser og det var ikke statistisk signifikante resultater med  $p > 0,05$  og 95% KI som krysser 0.

**Tabell 13:** Logistisk regresjonsmodell for å estimere oddsen for lav CMS skår med uavhengige bakgrunnsvariabler (n=73)

Uavhengig variabel	N=73		N=73		N=71	
	Ujustert OR (95%KI)	p- verdi	Justert OR (95%KI)	p- verdi	Justert OR (95%KI)	p- verdi
Alder	1,03 (0,98-1,08)	0,27	0,96 (0,88-1,05)	0,35	0,94 (0,86-1,03)	0,21
Kjønn (kvinne)	1,71 (0,50-5,83)	0,39	1,19 (0,28-5,10)	0,81	1,04 (0,22-4,84)	0,96
BMI	0,99 (0,90-1,09)	0,83	0,98 (0,87-1,10)	0,67	0,96 (0,86-1,08)	0,51
Røyk	1,03 (0,19-5,48)	0,97	0,71 (0,10-4,86)	0,73	0,66 (0,09-4,81)	0,68
Pensjonist/ufør	2,06 (0,75-5,56)	0,16	1,99 (0,45- 8,90)	0,37	2,84^ (0,58-13,90)	0,20
ASA 1	(ref.)	0,05	(ref.)	0,07		
ASA 2	2,24 (0,80-6,28)	0,13	2,87 (0,77- 10,71)	0,12	3,78^ (0,95-15,01)	0,06
ASA 3	8,00 (1,40-45,76)	0,02*	10,70 (1,40- 80,90)	0,02*	30,63^ (2,44-384,34)	0,01*

Avhengig dikotom kategorisk variabel: CMS skår >56=0, <56=1

CMS, Constant Murley Score.

\* Signifikant p<0,05

^ >15% økning i estimatene uten to observasjoner med høy Cook's D verdi

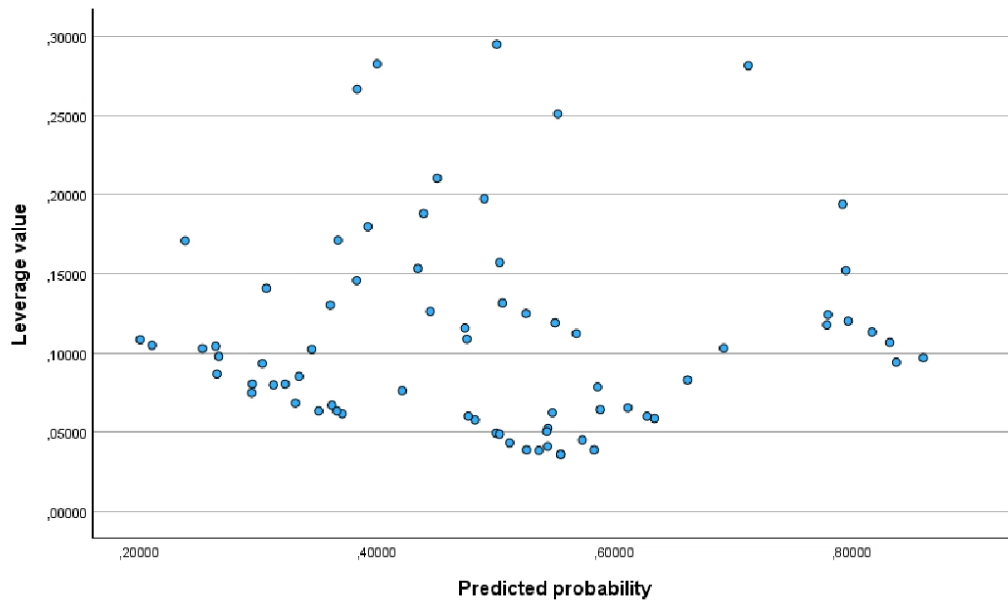
OR, odds ratio

KI, konfidensintervall

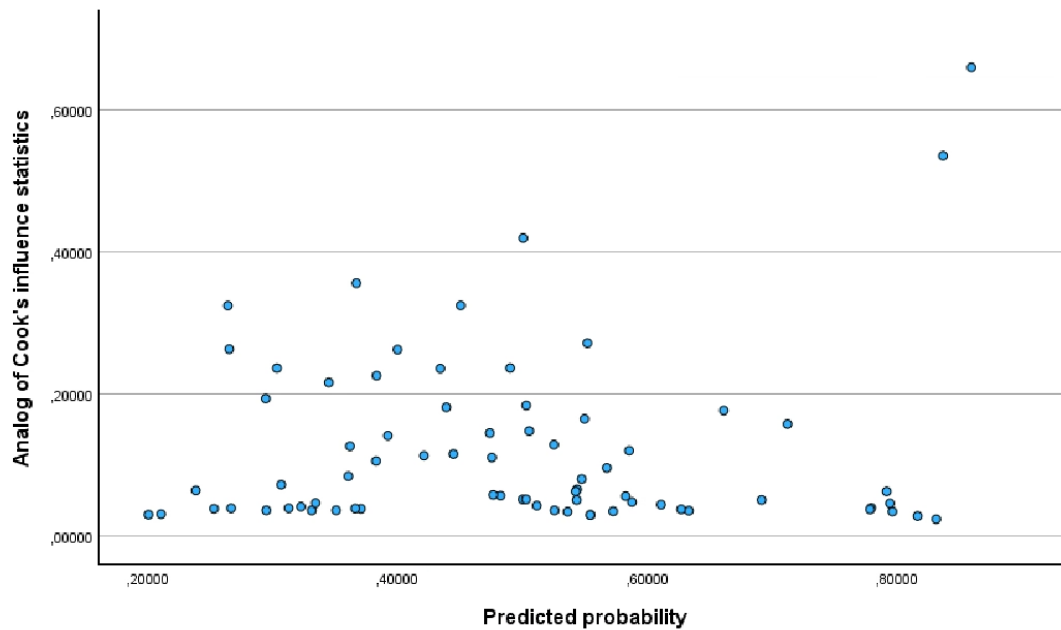
Ref, referansegruppe



**Figur 24:** Spredningsdiagram mellom uavhengig variabel Leverage verdi og avhengig variabel predikert sannsynlighet (n=73)



**Figur 25:** Spredningsdiagram mellom uavhengig variabel Cook's D og avhengig variabel predikert sannsynlighet (n=73)



## 5.0 Diskusjon

### 5.1 Metodediskusjon

Metodekvaliteten er avgjørende for om resultatene i masteroppgaven er til å stole på (Carter & Lubinsky, 2016, s. 76). I metodediskusjonen vil masteroppgavens metodiske kvalitet drøftes ut ifra intern validitet, statistisk validitet og ekstern validitet. Intern validitet vurderes ut ifra risiko for systematiske feil (seleksjonsskjevhet, måleskjevhet og konfunderende variabler), statistisk validitet vurderes ut ifra om det er benyttet hensiktsmessige statistiske analysemetoder for å besvare problemstillingen og ekstern validitet vurderes ut ifra om resultatene er generaliserbare til andre enn de som ble målt (populasjonen) (Carter & Lubinsky, 2016, s. 76). Reliabiliteten vurderes også ut ifra hvilke målemetoder som er benyttet.

#### 5.1.1 Studiedesign

Masteroppgaven er en tverrsnittstudie, og i en tverrsnittstudie er fenomenet målt på ett tidspunkt (Polit & Beck, 2017, s. 168). Tverrsnittstudier kan gi deskriptiv kunnskap om et fenomen og kunnskap om sammenheng mellom fenomener på et måletidspunkt, men ikke kunnskap om årsaksforhold (Carter & Lubinsky, 2016, s. 143; Polit & Beck, 2017, s. 168). For å besvare problemstillingene om postoperativ funksjon, undersøke eventuelle tak- eller gulveffekter ved måleinstrumenter, sammenheng mellom ulike måleinstrumenter, og sammenheng mellom uavhengige variabler og en avhengig variabel på ett måletidspunkt er tverrsnittstudie ett egnet design. Å identifisere sammenhenger mellom variabler, kan også foreslå områder for utforskning for fremtidige eksperimentelle studier hvor kunnskap om årsakssammenhenger kan identifiseres (Carter & Lubinsky, 2016, s. 148).

Gitt at masteroppgaven hadde hatt tilgjengelig data for flere måletidspunkter hos utvalget, kunne blant annet kunnskap om måleinstrumentenes reliabilitet og responsivitet blitt undersøkt i et longitudinelt studiedesign (Carter & Lubinsky, 2016, s. 255). Imidlertid var ikke to års oppfølging i datainnsamlingen fullført før masteroppgaven, så data fra flere måletidspunkter var ikke tilgjengelig for masterkandidaten. I lys av problemstillingene og tidsbegrensningen var derfor tverrsnittstudie et egnet design.

### 5.1.2 Utvalg og data

Utvalget i denne oppgaven bestod av flest kvinner (82%) og er sammenlignbart med andre studier som omhandler av PHF, hvor andelen kvinner er rapportert fra 67-90% (Bahrs et al., 2014; Bergdahl et al., 2016; Clement, Duckworth, et al., 2014; Fraser et al., 2020; Geiger et al., 2010; Gracitelli et al., 2016; Olerud et al., 2011; Plath et al., 2019; Sumrein et al., 2017). Plath et al. (2019) rapporterte 75% kvinner, Gracitelli et al. (2016) 72% kvinner og Geiger et al. (2010) 71% kvinner, imidlertid bestod utvalget henholdsvis av to-fire parts, to og tre-parts og to-fire parts PHF. Olerud et al. (2011) inkluderte kun fire-parts PHF og fant 85% kvinner i utvalget, mens Fraser et al. (2020) fant 90% kvinner i utvalg bestående av 3 og 4-parts i Norge. Årsaken til at masteroppgavens utvalg hadde spesielt stor kjønnskjevhet, kan være at kun tre og fire parts PHF ble inkludert, og det er vist at sannsynligheten for tre og fire-parts PHF er henholdsvis 5.3 og 5.1 ganger høyere hos kvinner over 60 år sammenlignet med menn over 60 år (Bahrs et al., 2014). Samtidig kan høy andel kvinner generelt også forklares med at osteoporose har vist sterk sammenheng med 70-80% av alle PHF (Court-Brown & Caesar, 2006) og det er kjent at 22% av kvinner over 50 år har osteoporose, sammenlignet med 7% blant menn over 50 år (Kanis et al., 2021). Høyere andel kvinner (90%) i Fraser et al. (2020) kan trolig forklares med høyere gjennomsnittsalder (75 år), da studien inkluderte kun personer over 65 år. Masteroppgavens høye andel kvinner virker tilfeldig og relativt lik andre sammenlignbare studier, så det ansees derfor ikke å være en trussel for den interne validiteten.

Gjennomsnittsalderen i utvalget var 66 år ved skadetidspunkt, og det var signifikant forskjell i alder mellom kvinner og menn (68 vs. 59). Aldersforskjellen kan skyldes skademekanismen fordi Bahrs et al. (2013) fant at forskjøvet PHF var hyppigst rapportert blant kvinner over 60 år og menn under 60 år, og resultatene viste en signifikant sammenheng mellom lavenergitraume og kvinner over 60 år og høyenergitraume og menn under 60 år. Gracitelli et al. (2016) fant at 80% av PHF var fra lavenergitraume, men differensierte ikke på kjønn (72% kvinner). Olerud et al. (2011) fant gjennomsnittsalder 77 år (79 for kvinner og 70 år menn), imidlertid inkluderte denne studien kun fire-parts PHF med alder over 55 år. Fraser et al. (2020) oppga gjennomsnittsalder 75 år hos tre og fire-parts PHF, men inkluderte kun deltakere over 65 år. Studier som har inkludert flere bruddtyper har rapportert sju til ti år høyere gjennomsnittsalder for kvinner, 62 vs. 72 (Sumrein et al., 2017), 62 vs. 71 (Bergdahl et al., 2016) og 63 vs. 56 (Geiger et al., 2010). Clement, Duckworth, et al. (2014) rapporterte en økt andel forskjøvet PHF hos de med høy alder, og foreslo økt andel med osteoporose i høyere alder som mulig årsak. Det er ikke tilgjengelig data om skademekanisme eller andel

med osteoporose i denne oppgaven, men på bakgrunn av høy gjennomsnittsalder og stor andel kvinner, kan det antas at en stor andel i utvalget har falt fra egen høyde og har osteoporose. I et lignende utvalg (tre og fire-parts PHF) i en studie fra Norge hadde majoriteten falt fra egen høyde (Fraser et al., 2020). Den høye gjennomsnittsalderen i utvalget og forskjellen i andel kvinner og menn i masteroppgaven virker tilfeldig og relativt lik sammenlignbare studier, men mangel på data fra mulig konfunderende variabler som skademekanisme og osteoporose gjør det usikkert om det er en trussel for den interne validiteten.

Utvalgets BMI var gjennomsnittlig 28 i masteroppgaven, dette samsvarer med en sammenlignbar studie med fire-parts PHF som rapporterte 27 hos utvalg over 55 år (Olerud et al., 2011). Røyking ble rapportert hos 8% i utvalget, og det er sammenlignbart med andel røykere i Norge (7%) (Helsedirektoratet, 2018), imidlertid har en lignende studie fra Norge rapportert 5% røykere (Fraser et al., 2020).

Andelen pensjonist eller ufør var 67% i masteroppgaven, og det var signifikant forskjell mellom kjønn, henholdsvis 75% blant kvinner og 31% blant menn. Denne forskjellen er trolig tilfeldig, og kan forklares med at det var flere menn i arbeidsfør alder sammenlignet med kvinner i utvalget. Clement, Duckworth, et al. (2014) fant at 88% ikke var i arbeid, men utvalget bestod av PHF over 65 år, så det er ikke sammenlignbart med masteroppgavens utvalg.

ASA skår en-tre ved operasjon var henholdsvis 37%, 36% og 10% i utvalget, og ikke signifikant forskjell mellom kjønn. En sammenlignbar studie av Plath et al. (2019) fant ASA 1 (25%), ASA 2 (40%) og ASA 3 (34%), imidlertid inkluderte denne studien kun personer over 60 år og det er kjent at komorbiditeter generelt øker med alderen (Rangan et al., 2015).

Masteroppgavens utvalg bestod av 73 personer. Utvalgsstørrelsen kan påvirke resultatene, og generelt er større utvalg mer representative for populasjonen sammenlignet med små utvalg, og større utvalg gir ofte høyere statistisk styrke (Carter & Lubinsky, 2016, s. 103-104). Strykeberegning ble ikke gjort i denne oppgaven da sekundærdata ble benyttet og utvalgsstørrelsen var dermed forhåndsbestemt. På en annen side har sammenlignbare studier utvalg bestående av 44 personer (Beastall et al., 2012) og 76 personer (Maurer et al., 2021), så det kan derfor antas at utvalgsstørrelsen er tilstrekkelig og ikke en trussel for den interne validiteten.

Frafall i oppgaven var 4%, da deltakerne ikke ville møte fysisk til seks måneders postoperativ oppfølging på grunn av covid-19. Frafallsskjevhet ansees å være til stede i tverrsnittstudier

hvis det er 60% etter inklusjon (Johnson & Wislar, 2012). Oppgavens høye responsrate kan forklares med at måleinstrumentene ble administrert ved fysisk oppmøte, men samtidig var dette begrensningen for 4% under pandemien. På grunn av lavt antall frafall i masteroppgaven, er trolig ikke frafallsskjevhet en trussel for den interne validiteten.

Oppsummert virker ikke systematisk seleksjonsskjevhet å være en trussel for masteroppgavens interne validitet, og utvalget er representativt for populasjonen og dette bedrer resultatenes generaliserbarhet (Polit & Beck, 2017, s. 224).

### **5.1.3 Målemetoder**

Målemetodene i masteroppgaven var forutbestemt, siden de inngikk i en randomisert studie. Måleinstrumentene er standardiserte, validerte og mye brukt for å måle skulder og armfunksjon hos pasienter med PHF (Richard et al., 2020). Måleinstrumentene i masteroppgaven måler flere domener og resultatene oppsummeres i en sumskår, og gir derfor ikke skår på et spesifikt konstrukt som for eksempel smerte, ADL eller ROM. Dette kan potensielt gi en måleskjevhet fordi informasjon om hvilke konstrukter som eventuelt skårer høyt og- eller lavt mangler. På en annen side vil slike måleinstrumenter kunne dekke flere aspekter i ICF-modellen og derfor kunne bedre den helhetlige forståelsen av funksjonsbegrensningen (van de Water et al., 2014). Imidlertid er måleinstrumentene i masteroppgaven kritisert for at de primært måler begrensninger i kroppsstrukturer, funksjon og aktiviteter, og i mindre grad deltakelse, miljø og personlige faktorer (Nowak et al., 2021).

I litteraturen er det diskrepans når det gjelder psykometriske egenskaper for måleinstrumentene benyttet for PHF (Nowak et al., 2019; van de Water et al., 2014; Vrotsou et al., 2018). DASH og OSS er oversatt og validert til norsk (Finsen, 2008; Fjalestad et al., 2014). CMS er ikke oversatt til norsk, og derfor ble PROMs delen oversatt muntlig av masterkandidaten. Bruk av ikke-oversatte PROMs kan ha ført til måleskjevhet eller feiltolkning som kan ha påvirket reliabiliteten og validiteten ved CMS (McKown et al., 2020). Til tross for at CROM-styrketest i CMS ble utført i tråd med Constant et al. (2008) retningslinjer, var det vanskelig å unngå måleskjevhet da pasientene benyttet stor grad av kompenserende bevegelser. En mulig årsak til kompenserende bevegelser kan være at testen var for krevende å utføre for pasientgruppen og dette viste seg også ved stort frafall på testen. CROM-styrketest av CMS er utsatt for stor kritikk på grunn av stort potensiale for gulveffekt og lav reliabilitet (Slobogean & Slobogean, 2011) noe som gjenspeiles i datainnsamlingen i

masteroppgaven. Det er rapportert usikkerhet når det gjelder CMS reliabilitet hos pasienter med PHF (Lillkrona, 2008; van de Water et al., 2014; Yian, Ramappa, Arneberg & Gerber, 2005). Manglende justering for kjønn og alder har blitt tilskrevet ansvar for instrumentets måleskjevhet, det ble derfor utviklet en relativ CMS skår justert for kjønn og alder basert på normative verdier hos den friske befolkningen (Constant, 1986; Katolik et al., 2005). Katolik et al. (2005) presenterte en formel for utregning av kjønn og aldersjustert CMS skår (Absolutt CMS skår/normativ gjennomsnittsskår) x 100). Formelen har blitt kritisert for å overvurdere funksjonsbedring hos eldre (Roy, MacDermid & Woodhouse, 2010). Bruk av en relativ CMS skår i publiserte studier har imidlertid resultert i vanskeligheter med å sammenligne resultater mellom studier (Ye et al., 2013). I litteraturen rapporteres oftest absolutt CMS skår (Richard et al., 2020). Overestimering av skulderstyrke i CMS er også rapportert som årsak til måleskjevheten (Lillkrona, 2008). I styrketesten kan det oppnås maksimalt 25 poeng, og det er vist at selv hos yngre og friske er det vanskelig å oppnå maksimal skår på styrkedomenet (Johansson & Adolfsson, 2005). De ulike domene (smerte, ADL, ROM og styrke) har ulik samleskår, og rasjonale bak denne fordelingen er ukjent og kan også bidra til måleskjevhet ved instrumentet benyttet hos pasienter med PHF (Lillkrona, 2008; van de Water et al., 2014; Yian et al., 2005).

Målemetodene i masteroppgaven omfatter både PROMs og CROMs. Mens CROMs kan gi mer nøyaktig mål på patoanatomiske og patofysiologiske endringer (Harreld, Clark, Downes, Virani & Frankle, 2013), gjenspeiler det ikke nødvendigvis pasientens dagligdagse funksjon og pasientens egen opplevelse av symptomer, og derfor kan PROMs gi bedre innsikt i pasientens utfordringer (van de Water et al., 2016). Erfaring fra datainnsamlingen var at noen pasienter oppnådde svake resultater på CROMs delene av CMS, men samtidig var smertefrie og hadde tilstrekkelig funksjon til å utføre sine dagligdagse aktiviteter. I en slik kontekst er ikke grad av styrke eller ROM meningsfylt for pasienten, og det påvirker måleinstrumentets innholds og konstruktvaliditet (van de Water et al., 2014).

Under datainnsamlingen hadde pasientene på t-skjorte (for å skjule operasjonsarr) slik at intervjuerskjevhet ikke var til stede og dermed intern validitet ivaretatt. Ved bruk av CROMs i intervensjonsstudier er det viktig at klinikerer er blindet for behandlingsmetode slik klinikerens eventuelle egeninteresser (intervjuerskjevhet) ikke gir skjevhet i målingene (Harreld et al., 2013). PROMs gir ikke skjevhet av en tredjeparts tolkning og kan derfor være et mer pålitelig mål (Copay et al., 2007), men kan gi skjevhet i målingene hvis pasient ikke ønsker å svare oppriktig, overvurderer egen funksjon eller har kognitiv svikt (Althubaiti,

2016; Polit & Beck, 2017, s. 172). Systematisk selvrappoteringskjevhet på grunn av egeninteresser er lite sannsynlig i dette utvalget, ettersom pasientene var blindet for operasjonsmetode og pasienter med kognitiv svikt ikke ble inkludert. På en annen side måler OSS funksjon siste fire uker og DASH funksjon siste uken. Det kan være vanskelig å huske tilbake i tid, spesielt hos den eldre befolkningen (Pais, Ruano, O & Barros, 2020), og på bakgrunn av masteroppgavens relativt høye gjennomsnittsalder er det usikkert om systematisk selvrappoteringskjevhet truer den interne validiteten.

Pasientbyrde ble ikke målt i masteroppgaven, men fortvilelse over gjentagende spørsmål kom stadig til uttrykk under datainnsamlingen, og det er vist at opplevd pasientbyrde ved PROMs kan være spesielt uttalt hos eldre (Stone, Shiffman, Atienza & Nebeling, 2007). Det mangler fortsatt kunnskap om hvordan aldersrelaterte endringer i kognitive og kommunikative funksjoner påvirker besvarelse av PROMs (Knäuper et al., 2016), men det er foreslått at kortere PROMs og mindre komplekse spørsmål ga mindre grad av pasientbyrde og påfølgende lavere selvrappoteringskjevhet (Fox, Sidani & Streiner, 2007). På en annen side var det ingen manglende data fra PROMs i masteroppgaven og lav andel manglende data kan være indikasjon på lav pasientbyrde og- eller at egnede målemetoder er benyttet (U.S. Department of Health, 2006). Imidlertid kan lav andel manglende data i masteroppgaven også forklares med at masterkandidaten kontrollerte alle spørreskjemaene etter besvarelse, og eventuelt manglende data ble deretter etterfylt. I tillegg hadde en stor del av pasientene med seg pårørende for assistanse til å besvare PROMs. Eventuell opplevd pasientbyrde på grunn av mange spørsmål fra PROMs, og- eller påvirkning fra pårørende i besvarelsen av PROMs kan ha gitt selvrappoteringskjevhet og er en mulig trussel for den interne validiteten. Se punkt 5.2.1.1- 5.2.1.3 i resultatdiskusjon for ytterligere diskusjon av måleinstrumentene.

#### **5.1.4 Generaliserbarhet**

Generaliserbarheten (ekstern validitet) av studien innebærer om resultatene fra studien kan generaliseres til andre grupper eller populasjoner (Carter & Lubinsky, 2016, s. 88).

Utvalget består kun av tre og fire-parts PHF og alle er behandlet med operasjon. Clement, Duckworth, et al. (2014) fant at tre og fire-parts PHF hos 109 personer hadde signifikant dårligere funksjon sammenlignet med en-parts PHF hos 278 personer ett år postoperativt målt med CMS (57 vs. 68;  $p < 0,01$ ), og tilsvarende resultater for konservativt behandlet og opererte (65 vs. 53,  $p < 0,01$ ). Utvalgets høye interne validitet (kun opererte tre og fire-parts PHF), går

på bekostning av resultatenes eksterne validitet (Pannucci & Wilkins, 2010), slik at masteroppgavens resultater ikke kan generaliseres til majoriteten av PHF, som er konservativt behandlet en-parts PHF.

Majoriteten i utvalget består av kvinner, med relativt høy gjennomsnittsalder, så resultatene kan trolig ikke generaliseres til yngre pasienter med tre-fire parts PHF og- eller menn uavhengige av alder. Imidlertid er flertallet av tre og fire-parts PHF kvinner over 65 år (Bergdahl et al., 2016; Iglesias-Rodríguez et al., 2021), så utvalget har god ekstern validitet og er representativt for populasjonen.

Seleksjonsskjevhet kan være til stede fordi alle pasientene ble rekruttert på samme sykehus, men samtidig styrker det ekstern validitet at sykehusets opptaksområde dekker et stort geografisk område, som inkluderer 24 ulike kommuner og til sammen nesten 600 000 innbyggere. Alle tre og fire-parts PHF som ble vurdert som operasjonskrevende ble forespeilet å delta i hovedstudien. Begge behandlingsalternativene var operasjon, dermed ble kun pasienter som samtykket til operasjon inkludert. I tillegg hadde hovedstudien en rekke inklusjons- og eksklusjonskriterier (Tabell 6). Det kan derfor ha blitt en seleksjonsskjevhet hvis et stort antall ikke ønsket operasjon, oppfylte eksklusjons eller ikke oppfylte inklusjonskriteriene (Pannucci & Wilkins, 2010). Det var ikke tilgjengelig data om de som ble ekskludert til bruk i masteroppgaven. Utvalget består blant annet ikke av pasienter med kognitiv svikt eller manglende norskkunnskaper, da pasienter med manglende evne til å lese og forstå norsk og pasienter som ikke var i stand til å forstå skriftlig samtykke ikke ble inkludert i hovedstudien. Sammenlignbare studier har brukt tilsvarende eksklusjonskriterier (Fraser et al., 2020; Gracitelli et al., 2016; Olerud et al., 2011; Plath et al., 2019). Dette kan forklares med at målemetodene er PROMs og at pasienter med kognitiv svikt ikke oppfyller kriteriene til å svare på PROMs (Kramer & Schwartz, 2017). Andelen med kognitiv svikt er ikke kjent hos pasienter med PHF, men kognitiv svikt er en kjent risikofaktor for fragile brudd generelt (Friedman & Mendelson, 2014). Derfor kan det antas at en andel PHF har kognitiv svikt, men resultater fra denne oppgaven kan ikke generaliseres til denne populasjonen. Manglende evne til å lese og forstå norsk kan også ha ekskludert en andel ikke-etnisk norske pasienter, men det er ikke tilgjengelig data til å kunne vite dette. Denne systematisk underrepresentasjonen av grupper som kan være relevant for forskningsspørsmålet gir en grad av seleksjonsskjevhet og er en trussel for den interne validiteten og dette påvirker generaliserbarheten (Polit & Beck, 2017, s. 251).



Måletidspunkt seks måneder postoperativt vil kanskje vurderes som kort sammenlignet med annen litteratur (Richard et al., 2020). Imidlertid fant Rangan et al. (2015) størst funksjonsbedringen ved operert forskjøvet PHF de første seks månedene hos 125 personer målt med OSS og ikke stor endring i OSS skår etter 12 måneder (38 vs. 39). På en annen side fant Olerud et al. (2011) funksjonsbedring ved operert fire-parts PHF mellom fire og 12 måneder målt med CMS (36 vs. 49) og DASH (43 vs. 32). Gracitelli et al. (2016) fant også funksjonsbedring mellom seks og 12 måneder målt med CMS (64 vs. 71) men ikke særlig funksjonsbedring målt med DASH (19 vs. 16). Det er derfor usikkert om resultatene fra masteroppgaven kan generaliseres til populasjoner med måletidspunkter over seks måneder postoperativt.

### 5.1.5 Statistiske analyser

På grunn av stor kjønnskjevhet i utvalget, ble deskriptive analyser av måleinstrumentene også utført gruppert på kjønn. Subgruppen menn hadde lav utvalgsstørrelse ( $n=13$ ), og stor forskjell mellom gruppene kan påvirke resultatenes eksterne og statistiske validitet (Dickinson, Adelson & Owen, 2012). Til tross for lavt antall menn var forutsetningene for parametrisk test oppfylt med normalfordelt data og Levene's test (lik varians i gruppene) (Carter & Lubinsky, 2016, s. 286). Gjennomsnittlig forskjell i skår på måleinstrumentene mellom kjønn ble testet med parametrisk to utvalgs t-test, og parametriske tester antas å være mer robuste enn ikke-parametriske tester (Carter & Lubinsky, 2016, s. 283).

Kjønnskjevheten i utvalget kan påvirke resultatenes eksterne validitet, derfor ble analyser også utført kun for kvinner. Det var ikke betydelig forskjell mellom resultatene fra Pearsons korrelasjon og- eller enkel lineær regresjon kun for kvinner sammenlignet med hele utvalget . Se resultater i vedlegg 8-12. Sammenlignbare studier med henholdsvis 80% (van de Water et al., 2014), 90 % (Fraser et al., 2020) og 70% kvinner (Ye et al., 2013) har ikke presentert analyser separat for kvinner.

Ved analyser av korrelasjon mellom data fra måleinstrumentene var forutsetninger for parametrisk Pearsons korrelasjon delvis oppfylt. Det var tilstrekkelig normalfordeling utvalgsstørrelsen tatt i betraktning og lineært forhold, men det var noen uteliggere (Vedlegg 4-6). Når forutsetningene ikke er oppfylt, skal i utgangspunktet ikke-parametrisk Spearman's rho benyttes (Carter & Lubinsky, 2016, s. 321). Analysene ble derfor utført med Spearman's rho og Pearsons korrelasjon med og uten uteliggere. Det var ubetydelig forskjell mellom

resultatene og det ble derfor valgt å oppgi resultatene fra Pearsons korrelasjon. Ved analyser av sammenheng mellom data fra måleinstrumentene var forutsetningene for enkel lineær regresjon usikker på grunn av usikker homoskedastisitet (residualer har kun delvis lik varians) og residualer over -3 for CMS og DASH i spredningsdiagram (Vedlegg 7). Analysene ble utført med og uten uteliggere, og på en måleskala fra 0-100 er ikke en differanse på 0,15 poeng av betydning og det ble derfor ansett som akseptabelt å benytte lineær regresjonsanalyse. Imidlertid vil disse vurderingene kunne påvirke masteroppgavens statistiske validitet og derfor også i hvilken grad en kan stole på resultatene (Carter & Lubinsky, 2016, s. 76).

Når det gjelder odds ratio for lav CMS skår forklart av uavhengige bakgrunnsvariabler, ble logistisk regresjon benyttet. Det finnes ingen kjent grense for antall uavhengige variabler, men det anbefales å bruke få variabler med kjent sammenheng (Polit & Beck, 2017, s. 417). Kravet til antall hendelser per uavhengig variabel og dens påvirkning på den statistiske validiteten i logistisk regresjon er diskutert i litteraturen. Peduzzi, Concato, Kemper, Holford og Feinstein (1996) fant at lavt antall hendelser per variabel påvirket styrken på estimatene, men også at uavhengige variabler med både lav og høyt antall hendelser kunne gi skjevhet i resultatene. Dette ble forklart med enten lav prevalens eller lav sammenheng kan forsterke effekten av lavt antall hendelser per uavhengig variabel. Sammenlignbare studier har likt antall eller flere uavhengige variabler, men høyere antall hendelser per uavhengig variabel og to-seks ganger større utvalgsstørrelse (Clement, Duckworth, et al., 2014; Miquel et al., 2021; Südkamp et al., 2011).

En «tommelfingerregel» er at det bør være minst 10 hendelser per uavhengig variabel (Vittinghoff & McCulloch, 2007). Den uavhengige variabelen «røyker» i regresjonsmodellen hadde et lavt antall hendelser (n=6) i masteroppgaven. Røyking er assosiert med høyere forekomst av manglende tilheling og postoperative infeksjoner etter bruddbehandling (Xu, Anderson, Park, Chen & Lee, 2021), og en kan derfor anta at røyking også påvirker postoperativ skulderfunksjon ved PHF. På bakgrunn av lavt antall hendelser i variabelen ble analysene også utført uten variabelen røyking, uten betydelig påvirkning av de andre estimatene i modellen (Vedlegg 13). Andre uavhengige variabler i analysen (alder, kjønn, BMI, arbeidsstatus, ASA skår (komorbiditet)) har vist sammenheng med dårlig skulderfunksjon etter PHF (Fisher, Driesman, Saleh, Egol & Konda, 2022; Südkamp et al., 2011).

Imidlertid er en begrensning med analysen at det ikke er kontrollert for eventuell sammenheng med lav CMS skår fra konfunderende variabler det ikke er tilgjengelig data for (Polit & Beck, 2017, s. 162). Det er vist at psykososiale faktorer som bevegelsesfrykt, smertetoleranse, sivilstatus og sosialt nettverk, og postoperative komplikasjoner (Jayakumar et al., 2019), sosial og fysisk selvstendighet før skadetidspunkt og behandlingsmetode (Clement, Duckworth, et al., 2014), diabetes og osteoporose (Kruithof, Formijne Jonkers, van der Ven, van Olden & Timmers, 2017) har sammenheng med skulderfunksjon etter PHF. For å styrke resultatene burde flere potensielt konfunderende variabler vært med i analysene, men masteroppgaven benytter sekundærdata fra en RCT og manglende relevante variabler er en ulempe ved å benytte data som er samlet inn til et annen formål, da det begrenser besvarelsen av problemstillingene i masteroppgaven (Hox & Boeije, 2005). Imidlertid kan påvirkning av ukjente eller manglende konfunderende variabler kontrolleres for bare ved randomisering av utvalg, som i en RCT (Pannucci & Wilkins, 2010).

Den største trusselen for den statistiske validiteten i den logistiske regresjonsmodellen var at forutsetningene ikke var oppfylt fordi det ble identifisert to observasjoner med høye Cook's D verdier, og da analysene ble utført uten disse observasjonene påvirket det tre andre estimater i modellen i betydelig grad (over 15% økning). Den statistiske validiteten i masteroppgavens er derfor truet. For å ivareta statistisk validitet burde det vært utprøvd transformasjon eller en robust modell, men det ble vurdert til å falle utenfor rammene for en masteroppgave. For lav utvalgsstørrelse, heterogent utvalg, potensielt manglende konfunderende variabler og eventuelt for liten andel hendelser i variablene ASA og Arbeid kan være årsaken til observasjonenes store påvirkning på estimatene. Urealistisk stort konfidensintervall kan indikere for lavt antall hendelser per uavhengige variabel (Vittinghoff & McCulloch, 2007), noe som gjenspeilet resultatene i masteroppgaven da det kun var 10 hendelser i variabelen ASA 3 og svært stort konfidensintervall. Imidlertid var det 24 hendelser i variabelen Arbeid, men der var også konfidensintervallet stort. Resultatene fra logistisk regresjon har lav statistikk validitet og resultatene er ikke til å stole på (Carter & Lubinsky, 2016, s. 73).

### **5.1.6 Styrker og svakheter**

Styrker og svakheter i masteroppgaven er diskutert i metodediskusjonen punkt 5.1.1-5.1.5. Oppsummert er masteroppgavens design egnet til å besvare problemstillingene, utvalget ansees som representativt for populasjonen og det var lavt frafall, imidlertid er det

metodologiske svakheter på grunn av mulig måleskjevhet, manglende potensielle konfunderende variabler og trussel av statistisk validitet på grunn av manglende oppfyllelse av forutsetninger i logistisk regresjon.

## **5.2 Resultatdiskusjon**

Hensikten med masteroppgaven var å øke kunnskapen om postoperativ funksjon, undersøke eventuelle tak- eller gulveffekter ved måleinstrumentene CMS, DASH og OSS, å sammenligne nevnte måleinstrumenter hos pasienter med tre og fire-parts PHF ved å undersøke sammenheng. På grunn av erfaringsmessig større pasient- og klinikerbyrde ved CROMs delen i CMS og manglende data på styrkedomenet, var det ønskelig med ytterligere kunnskap om dette måleinstrumentet for å bedre kunne vurdere instrumentets egnethet for pasienter operert for tre og fire-parts PHF ved å undersøke sammenheng mellom uavhengige bakgrunnsvariabler og lav CMS skår. Resultatene viste gjennomsnittlig dårlig skulderfunksjon målt med CMS og gjennomsnittlig god skulderfunksjon målt med DASH og OSS (Tabell 8). Det ble også identifisert mulig takeffekt hos menn ved bruk av OSS og mulig gulveffekt på styrkedomenet ved CMS uavhengig av kjønn (Tabell 10). Det var sterk signifikant korrelasjon mellom måleinstrumentene (Tabell 11) og signifikant negativ sammenheng mellom CMS og DASH, signifikant negativ sammenheng mellom DASH og OSS og signifikant positiv sammenheng mellom CMS og OSS (Tabell 12). Sammenhengen var sterkest mellom PROMs. Til tross for høy sammenheng mellom måleinstrumentene var det lav enighet mellom CROMs og PROMs (Tabell 8 og 9). ASA skår 3 var eneste uavhengige variabel som viste signifikant høyere odds ratio for lav CMS skår i justert regresjonsmodell (Tabell 13).

### **5.2.1 Deskriptiv statistikk av data fra måleinstrumentene**

#### **5.2.1.1 Postoperativ funksjon målt med CMS**

Resultatene fra deskriptive analyser viste at CMS gjennomsnittsskår var 54, gjennomsnittlig skår blant kvinner var signifikant lavere enn blant menn (52 vs. 61,  $p < 0,05$ ). Resultatene i masteroppgaven viste dårligere gjennomsnittsskår sammenlignet med andre lignende studier som har rapportert CMS skår 59, 64 og 90 (Fraser et al., 2020; Gracitelli et al., 2016; Zhu et

al., 2011). Dette kan skyldes ulikheter i sammensetningen av utvalg, siden noen hadde mindre andel kvinner (Gracitelli et al., 2016; Zhu et al., 2011) og referanseverdiene (Constant, 1986) er vist å være lavere for kvinner sammenlignet med menn (Tabell 1) og det er en kjent svakhet at måleinstrumentet ikke har tilstrekkelig justering for kjønn (Katolik et al., 2005; Yian et al., 2005). På en annen side hadde Fraser et al. (2020) utvalg bestående av tre og fire-parts PHF høyere andel kvinner (90%), som trolig kan skyldes at kun personer over 65 ble inkludert, da det er vist økende andel tre og fire-parts PHF blant kvinner over 60 år (Bahrs et al., 2013; Bahrs et al., 2014). Høyere gjennomsnittsskår i sammenlignbare studier (Gracitelli et al., 2016; Zhu et al., 2011) kan også forklares med at utvalgene hadde andre bruddtyper da det er vist at fire-parts PHF har dårligere funksjon målt med CMS etter ett år, sammenlignet med to og tre-parts PHF (53 vs. 61) (Clement, Duckworth, et al., 2014). CMS gjennomsnittsskår lavere enn masteroppgavens ble presentert i Olerud et al. (2011), hvor utvalget bestod av kun fire-parts PHF (operert med halvprotese), og dette underbygger antagelsene om at bruddtypen påvirker CMS gjennomsnittsskår.

Gracitelli et al. (2016) og Fraser et al. (2020) spesifiserte ikke av hvordan styrketesten ble utført, og det kan også bidratt til skjevhet i resultatene hvis ikke retningslinjene til Constant et al. (2008) ble fulgt, imidlertid spesifiserte Zhu et al. (2011) at de fulgte retningslinjene på bevegelses og styrkedomenet av (Constant et al., 2008). Høyere gjennomsnittsalder i masteroppgavens utvalg sammenlignet med Zhu et al. (2011) kan også ha bidratt til lavere gjennomsnittsskår siden referanseverdiene (Constant, 1986) er vist å være lavere ved økende alder. Imidlertid hadde andre sammenlignbare studier (Fraser et al., 2020; Gracitelli et al., 2016) tilsvarende og høyere gjennomsnittsalder. Resultatene fra Zhu et al. (2011) er sammenlignbare med en studie (Plath et al., 2019) som viste kjønns- og aldersjustert CMS gjennomsnittsskår, men det fremkommer ikke i Zhu et al. (2011) at resultatene er kjønns- og aldersjustert.

Noe bedre gjennomsnittsskår i Fraser et al. (2020) kan skyldes operasjonsmetoden da resultatene fra studien viste signifikant forskjell i funksjon målt med CMS mellom gruppene plating og reversert protese (63 vs. 54). Imidlertid beskrev Handoll et al. (2022) lav evidens for at reversert protese gir bedre skulderfunksjon i en Cochrane-oversikt og de andre sammenlignbare studiene med samme operasjonsmetode som i masteroppgaven (Gracitelli et al., 2016; Zhu et al., 2011), rapporterte høyere gjennomsnittlig CMS skår sammenlignet med Fraser et al. (2020). Plath et al. (2019) rapporterte median CMS skår på 57 seks måneder

postoperativt, hvor samme operasjonsmetoder var benyttet, men det er ikke sammenlignbart med masteroppgavens gjennomsnittsskår.

Forskjellen i gjennomsnittsskår kan trolig ikke tilskrives måletidspunkt, da Zhu et al. (2011) ikke fant forskjell av betydning mellom ett og tre år, og Rangan et al. (2015) har rapportert størst funksjonsbedringen hos forskjøvet PHF de første seks månedene postoperativt målt med OSS.

Masteroppgavens gjennomsnittsskår kan også være påvirket av postoperative komplikasjoner i utvalget, da det er vist betydelig dårligere funksjon målt med CMS hos subgrupper med komplikasjoner sammenlignet med subgrupper uten komplikasjoner (50 vs. 71) (Ye et al., 2013). Det er ikke tilgjengelig data for andel komplikasjoner i masteroppgaven, men studier med samme operasjonsmetoder har rapportert 11-33% komplikasjoner (Gracitelli et al., 2016; Plath et al., 2019; Zhu et al., 2011), så er det sannsynlig at masteroppgavens utvalg også har en betydelig andel komplikasjoner. Fraser et al. (2020) fant høyere andel komplikasjoner hos tre og fire-parts operert med plating (20%), sammenlignet med reversert skulderprotese (11%) noe som underbygger antagelsen om betydelig andel komplikasjoner i masteroppgavens utvalg.

Andelen med det som ansees som dårlig CMS skår (under 56) var 48% i utvalget, andelen med det som ansees som moderat CMS skår (56-70) var 43% og ingen oppnådde det som ansees som svært god CMS skår (over 86) (Tabell 9). Få sammenlignbare studier har rapportert kategorisert CMS skår, men en relativt stor andel med det som ansees som dårlig CMS skår samsvarer med disse. Geiger et al. (2010) fant at 39% hadde dårlig CMS skår gjennomsnittlig 25 måneder postoperativt og Ye et al. (2013) rapporterte at 60% av utvalget hadde moderat til dårlig funksjon (under 70) målt med CMS ett år postoperativt hos tre og fire-parts PHF operert med plating. Resultatene i masteroppgaven viste også at kun 56% av de med lav CMS skår (under 56) oppfylte kriteriene til å utføre styrketesten, til sammenligning var andelen 100% hos de som hadde CMS skår 56 eller høyere (Figur 15). Forskjellen var signifikant og resultatene kan indikere at styrketestens kriterier (smertefrihet og elevasjon av arm over skulderhøyde) og tap av poeng (maksimalt 25) ved å ikke utføre testen er svært utslagsgivende på CMS summskår hos pasienter med tre eller fire-parts PHF operert med nagle eller plating seks måneder postoperativt.

Oppsummert fremkommer det ulike resultater i sammenlignbare studier, noe som trolig skyldes at både kjønn, alder, operasjonsmetode, måletidspunkt og eventuelt manglende

standardisering av CROMs delen påvirker resultatene på CMS. Heterogene utvalg, ulike måletidspunkter og operasjonsmetoder gjør det vanskelig å sammenligne resultatene. Den store variasjonen i resultater på tvers av sammenlignbare studier kan muligens tilskrives måleinstrumentets usikre psykometriske egenskaper for pasienter med PHF. I litteraturen har usikre psykometriske egenskaper ført til at det frarådes bruk av CMS som primærutfallsmål i studier som omhandler PHF (Lillkrona, 2008) og i systematiske oversikter er det ingen som anbefaler bruk av CMS for å måle effekt av behandling hos pasienter med PHF (Nowak et al., 2019; Richard et al., 2020; van de Water et al., 2011; Vrotsou et al., 2018). Resultatene fra masteroppgaven og sammenlignbare studier kan indikere at CMS er mindre egnet til å evaluere funksjon hos pasienter med tre og fire-parts PHF operert med plating eller nagle seks måneder postoperativt.

### **5.2.1.2 Postoperativ funksjon målt med DASH**

Resultatene fra masteroppgaven viste at DASH gjennomsnittsskår var 17,8, og gjennomsnittsskår blant kvinner var lavere enn blant menn (19,2 vs. 10,9,  $p=0,05$ ). Resultatene er sammenlignbare med lignende studier som fant at gjennomsnittlig DASH skår på 19 seks måneder postoperativt hos to og tre-parts PHF (Gracitelli et al., 2016) og gjennomsnittlig DASH skår 20 ett år postoperativt hos tre og fire-parts PHF (Ye et al., 2013). Imidlertid fant Olerud et al. (2011) gjennomsnittlig DASH skår 32 ett år postoperativt hos PHF operert med halvprotese. Forskjellen fra resultatene i masteroppgaven, kan forklares med at kun fire-parts PHF ble inkludert i Olerud et al. (2011) og utvalget hadde større andel kvinner (85%), da det er vist at kvinner og fire-parts PHF har dårligere skulderfunksjon sammenlignet med menn og andre bruddtyper målt med CMS (Clement, Duckworth, et al., 2014). Normative gjennomsnittsverdier (Tabell 2) viser også lavere forventet DASH skår ved høyere alder og hos kvinner sammenlignet med menn (Aasheim & Finsen, 2014).

Operasjonsmetode benyttet kan også være av betydning for skulderfunksjon, men det er veldig lav eller ingen evidens for dårligere DASH skår ved halvprotese sammenlignet med plating eller nagling (Handoll et al., 2022). Plath et al. (2019) rapporterte median 48 hos PHF operert med plating eller nagle over 60 år, hvor majoriteten var tre-parts PHF seks måneder postoperativ, men det er ikke sammenlignbart med masteroppgavens gjennomsnittsskår.

Ytterligere deskriptive analyser i masteroppgaven viste at 77% av utvalget hadde det som ansees som svært god funksjon målt med DASH, 19% hadde god funksjon og 4% hadde

moderat funksjon. Ingen hadde det som ansees som dårlig funksjon. Hele 92% hadde svært god funksjon blant menn og 77% blant kvinner, men det var ikke signifikant forskjell mellom kjønn. Resultatene samsvarer med en lignende studie hvor 80% av utvalget (operert tre- og fire-parts PHF) hadde svært god funksjon målt med DASH (Ye et al., 2013).

Masteroppgavens resultater og sammenlignbare studier kan det indikere at DASH skår ikke påvirkes i stor grad av bruddtype, alder og kjønn hos pasienter med PHF. Nowak et al. (2019) og Padua et al. (2021) mente at DASH var mest egnet for å måle symptomer og begrensinger i ADL hos pasienter med PHF. DASH har vist svært god reliabilitet og moderat validitet hos pasienter med PHF (Nowak et al., 2019). Resultatene fra masteroppgaven, sammenlignbare studier og litteraturen kan indikere at DASH er akseptabel for å evaluere funksjon hos pasienter med tre og fire-parts PHF operert med plating eller nagle seks måneder postoperativt.

### **5.2.1.3 Postoperativ funksjon målt med OSS**

Resultatene i denne masteroppgaven viste at gjennomsnittlig OSS skår var 39, kvinner hadde lavere skår enn menn (38,2 vs. 42,6,  $p=0,05$ ). En sammenlignbar studie fant gjennomsnittlig OSS skår 38 hos gruppen som var operert seks måneder postoperativt (Rangan et al., 2015). Studien hadde tilnærmet lik kjønnsfordeling og gjennomsnittsalder, men inkluderte også to-parts PHF og hadde et større utvalg. En annen sammenlignbar studie av Fraser et al. (2020) fant OSS skår 39 to år postoperativt for hele utvalget (reversert protese og plating) bestående av 3 og 4-parts PHF, hvor det var høy andel kvinner (90%), høy gjennomsnittsalder (75 år) og utvalget var randomisert til plating eller reversert protese. På bakgrunn av sammenlignbare resultater kan det indikere at OSS skår ikke påvirkes i stor grad av bruddtype, alder og kjønn hos pasienter med PHF. Imidlertid viser normative gjennomsnittsverdier (Tabell 3) dårligere funksjon med øktende alder og hos kvinner sammenlignet med menn (Clement & Court-Brown, 2014).

Fordi OSS er et skulderspesifikt måleinstrument, og derav trolig mindre påvirkning av andre eventuelle komorbiditeter, kan dette gi høyere validitet når det gjelder mål av postoperativ funksjon i skulderen (Dawson et al., 2009).



Oppsummert kan resultatene fra masteroppgaven og sammenlignbare studier kan indikere at OSS er egnet til å evaluere funksjon med tre og fire-parts PHF operert med plating eller nagle seks måneder postoperativt.

### **5.2.2 Tak- eller gulveffekter (eller manglende skårer) for de tre måleinstrumentene**

Resultatene fra masteroppgaven viste at 22% av utvalget ikke oppfylte kriteriene for utføre CMS styrketesten og oppnådde derfor null poeng i dette domenet. Det var liten forskjell mellom kvinner og menn (22% vs. 23%). Styrkedomenet av CMS kan derfor indikere gulveffekt, da det ansees å være tilstede hvis mer enn 15% av utvalget oppnår laveste skår (McHorney & Tarlov, 1995). Christie, Hagen, Mowinckel og Dagfinrud (2009) rapporterte gulveffekt på styrkedomenet hos pasienter med revmatiske lidelser, da 52 % ikke kunne utføre testen. Tilsvarende funn er gjort ved kapsulitt i skulder (Othman & Taylor, 2004). En stor andel manglende data og frafall er mulig indikasjoner på høy pasientbyrde eller uegnede målemetoder. Det er få studier som har rapportert resultater fra styrkedomenet alene. Ye et al. (2013) rapporterte gjennomsnittsskår 11 på styrkedomenet hos pasienter med tre og fire-parts PHF ett år etter plating, men oppga imidlertid ikke andel som eventuelt uteble fra styrketesten. Fraser et al. (2020) rapporterte også 11 i gjennomsnittsverdi på styrkedomenet hos opererte tre og fire-parts PHF, men oppga heller ikke andelen som ikke utførte testen.

Det er bemerkelsesverdig at andre sammenlignbare studier ikke hadde manglende data på styrkedomenet (Fraser et al., 2020; Gracitelli et al., 2016; Plath et al., 2019; Ye et al., 2013; Zhu et al., 2011), da andre studier (Christie et al., 2009; Othman & Taylor, 2004) og masteroppgaven hadde en betydelig andel som ikke oppfylte kriteriene. Kun en sammenlignbar studie (Zhu et al., 2011) oppga at retningslinjene til Constant et al. (2008) ble fulgt. Mulige årsaker til at det ikke rapporteres manglende data på styrkedomenet i sammenlignbare studier kan være at retningslinjene til Constant et al. (2008) ikke ble fulgt, og dermed ingen som ikke oppfylte kriteriene eller eventuelt at det ikke er prioritert å beskrive i studiene. Hvis retningslinjene av Constant et al. (2008) for utførelse av styrketesten ikke ble fulgt i sammenlignbare studier påvirker dette reliabiliteten og det kan gi skjevhet i resultatene og påvirke sammenlignbarheten mellom studiene.

Det ble ikke registrert tak- eller gulveffekt på DASH hverken for kvinner eller menn.

Resultatene viste at 11% skåret best mulig OSS skår (48), og for gruppen menn var andelen 23%. Dette gjaldt imidlertid kun tre personer (Tabell 10). Resultatene kan allikevel indikere takeffekt for måleinstrumentet hos menn operert for tre og fire-parts PHF med plating eller nagle seks måneder postoperativt. Goudie, MacDonald og Robinson (2022) fant takeffekt hos 230 personer (31%) med PHF behandlet konservativt ved bruk av OSS ett år etter skadetidspunkt. Hvis takeffekt oppstår er det sannsynlig at ekstreme verdier mangler i den nedre skalaen på måleinstrumentet (Terwee et al., 2007). Dette kan indikere dårlig innholdsvaliditet ved instrumentet og i tillegg vil ikke utvalget med høyest skår kunne skilles fra hverandre og dermed vil reliabiliteten på instrumentet reduseres. Responsiviteten er også redusert fordi endring ikke kan måles hos gruppen med maksimal skår (Terwee et al., 2007).

### **5.2.3 Sammenheng mellom data fra måleinstrumentene**

Resultatene i masteroppgaven viste sterk signifikant korrelasjon mellom måleinstrumentene og sterkest korrelasjon mellom PROMs (Tabell 11). Resultatene samsvarer med lignende studier (Beastall et al., 2012; Maurer et al., 2021; van de Water et al., 2014; Ye et al., 2013) (Tabell 4). Sterk korrelasjon mellom måleinstrumentene i mastoppgavens resultater og sammenlignbare studier, spesielt mellom PROMs, kan indikere at instrumentene måler det samme og har gode psykometriske egenskaper (konstruktvaliditet) hos den gitte pasientgruppen (Choi, Schalet, Cook & Cella, 2014; van de Water et al., 2014). I et forskerperspektiv kan høy sammenheng mellom to eller flere uavhengige variabler (multikollinearitet) hindre at analysene blir pålitelige (Daoud, 2017) og i slike tilfeller må eventuelt uavhengige variabler slås sammen eller en eller flere elimineres. Fra et etisk perspektiv skal all klinisk forskning beskytte mennesker, og det omhandler også å minimere belastningen på pasientene når det gjelder bruk av målemetoder (Lingler, Schmidt, Gentry, Hu & Terhorst, 2014), og bruk av mange målemetoder kan føre til stor pasientbyrde, som også kan påvirke kvaliteten på datainnsamlingen (U.S. Department of Health, 2006; Yan, Fricker & Tsai, 2020, s. 194). CMS er vurdert til å ha uakseptabel kliniker og pasientbyrde (Nowak et al., 2019), og fra et klinikerperspektiv kan bruk av multiple måleinstrumenter med høy sammenheng anses som overflødig, og bør derfor ikke brukes samtidig av hensyn til kliniker og pasientbyrden (Baumgarten, Barthman & Chang, 2021). På en annen side har en systematisk oversikt av Richard et al. (2020) beskrevet sammenheng mellom høyere evidensnivå ved bruk av høyere antall måleinstrumenter hos pasienter med PHF, og derfor

foreslått å bruke minst tre måleinstrumenter i fremtidige studier. På bakgrunn av populasjonens funksjonsnivå anbefalte Maurer et al. (2021) bruk av DASH og OSS, fordi PROMs kan gi mindre pasientbyrde for pasienter med PHF hvis de kan besvares hjemme.

Resultatene i masteroppgaven viser noe svakere sammenheng mellom PROMs (OSS og DASH) og CROMs (CMS), sammenlignet med PROMs (DASH) og PROMs (OSS), samme resultater vises også i de sammenlignbare studiene (Maurer et al., 2021; van de Water et al., 2014). Ye et al. (2013) fant moderat sammenheng mellom styrkedomenet ved CMS og DASH ( $r_s = -0,56$ ). Dette kan skyldes at CROMs og PROMs måler ulike dimensjoner av fenomenet (CMS måler også ROM og styrke), eller at CMS har dårligere psykometriske egenskaper for pasienter operert for PHF (Harreld et al., 2013).

Til tross for sterk signifikant sammenheng mellom data fra måleinstrumentene i masteroppgaven hadde 48% det som ansees som dårlig funksjon målt med CMS (under 56). Til sammenligning hadde ingen i utvalget det som ansees som dårlig funksjon målt med DASH, imidlertid hadde 77% det som ansees som svært god skulderfunksjon målt med DASH. Til tross for at det ikke er kjent kategorisering av OSS, indikerer gjennomsnittsskår 39 og mulig takeffekt generelt god skulderfunksjon i utvalget målt med OSS. Lav enighet i resultatene samsvarer med andre studier som har benyttet samme måleinstrumenter hos pasienter med PHF (Beastall et al., 2012; Südkamp et al., 2011; Ye et al., 2013). Ye et al. (2013) fant at 60% av utvalget hadde moderat til dårlig funksjon målt med CMS ett år postoperativt og 80% av utvalget hadde det som ansees som svært god DASH skår, og derfor vurdert til lav enighet mellom måleinstrumentene hos tre og fire-parts PHF operert med plating. Beastall et al. (2012) fant sterk sammenheng mellom CMS og OSS hos opererte PHF, men samtidig viste gjennomsnittsskår for OSS mye bedre skulderfunksjon enn gjennomsnittsskår for CMS. Baker, Nanda, Goodchild, Finn og Rangan (2008) rapporterte også lav enighet mellom CMS og OSS til tross for sterk sammenheng hos PHF konservativt behandlet.

Resultatene kan tyde på at de objektive domene i CMS påvirker sumskåren i stor grad. En annen forklaring kan være at de objektive målene ikke samsvarer med pasientens egen opplevelse eller måler det som har betydning for pasienten, eller på en annen side at pasienter overvurderer egne evner til å utføre oppgaver i PROMs (Kruger & Dunning, 1999).

Resultatene i masteroppgaven er fra samme utvalg og samme måletidspunkt, men vurdering av behandlingseffekt kan tolkes svært ulikt basert på resultatene og ulike konklusjoner kan bli gitt ved bruk av måleinstrumentene separat i studier. I følge Ye et al. (2013) eksisterer

uenigheten mellom PROMs DASH og delvis CROMs CMS trolig fordi CMS er utsatt for måleskjevheter og kanskje ikke representerer pasientens perspektiv tilstrekkelig. Richard et al. (2020) og Padua et al. (2021) kritiserte den utbredte bruken av CMS, på bakgrunn av manglende psykometriske egenskaper og redusert pasientperspektiv. I en kvalitativ studie med semistrukturerte intervjuer av O'Hara et al. (2017) rapporterte deltakerne begrensninger i dagligdagse- og fritidsaktiviteter, smerte og søvnforstyrrelser etter PHF og PROMs anbefales når pasienten er eneste tilgjengelig kilde til informasjonen eller når CROMs ikke samsvarer med pasientens egen opplevelse (Copay et al., 2007). Det siste tiåret har man sett et paradigmeskifte når det gjelder valg av måleinstrumenter, med økende fokus på pasientens perspektiv og bruk av PROMs (Hamilton, Giesinger & Giesinger, 2017) og det gir lavere kostnad og mindre byrde for alle berørte å gjennomføre sammenlignet med CROMs (Dawson et al., 2009). Unødvendig fysisk, følelsesmessig eller kognitiv belastning vil generelt redusere kvantiteten og kvaliteten på datamaterialet (U.S. Department of Health, 2006). Å identifisere hvordan pasienter oppfatter og rapporterer symptomer og funksjonsbegrensninger best mulig er svært viktig for å kunne avgjøre om en behandling har hatt effekt (Harreld et al., 2013).

På bakgrunn av mangelfull psykometrisk data om de ulike måleinstrumentene for pasienter med PHF, er det vanskelig å vurdere hvilket som er best egnet (van de Water et al., 2011). Når det gjelder CROMs, blant annet CMS, hadde ingen av måleinstrumentene akseptable psykometriske egenskaper ifølge Nowak et al. (2019). Når det gjaldt PROMs, fant forfatterne akseptable psykometriske egenskaper ved DASH og OSS. Det er vist at PHF er knyttet til begrensninger i psykososiale forhold og dagligdagse aktiviteter, og det anbefales derfor bruk av PROMs for å fange pasientens perspektiv (Jayakumar et al., 2018). Manglende justering for alder og kjønn ved CMS, gjør instrumentet mindre egnet til å måle heterogene utvalg som PHF (Maurer et al., 2021). I lys av ICF-modellen måler DASH (77%) og OSS (75%) i mye større grad aktivitet- og deltakelses begrensninger, sammenlignet med CMS (20%) (Nowak et al., 2021).

Oppsummert kan resultatene fra masteroppgaven støttet av andre studier indikere at CMS og OSS er mindre egnede måleinstrumenter for pasienter med tre og fire-parts PHF operert med plating eller nagle sammenlignet med DASH.

#### 5.2.4 Sammenheng mellom lav CMS skår og uavhengige bakgrunnsvariabler

Regresjonsanalysen viste signifikant høyere odds ratio for lav CMS skår ved ASA skår 3 sammenlignet med ASA 1 og 2 ujustert og justert for konfunderende variabler (Tabell 13). Konfidensintervallet var imidlertid svært stort (1,40-80,90) og forutsetningene for analysene var ikke oppfylt på grunn av to observasjoner med høy Cook's D verdi som påvirket estimatene i betydelig grad og derfor er det stor usikkerhet i estimatene.

Når det gjelder de andre uavhengige variablene viste det å være pensjonist eller ufør høyere odds ratio for lav CMS skår i justerte analyser, men resultatene var ikke signifikante. Utvalgets andel i arbeid bestod av 69% menn og utvalgets andel pensjonist/ufør bestod av 75% kvinner, så resultatene kan ha sammenheng med kjønn. Det var uventet at de uavhengige variablene kjønn (kvinner) og økende alder ikke viste signifikant høyere odds ratio for lav CMS skår, da det er kjent at absolutt CMS skår ikke justerer tilstrekkelig for kjønns- og aldersforskjeller (Katolik et al., 2005; Yian et al., 2005), normative gjennomsnittsverdier er lavere for kvinner sammenlignet med menn og ved økende alder (Yian et al., 2005). I tillegg er kjønn og alder indentifiserte risikofaktorer for dårlig skulderfunksjon etter PHF (Fisher et al., 2022). Imidlertid fant en sammenlignbar studie av Miquel et al. (2021) som undersøkte 126 personer med en-fire parts PHF, hvor de fleste var behandlet konservativt, signifikant negativ sammenheng mellom lav CMS skår og økende alder. En annen sammenlignbar studie av Clement, Duckworth, et al. (2014) undersøkte 637 personer med en-fire parts PHF, hvor majoriteten var behandlet konservativt, fant også signifikant negativ sammenheng mellom økende alder og lav CMS skår etter ett år. Südkamp et al. (2011) fant sammenheng med lav CMS skår for kvinner og personer over 80 år hos utvalg bestående av 463 personer operert for tre og fire-parts PHF. Men andre sammenlignbare studier fant heller ikke sammenheng mellom kvinner og lav CMS skår etter ett år (Clement, Duckworth, et al., 2014; Miquel et al., 2021).

De sammenlignbare studiene har en betydelig høyere utvalgsstørrelse sammenlignet med masteroppgaven, og det kan forklare hvorfor signifikante resultater i sammenlignbare studier er oppnådd (Dabija & Jain, 2019). Statistisk styrke kan være lav ved analyser av små utvalg, eller analyser av mindre subgrupper i utvalget, og kan resultere i manglende sammenheng mellom avhengig og uavhengig variabel til tross for at det faktisk er en sammenheng (Polit & Beck, 2017, s. 221). På en annen side er ikke utvalgene i de sammenlignbare studiene like

masteroppgavens utvalg når det gjelder kjønns sammensetning, bruddtyper, gjennomsnittsalder, behandlingsmetode eller måletidspunkt.

Til tross for stor usikkerhet i estimatene for høyere odds ratio for lav CMS skår ved ASA 3, er økende grad av komorbiditet identifisert risikofaktor for dårligere postoperativ skulderfunksjon målt med DASH (Fisher et al., 2022). Tilsvarende funn er gjort for høyere BMI (Werner, Griffin, Yang, Brockmeier & Gwathmey, 2015) og alder over 65 år (Shulman et al., 2013). Imidlertid fant ikke Miquel et al. (2021) sammenheng mellom høyere BMI og lav CMS skår. Goudie et al. (2022) prospektive kohort bestående av 774 PHF behandlet konservativt rapporterte at komorbiditet og psykososiale faktorer som grad av selvstendighet, grad av sosial deltakelse og affektive lidelser var de sterkeste uavhengige prediktorene på skulderfunksjon ett år etter skadetidspunkt målt med OSS. Andre studier har også rapportert psykososiale faktorer og grad av sosial deltakelse som sterkeste uavhengige prediktorene på skulderfunksjon etter PHF (Clement, McQueen & Court-Brown, 2014; Jayakumar et al., 2019), mens andre har rapportert manglende bruddtilheling som årsak til redusert skulderfunksjon hos konservativt og operativt behandlet PHF (Court-Brown & McQueen, 2008) eller grad av bruddforskyvning hos konservativt behandlede PHF (Foruria, De Gracia, Larson, Munuera & Sanchez-Sotelo, 2011).

### **5.3 Kliniske implikasjoner og videre forskning**

Resultatene fra masteroppgaven har bidratt til økt kunnskap om hvilken skulderfunksjon pasienter med tre og fire-parts PHF operert med plating eller nagle har oppnådd seks måneder postoperativt, samt forskjell mellom ulike subgrupper. Basert på resultatene har majoriteten pasientene av pasientene oppnådd god skulderfunksjon målt med DASH og OSS, og nesten halvparten dårlig funksjon målt med CMS. Resultatene viser at det er viktig å være oppmerksom på hvilket måleinstrument som er benyttet når resultater fra studier som omhandler opererte tre og fire-parts PHF tolkes, da samme utvalgs skulderfunksjon kan tolkes svært ulikt avhengig av måleinstrumentet. På bakgrunn av resultatene kan man anta at lavere skår ved CMS ikke nødvendigvis betyr dårlig effekt av behandling, ettersom flertallet i tilsvarende utvalg skåret god funksjon på DASH og OSS. Resultatene fra denne masteroppgaven vil derfor også kunne bidra med viktig informasjon om hvordan resultatene fra hovedstudien kan tolkes og forstås.

I lys av sterk sammenheng mellom måleinstrumenter i masteroppgaven, potensiell pasientbyrde og helsevesenets fremtidige kapasitetsutfordringer bør kostnadseffektive valg gjøres (Harreld et al., 2013). Det kan derfor ansees om unødvendig å bruke måleinstrumentene samtidig i klinisk praksis eller forskning. Det er ønskelig å gjøre forskningsdeltakelse mer praktisk og pasientsentrert, og nøye vurdering av antall måleinstrumenter og antall måletidspunkter kan bidra til dette (Bodart et al., 2019).

Resultatene kan bidra til kunnskap om valg av målemetode for evaluering av opererte tre og fire-parts PHF i klinisk praksis eller i forskning, fordi resultatene kan indikere at DASH er bedre egnet måleinstrument enn OSS og CMS, da resultatene indikerer at OSS har takeffekt hos menn og styrkedomenet av CMS har høy andel manglende data. Resultatene støttes av systematiske oversikter som har undersøkt målemetoder for alle typer PHF da forfatterne mente at DASH muligens var det mest egnede instrumentet for å måle symptomer og begrensinger i ADL (Nowak et al., 2019; Padua et al., 2021). Valg av måleinstrumenter bør baseres på psykometriske egenskaper for den tiltenkte pasientgruppen, bør være lett tilgjengelig og gi minst byrde for pasient og kliniker (Richard et al., 2020).

Når det gjelder hvilke bakgrunnsvariabler som er assosiert med lav CMS skår, kan ikke masteroppgavens resultater gi ny kunnskap på grunn av lav statistisk validitet. Det er diskrepans mellom sammenlignbare studiers resultater og andre studier har undersøkt sammenheng med andre bakgrunnsvariabler, så det er ikke sikker kunnskap om hvilke bakgrunnsvariabler som har sammenheng med lav CMS skår hos pasienter operert med plating eller nagle hos tre og fire-parts PHF. For å kunne oppnå slik kunnskap bør det i fremtiden utføres større tverrsnittstudier/kohortstudier som inkluderer kun opererte tre og fire-parts PHF, og alle relevante forklaringsvariabler må samles inn.

Masteroppgaven belyser også usikker og manglende kunnskap om psykometriske egenskaper ved måleinstrumenter benyttet for evaluering av PHF. Manglende gullstandard for måleinstrumenter, og dermed stor variasjon i bruk, har skapt utfordringer når de gjelder kunnskap om behandling av PHF og det er etterlyst økt kunnskap om eksisterende målemetoder (Richard et al., 2020). Mangelfull psykometrisk data om de ulike måleinstrumentene indikerer at videre forskning bør fokusere på bedre kunnskapen om måleinstrumentenes validitet, reliabilitet, responsivitet og pasientbyrde benyttet for evaluering av opererte tre og fire-parts PHF.

Det er foreslått å etablere et kjernesett med måleinstrumenter for pasientgruppen for å bedre sammenlignbarheten mellom resultater i enkeltstudier (Jayakumar et al., 2018; Nowak et al., 2019). Patient-Reported Outcome Measurement Information System (PROMIS) er et initiativ som utvikler og validerer PROMs for å bedre måleegenskapene, bedre sammenlignbarheten mellom studier og redusere pasient og klinikerbyrden (Cella et al., 2010). Morgan, Kallen, Okike, Lee og Vrahas (2015) fant at PROMIS Physical Function Computer Adaptive test kan være egnet for evaluering av pasienter med PHF.



## 6.0 Konklusjon

Denne masteroppgaven har undersøkt skulderfunksjon med ulike måleinstrumenter i et tverrsnittsdesign blant 73 pasienter med tre og fire-parts PHF som er operert med plating eller nagle. Resultatene viser at disse pasientene gjennomsnittlig har dårlig skulderfunksjon målt med CMS, og gjennomsnittlig god skulderfunksjon målt med DASH og OSS seks måneder postoperativt. Kvinner har dårligere skulderfunksjon sammenlignet med menn uavhengig av måleinstrument, men det var kun for CMS at det var en statistisk signifikant forskjell.

Resultatene indikerer en takeffekt ved bruk av OSS for menn og viser en stor andel manglende data på styrkedomenet ved CMS, noe som kan indikere at disse to instrumentene er mindre egnet enn DASH til å evaluere funksjon for pasienter med tre og fire-parts PHF seks måneder postoperativt.

Det fremkommer av masteroppgavens resultater at det var en sterk og signifikant sammenheng mellom måleinstrumentene, men lav enighet mellom PROMs og CROMs.

Resultatene viste signifikant høyere odds ratio for lav CMS skår hos personer med ASA 3, sammenlignet med ASA 1 og 2 i justerte og ujusterte analyser, men disse resultatene må tolkes med forsiktighet på grunn av metodologiske svakheter i de statistiske analysene.

## 7.0 Referanseliste

- Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc*, 9, 211-217. <https://doi.org/10.2147/jmdh.S104807>
- Bahrs, C., Badke, A., Rolauffs, B., Weise, K., Zipplies, S., Dietz, K. & Eingartner, C. (2010). Long-term results after non-plate head-preserving fixation of proximal humeral fractures. *International Orthopaedics*, 34(6), 883-889. <https://doi.org/10.1007/s00264-009-0848-4>
- Bahrs, C., Bauer, M., Blumenstock, G., Eingartner, C., Bahrs, S. D., Tepass, A., ... Rolauffs, B. (2013). The complexity of proximal humeral fractures is age and gender specific. *Journal of Orthopaedic Science*, 18(3), 465-470. <https://doi.org/10.1007/s00776-013-0361-x>
- Bahrs, C., Stojicevic, T., Blumenstock, G., Brorson, S., Badke, A., Stöckle, U., ... Freude, T. (2014). Trends in epidemiology and patho-anatomical pattern of proximal humeral fractures. *International Orthopaedics*, 38(8), 1697-1704. <https://doi.org/10.1007/s00264-014-2362-6>
- Baker, P., Nanda, R., Goodchild, L., Finn, P. & Rangan, A. (2008). A comparison of the Constant and Oxford Shoulder Scores in patients with conservatively treated proximal humeral fractures. *Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ... [et al.]*, 17, 37-41. <https://doi.org/10.1016/j.jse.2007.04.019>
- Bakhsh, W. & Nicandri, G. (2018). Anatomy and Physical Examination of the Shoulder. *Sports Med Arthrosc Rev*, 26(3), e10-e22. <https://doi.org/10.1097/jsa.0000000000000202>
- Bassett, R. (2023). Proximal humeral fractures in adults. I M. Gammons & C. Asplund (Red.), *UpToDate*, . UpToDate (lest 3.10.2023):
- Baumgarten, K. M., Barthman, B. J. & Chang, P. S. (2021). The American Shoulder and Elbow Score Is Highly Correlated With the Western Ontario Rotator Cuff Index and Has Less Responder and Administrator Burden. *Arthrosc Sports Med Rehabil*, 3(6), e1637-e1643. <https://doi.org/10.1016/j.asmr.2021.07.019>
- Beastall, J. E., Fielding, S., Christie, E. & Johnstone, A. J. (2012). Shoulder outcome measures: is there a right answer? *European Journal of Trauma and Emergency Surgery*, 38(6), 659-664. <https://doi.org/10.1007/s00068-012-0220-z>
- Beks, R. B., Ochen, Y., Frima, H., Smeeing, D. P. J., van der Meijden, O., Timmers, T. K., ... Houwert, R. M. (2018). Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. *J Shoulder Elbow Surg*, 27(8), 1526-1534. <https://doi.org/10.1016/j.jse.2018.03.009>
- Bergdahl, C., Ekholm, C., Wennergren, D., Nilsson, F. & Möller, M. (2016). Epidemiology and patho-anatomical pattern of 2,011 humeral fractures: data from the Swedish Fracture Register. *BMC Musculoskeletal Disorders*, 17, 159. <https://doi.org/10.1186/s12891-016-1009-8>

- Bergdahl, C., Wennergren, D., Ekelund, J. & Möller, M. (2020). Mortality after a proximal humeral fracture. *Bone Joint J*, 102-b(11), 1484-1490. <https://doi.org/10.1302/0301-620x.102b11.Bjj-2020-0627.R1>
- Bjorgul, K., Novicoff, W. M. & Saleh, K. J. (2010). American Society of Anesthesiologist Physical Status score may be used as a comorbidity index in hip fracture surgery. *Journal of Arthroplasty*, 25(6 Suppl), 134-137. <https://doi.org/10.1016/j.arth.2010.04.010>
- Bodart, S., Byrom, B., Crescioni, M., Eremenco, S. & Flood, E. (2019). Perceived Burden of Completion of Patient-Reported Outcome Measures in Clinical Trials:: Results of a Preliminary Study. *Ther Innov Regul Sci*, 53(3), 318-323. <https://doi.org/10.1177/2168479018788053>
- Bondevik, H. & Engebretsen, E. (2020). Kunnskapstyper og standarisering. I *Vitenskapsteori for sosial- og helsefag*. Gyldendal Akademisk.
- Carter, R. E. & Lubinsky, J. (2016). *Rehabilitation research : principles and applications* (Fifth edition. utg.). St. Louis, Missouri: Elsevier.
- Castro-Franco, A. D., Mendoza-Muñoz, I., González-Ángeles, Á., Cruz-Sotelo, S. E., Castañeda, A. M. & Siqueiros-Hernández, M. (2020). Trends in the Characterization of the Proximal Humerus in Biomechanical Studies: A Review. *Applied Sciences*, 10(18), 6514. Hentet fra <https://www.mdpi.com/2076-3417/10/18/6514>
- Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., ... Choi, S. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*, 63(11), 1179-1194.
- Choi, S. W., Schalet, B., Cook, K. F. & Cella, D. (2014). Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychological Assessment*, 26(2), 513-527. <https://doi.org/10.1037/a0035768>
- Christie, A. M., Hagen, K. B. P., Mowinckel, P. M. & Dagfinrud, H. P. (2009). Methodological properties of six shoulder disability measures in patients with rheumatic diseases referred for shoulder surgery. *J Shoulder Elbow Surg*, 18(1), 89-95. <https://doi.org/10.1016/j.jse.2008.07.008>
- Chu, S. P., Kelsey, J. L., Keegan, T. H. M., Sternfeld, B., Prill, M., Quesenberry, C. P. & Sidney, S. (2004). Risk Factors for Proximal Humerus Fracture. *American Journal of Epidemiology*, 160(4), 360-367. <https://doi.org/10.1093/aje/kwh224>
- Clement, N. D. & Court-Brown, C. M. (2014). Oxford shoulder score in a normal population. *International Journal of Shoulder Surgery*, 8(1), 10-14. <https://doi.org/10.4103/0973-6042.131849>

- Clement, N. D., Duckworth, A. D., McQueen, M. M. & Court-Brown, C. M. (2014). The outcome of proximal humeral fractures in the elderly. *The Bone & Joint Journal*, 96-B(7), 970-977. <https://doi.org/10.1302/0301-620x.96b7.32894>
- Clement, N. D., McQueen, M. M. & Court-Brown, C. M. (2014). Social deprivation influences the epidemiology and outcome of proximal humeral fractures in adults for a defined urban population of Scotland. *European Journal of Orthopaedic Surgery & Traumatology*, 24, 1039-1046.
- Constant, C. R. (1986). Age related recovery of shoulder function after injury. *Thesis, University College*.
- Constant, C. R., Gerber, C., Emery, R., Sjøbjerg, J., Gohlke, F. & Pascal, B. (2008). A review of the Constant score: Modifications and guidelines for its use. *Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ... [et al.]*, 17, 355-361. <https://doi.org/10.1016/j.jse.2007.06.022>
- Constant, C. R. & Murley, A. H. (1987). A clinical method of functional assessment of the shoulder. *Clinical Orthopaedics and Related Research*, (214), 160-164.
- Copay, A. G., Subach, B. R., Glassman, S. D., Polly, D. W., Jr. & Schuler, T. C. (2007). Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*, 7(5), 541-546. <https://doi.org/10.1016/j.spinee.2007.01.008>
- Cosman, F., de Beur, S. J., LeBoff, M. S., Lewiecki, E. M., Tanner, B., Randall, S. & Lindsay, R. (2014). Clinician's Guide to Prevention and Treatment of Osteoporosis. *Osteoporosis International*, 25(10), 2359-2381. <https://doi.org/10.1007/s00198-014-2794-2>
- Court-Brown, C. M. & Caesar, B. (2006). Epidemiology of adult fractures: A review. *Injury*, 37(8), 691-697. <https://doi.org/10.1016/j.injury.2006.04.130>
- Court-Brown, C. M., Garg, A. & McQueen, M. M. (2001). The epidemiology of proximal humeral fractures. *Acta Orthopaedica Scandinavica*, 72(4), 365-371. <https://doi.org/10.1080/000164701753542023>
- Court-Brown, C. M. & McQueen, M. M. (2008). Nonunions of the proximal humerus: their prevalence and functional outcome. *Journal of Trauma*, 64(6), 1517-1521. <https://doi.org/10.1097/TA.0b013e3181469840>
- Dabija, D. I. & Jain, N. B. (2019). Minimal Clinically Important Difference of Shoulder Outcome Measures and Diagnoses: A Systematic Review. *American Journal of Physical Medicine and Rehabilitation*, 98(8), 671-676. <https://doi.org/10.1097/phm.0000000000001169>
- Dahl, H. A. & Rinvik, E. (2010). *Menneskets funksjonelle anatomi : med hovedvekt på bevegelsesapparatet* (3. utg. utg.). Oslo: Cappelen akademisk.
- Daoud, J. I. (2017). Multicollinearity and regression analysis. *Journal of Physics: Conference Series* (s. 012009): IOP Publishing.

- Dawson, J., Fitzpatrick, R. & Carr, A. (1996). Questionnaire on the perceptions of patients about shoulder surgery. *Journal of Bone and Joint Surgery (British Volume)*, 78(4), 593-600.
- Dawson, J., Rogers, K., Fitzpatrick, R. & Carr, A. (2009). The Oxford shoulder score revisited. *Archives of Orthopaedic and Trauma Surgery*, 129(1), 119-123.  
<https://doi.org/10.1007/s00402-007-0549-7>
- de Vet, H. C., Terwee, C. B., Ostelo, R. W., Beckerman, H., Knol, D. L. & Bouter, L. M. (2006). Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes*, 4, 54.  
<https://doi.org/10.1186/1477-7525-4-54>
- Dickinson, E. R., Adelson, J. L. & Owen, J. (2012). Gender Balance, Representativeness, and Statistical Power in Sexuality Research Using Undergraduate Student Samples. *Archives of Sexual Behavior*, 41(2), 325-327. <https://doi.org/10.1007/s10508-011-9887-1>
- Enger, M., Skjaker, S. A., Melhuus, K., Nordsletten, L., Pripp, A. H., Moosmayer, S. & Brox, J. I. (2018). Shoulder injuries from birth to old age: A 1-year prospective study of 3031 shoulder injuries in an urban population. *Injury*, 49(7), 1324-1329.  
<https://doi.org/10.1016/j.injury.2018.05.013>
- Finsen, V. (2008). [Norwegian version of the DASH questionnaire for examination of the arm shoulders and hand]. *Tidsskrift for den Norske Laegeforening*, 128(9), 1070.
- Fisher, N. D., Driesman, A., Saleh, H., Egol, K. A. & Konda, S. R. (2022). The Proximal Humerus Outcome Score at One Year (POSY) Predicts Which Patients Have Poor Functional Outcomes Following Operative Fixation of Proximal Humerus Fractures. *Curēus (Palo Alto, CA)*, 14(7), e26631-e26631. <https://doi.org/10.7759/cureus.26631>
- Fjalestad, T., Iversen, P., Hole, M., Smedsrud, M. & Madsen, J. E. (2014). Clinical investigation for displaced proximal humeral fractures in the elderly: a randomized study of two surgical treatments: reverse total prosthetic replacement versus angular stable plate Philos (The DELPHI-trial). *BMC Musculoskeletal Disorders*, 15, 323. <https://doi.org/10.1186/1471-2474-15-323>
- Foruria, A. M., De Gracia, M. M., Larson, D. R., Munuera, L. & Sanchez-Sotelo, J. (2011). The pattern of the fracture and displacement of the fragments predict the outcome in proximal humeral fractures. *Journal of Bone and Joint Surgery (British Volume)*, 93(3), 378-386.  
<https://doi.org/10.1302/0301-620X.93B3.25083>
- Fox, M. T., Sidani, S. & Streiner, D. (2007). Using standardized survey items with older adults hospitalized for chronic illness. *Research in Nursing and Health*, 30(4), 468-481.
- Franchignoni, F., Vercelli, S., Giordano, A., Sartorio, F., Bravini, E. & Ferriero, G. (2014). Minimal clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (QuickDASH). *Journal of Orthopaedic and Sports Physical Therapy*, 44(1), 30-39. <https://doi.org/10.2519/jospt.2014.4893>

- Fraser, A. N., Bjørdal, J., Wagle, T. M., Karlberg, A. C., Lien, O. A., Eilertsen, L., ... Fjalestad, T. (2020). Reverse Shoulder Arthroplasty Is Superior to Plate Fixation at 2 Years for Displaced Proximal Humeral Fractures in the Elderly: A Multicenter Randomized Controlled Trial. *JBJS*, *102*(6), 477-485. <https://doi.org/10.2106/jbjs.19.01071>
- Friedman, S. M. M. D. M. P. H. & Mendelson, D. A. M. D. M. S. (2014). Epidemiology of Fragility Fractures. *Clinics in Geriatric Medicine*, *30*(2), 175-181. <https://doi.org/10.1016/j.cger.2014.01.001>
- Geiger, E. V., Maier, M., Kelm, A., Wutzler, S., Seebach, C. & Marzi, I. (2010). Functional outcome and complications following PHILOS plate fixation in proximal humeral fractures. *Acta Orthopaedica et Traumatologica Turcica*, *44*(1), 1-6.
- Goudie, E. B., MacDonald, D. J. & Robinson, C. M. (2022). Functional Outcome After Nonoperative Treatment of a Proximal Humeral Fracture in Adults. *Journal of Bone and Joint Surgery (American Volume)*, *104*(2), 123-138. <https://doi.org/10.2106/jbjs.20.02018>
- Gracitelli, M. E. C. P., Malavolta, E. A. P., Assunção, J. H. M. D., Kojima, K. E. P., dos Reis, P. R. M. D., Silva, J. S. P., ... Hernandez, A. J. P. (2016). Locking intramedullary nails compared with locking plates for two- and three-part proximal humeral surgical neck fractures: a randomized controlled trial. *J Shoulder Elbow Surg*, *25*(5), 695-703. <https://doi.org/10.1016/j.jse.2016.02.003>
- Hamilton, D. F., Giesinger, J. M. & Giesinger, K. (2017). It is merely subjective opinion that patient-reported outcome measures are not objective tools. *Bone & Joint Research*, *6*(12), 665-666. <https://doi.org/10.1302/2046-3758.612.Bjr-2017-0347>
- Handoll, H. H. G., Elliott, J., Thillemann, T. M., Aluko, P. & Brorson, S. (2022). Interventions for treating proximal humeral fractures in adults. *Cochrane Database of Systematic Reviews*, (6). <https://doi.org/10.1002/14651858.CD000434.pub5>
- Harreld, K., Clark, R., Downes, K., Virani, N. & Frankle, M. (2013). Correlation of subjective and objective measures before and after shoulder arthroplasty. *Orthopedics*, *36*(6), 808-814. <https://doi.org/10.3928/01477447-20130523-29>
- Helsedirektoratet. (2018). *Tobacco Control in Norway*. Oslo: Helsedirektoratet.
- Hinkle, D. E., Jurs, S. G. & Wiersma, W. (2003). *Applied statistics for the behavioral sciences* (5th. utg.). Boston: Houghton Mifflin.
- Hodgson, S. A., Mawson, S. J., Saxton, J. M. & Stanley, D. (2007). Rehabilitation of two-part fractures of the neck of the humerus (two-year follow-up). *J Shoulder Elbow Surg*, *16*(2), 143-145. <https://doi.org/10.1016/j.jse.2006.06.003>
- Hox, J. J. & Boeije, H. R. (2005). Data collection, primary versus secondary.
- Hudak, P. L., Amadio, P. C. & Bombardier, C. (1996). Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. *The Upper*

- Extremity Collaborative Group (UECG). *American Journal of Industrial Medicine*, 29(6), 602-608. [https://doi.org/10.1002/\(sici\)1097-0274\(199606\)29:6<602::Aid-ajim4>3.0.Co;2-1](https://doi.org/10.1002/(sici)1097-0274(199606)29:6<602::Aid-ajim4>3.0.Co;2-1)
- Iglesias-Rodríguez, S., Domínguez-Prado, D. M., García-Reza, A., Fernández-Fernández, D., Pérez-Alfonso, E., García-Piñeiro, J. & Castro-Menéndez, M. (2021). Epidemiology of proximal humerus fractures. *Journal of Orthopaedic Surgery and Research*, 16(1), 402. <https://doi.org/10.1186/s13018-021-02551-x>
- Jayakumar, P., Overbeek, C. L., Lamb, S., Williams, M., Funes, C. J., Gwilym, S., ... Vranceanu, A. M. (2018). What Factors Are Associated With Disability After Upper Extremity Injuries? A Systematic Review. *Clinical Orthopaedics and Related Research*, 476(11), 2190-2215. <https://doi.org/10.1097/corr.0000000000000427>
- Jayakumar, P., Teunis, T., Williams, M., Lamb, S. E., Ring, D. & Gwilym, S. (2019). Factors associated with the magnitude of limitations during recovery from a fracture of the proximal humerus: predictors of limitations after proximal humerus fracture. *Bone Joint J*, 101-b(6), 715-723. <https://doi.org/10.1302/0301-620x.101b6.Bjj-2018-0857.R1>
- Johansson, K. M. & Adolphson, L. E. (2005). Intraobserver and interobserver reliability for the strength test in the Constant-Murley shoulder assessment. *J Shoulder Elbow Surg*, 14(3), 273-278. <https://doi.org/10.1016/j.jse.2004.08.001>
- Johnson, T. P. & Wislar, J. S. (2012). Response rates and nonresponse errors in surveys. *JAMA*, 307(17), 1805-1806.
- Justesen, L. (2012). Theoretical perspectives and qualitative methodology. I N. Mik-Meyer (Red.), *Qualitative research methods in organisation studies* (s. 11-34). København: Hans Reitzels Forl.
- Kanis, J. A., Norton, N., Harvey, N. C., Jacobson, T., Johansson, H., Lorentzon, M., ... Borgström, F. (2021). SCOPE 2021: a new scorecard for osteoporosis in Europe. *Arch Osteoporos*, 16(1), 82. <https://doi.org/10.1007/s11657-020-00871-9>
- Kanis, J. A., Oden, A., Johnell, O., Jonsson, B., de Laet, C. & Dawson, A. (2001). The burden of osteoporotic fractures: a method for setting intervention thresholds. *Osteoporosis International*, 12(5), 417-427. <https://doi.org/10.1007/s001980170112>
- Katolik, L. I., Romeo, A. A., Cole, B. J., Verma, N. N., Hayden, J. K. & Bach, B. R. (2005). Normalization of the Constant score. *Journal of Shoulder and Elbow Surgery*, 14(3), 279-285. <https://doi.org/https://doi.org/10.1016/j.jse.2004.10.009>
- Kim, D. M., Park, D., Kim, H., Lee, E. S., Shin, M. J., Jeon, I. H. & Koh, K. H. (2020). Risk Factors for Severe Proximal Humerus Fracture and Correlation Between Deltoid Tuberosity Index and Bone Mineral Density. *Geriatr Orthop Surg Rehabil*, 11, 2151459320938571. <https://doi.org/10.1177/2151459320938571>
- Knäuper, B., Carrière, K., Chamandy, M., Xu, Z., Schwarz, N. & Rosen, N. O. (2016). How aging affects self-reports. *Eur J Ageing*, 13(2), 185-193. <https://doi.org/10.1007/s10433-016-0369-0>

- Kramer, J. M. P. O. T. R. L. & Schwartz, A. M. S. O. T. R. L. (2017). Reducing Barriers to Patient-Reported Outcome Measures for People With Cognitive Impairments. *Archives of Physical Medicine and Rehabilitation*, 98(8), 1705-1715. <https://doi.org/10.1016/j.apmr.2017.03.011>
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. <https://doi.org/10.1037//0022-3514.77.6.1121>
- Kruithof, R. N., Formijne Jonkers, H. A., van der Ven, D. J. C., van Olden, G. D. J. & Timmers, T. K. (2017). Functional and quality of life outcome after non-operatively managed proximal humeral fractures. *Journal of Orthopaedics and Traumatology*, 18(4), 423-430. <https://doi.org/10.1007/s10195-017-0468-5>
- Kukkonen, J. M. D., Kauko, T. B., Vahlberg, T. B., Joukainen, A. M. D. P. & Äärimaa, V. M. D. P. (2013). Investigating minimal clinically important difference for Constant score in patients undergoing rotator cuff surgery. *J Shoulder Elbow Surg*, 22(12), 1650-1655. <https://doi.org/10.1016/j.jse.2013.05.002>
- Kyte, D. G., Calvert, M., van der Wees, P. J., ten Hove, R., Tolan, S. & Hill, J. C. (2015). An introduction to patient-reported outcome measures (PROMs) in physiotherapy. *Physiotherapy*, 101(2), 119-125. <https://doi.org/https://doi.org/10.1016/j.physio.2014.11.003>
- Launonen, A. P., Lepola, V., Saranko, A., Flinkkilä, T., Laitinen, M. & Mattila, V. M. (2015). Epidemiology of proximal humerus fractures. *Arch Osteoporos*, 10, 209. <https://doi.org/10.1007/s11657-015-0209-4>
- Lillkrona, U. (2008). How should we use the Constant Score?--A commentary. *J Shoulder Elbow Surg*, 17(2), 362-363. <https://doi.org/10.1016/j.jse.2007.06.013>
- Lingler, J. H., Schmidt, K. L., Gentry, A. L., Hu, L. & Terhorst, L. A. (2014). A New Measure of Research Participant Burden: Brief Report. *J Empir Res Hum Res Ethics*, 9(4), 46-49. <https://doi.org/10.1177/1556264614545037>
- Marongiu, G., Leinardi, L., Congia, S., Frigau, L., Mola, F. & Capone, A. (2020). Reliability and reproducibility of the new AO/OTA 2018 classification system for proximal humeral fractures: a comparison of three different classification systems. *Journal of Orthopaedics and Traumatology*, 21(1), 4. <https://doi.org/10.1186/s10195-020-0543-1>
- Matsumura, N., Furuhashi, R., Seto, T., Takada, Y., Shirasawa, H., Oki, S., ... Shiono, S. (2020). Reproducibility of the modified Neer classification defining displacement with respect to the humeral head fragment for proximal humeral fractures. *Journal of Orthopaedic Surgery and Research*, 15, 1-7.
- Mattyasovszky, S. G., Burkhart, K. J., Ahlers, C., Proschek, D., Dietz, S. O., Becker, I., ... Rommens, P. M. (2011). Isolated fractures of the greater tuberosity of the proximal humerus: a long-term retrospective study of 30 patients. *Acta Orthopaedica*, 82(6), 714-720. <https://doi.org/10.3109/17453674.2011.618912>



- Maurer, E., Bahrs, C., Kühle, L., Ziegler, P., Gonser, C. & Stollhof, L. E. (2021). Comparability of Patient-reported Outcome Measures and Clinical Assessment Tools for Shoulder Function in Patients with Proximal Humeral Fracture. *Zeitschrift für Orthopädie und Unfallchirurgie*, 159(06), 638-648.
- McHorney, C. A. & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of Life Research*, 4(4), 293-307. <https://doi.org/10.1007/bf01593882>
- McKown, S., Acquadro, C., Anfray, C., Arnold, B., Eremenco, S., Giroudet, C., ... Weiss, D. (2020). Good practices for the translation, cultural adaptation, and linguistic validation of clinician-reported outcome, observer-reported outcome, and performance outcome measures. *Journal of Patient-Reported Outcomes*, 4(1), 89. <https://doi.org/10.1186/s41687-020-00248-z>
- Miquel, J., Elisa, C., Fernando, S., Alba, R. & Torrens, C. (2021). Non-medical patient-related factor influence in proximal humeral fracture outcomes: a multicentric study. *Archives of Orthopaedic and Trauma Surgery*, 141(11), 1919-1926. <https://doi.org/10.1007/s00402-020-03643-9>
- Mokkink, L., Terwee, C. & de Vet, H. (2021). Key concepts in clinical epidemiology: Responsiveness, the longitudinal aspect of validity. *Journal of Clinical Epidemiology*, 140, 159-162. <https://doi.org/10.1016/j.jclinepi.2021.06.002>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... de Vet, H. C. (2010a). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19(4), 539-549. <https://doi.org/10.1007/s11136-010-9606-8>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... de Vet, H. C. (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737-745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>
- Morgan, J. H., Kallen, M. A., Okike, K., Lee, O. C. & Vrahas, M. S. (2015). PROMIS Physical Function Computer Adaptive Test Compared With Other Upper Extremity Outcome Measures in the Evaluation of Proximal Humerus Fractures in Patients Older Than 60 Years. *J Orthop Trauma*, 29(6), 257-263. <https://doi.org/10.1097/bot.0000000000000280>
- Mostafa, E., Imonugo, O. & Varacallo, M. (2023). Anatomy, Shoulder and Upper Limb, Humerus. I *StatPearls*. Treasure Island (FL): StatPearls Publishing. Copyright © 2023, StatPearls Publishing LLC.
- Neer, C. S., 2nd. (1970). Displaced proximal humeral fractures. I. Classification and evaluation. *Journal of Bone and Joint Surgery (American Volume)*, 52(6), 1077-1089.

- Neer, C. S., 2nd. (2002). Four-segment classification of proximal humeral fractures: purpose and reliable use. *J Shoulder Elbow Surg*, 11(4), 389-400. <https://doi.org/10.1067/mse.2002.124346>
- Nowak, L. L., Davis, A. M., Mamdani, M., Beaton, D., Kennedy, C. & Schemitsch, E. H. (2019). A Systematic Review and Standardized Comparison of Available Evidence for Outcome Measures Used to Evaluate Proximal Humerus Fracture Patients. *J Orthop Trauma*, 33(7), e256-e262. <https://doi.org/10.1097/bot.0000000000001462>
- Nowak, L. L., Davis, A. M., Mamdani, M., Beaton, D. & Schemitsch, E. H. (2021). A concept analysis and overview of outcome measures used for evaluating patients with proximal humerus fractures. *Disabil Rehabil*, 43(10), 1450-1462. <https://doi.org/10.1080/09638288.2019.1649728>
- Nowak, L. L., Vicente, M. R., McKee, M. D., Hall, J. A., Nauth, A. & Schemitsch, E. H. (2017). Orthopaedic surgeons' opinions surrounding the management of proximal humerus fractures: an international survey. *International Orthopaedics*, 41(9), 1749-1755. <https://doi.org/10.1007/s00264-017-3569-0>
- O'Hara, N. N., Garibaldi, A., Sprague, S., Jackson, J., Kwok, A. K., Beaton, D. E., ... Slobogean, G. P. (2017). Rehabilitation, not injury or treatment details, dominate proximal humeral fracture patient concerns: a thematic analysis. *European Journal for Person Centered Healthcare*, 5(3), 351-356.
- Olerud, P. M. D., Ahrengart, L. M. D. P., Ponzer, S. M. D. P., Saving, J. M. D. & Tidermark, J. M. D. P. (2011). Hemiarthroplasty versus nonoperative treatment of displaced 4-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg*, 20(7), 1025-1033. <https://doi.org/10.1016/j.jse.2011.04.016>
- Othman, A. & Taylor, G. (2004). Is the Constant score reliable in assessing patients with frozen shoulder? 60 shoulders scored 3 years after manipulation under anaesthesia. *Acta Orthopaedica Scandinavica*, 75(1), 114-116. <https://doi.org/10.1080/00016470410001708230>
- Padua, R., de Girolamo, L., Grassi, A. & Cucchi, D. (2021). Choosing patient-reported outcome measures for shoulder pathology. *EFORT Open Rev*, 6(9), 779-787. <https://doi.org/10.1302/2058-5241.6.200109>
- Pais, R., Ruano, L., O, P. C. & Barros, H. (2020). Global Cognitive Impairment Prevalence and Incidence in Community Dwelling Older Adults-A Systematic Review. *Geriatrics (Basel)*, 5(4). <https://doi.org/10.3390/geriatrics5040084>
- Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS (7th ed.)* Allen & Unwin.
- Palvanen, M., Kannus, P., Niemi, S. & Parkkari, J. (2006). Update in the epidemiology of proximal humeral fractures. *Clinical Orthopaedics and Related Research*, 442, 87-92. <https://doi.org/10.1097/01.blo.0000194672.79634.78>

- Palvanen, M., Kannus, P., Parkkari, J., Pitkääjärvi, T., Pasanen, M., Vuori, I. & Järvinen, M. (2000). The injury mechanisms of osteoporotic upper extremity fractures among older adults: a controlled study of 287 consecutive patients and their 108 controls. *Osteoporosis International*, 11(10), 822-831. <https://doi.org/10.1007/s001980070040>
- Pannucci, C. J. & Wilkins, E. G. (2010). Identifying and avoiding bias in research. *Plastic and Reconstructive Surgery*, 126(2), 619-625. <https://doi.org/10.1097/PRS.0b013e3181de24bc>
- Peck, W. A., Burckhardt, P., Christiansen, C., Fleisch, H. A., Genant, H. K., Gennari, C., ... et al. (1993). Consensus development conference: Diagnosis, prophylaxis, and treatment of osteoporosis. *American Journal of Medicine*, 94(6), 646-650. [https://doi.org/10.1016/0002-9343\(93\)90218-E](https://doi.org/10.1016/0002-9343(93)90218-E)
- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49(12), 1373-1379. [https://doi.org/10.1016/S0895-4356\(96\)00236-3](https://doi.org/10.1016/S0895-4356(96)00236-3)
- Perka, C., Arnold, U. & Buttgerit, F. (2000). Influencing factors on perioperative morbidity in knee arthroplasty. *Clinical Orthopaedics and Related Research (1976-2007)*, 378, 183-191.
- Plath, J. E., Kerschbaum, C., Seebauer, T., Holz, R., Henderson, D. J., Förch, S. & Mayr, E. (2019). Locking nail versus locking plate for proximal humeral fracture fixation in an elderly population: a prospective randomised controlled trial. *BMC Musculoskeletal Disorders*, 20(1), 1-13.
- Polit, D. F. & Beck, C. T. (2017). *Nursing Research: Generating and Assessing Evidence for Nursing Practice* Wolters Kluwer Health.
- Rangan, A., Handoll, H., Brealey, S., Jefferson, L., Keding, A., Martin, B. C., ... Torgerson, D. (2015). Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA*, 313(10), 1037-1047. <https://doi.org/10.1001/jama.2015.1629>
- Richard, G. J., Denard, P. J., Kaar, S. G., Bohsali, K. I., Horneff, J. G., Carpenter, S., ... King, J. J. (2020). Outcome measures reported for the management of proximal humeral fractures: a systematic review. *J Shoulder Elbow Surg*, 29(10), 2175-2184. <https://doi.org/10.1016/j.jse.2020.04.006>
- Robinson, B. C., Athwal, G. S., Sanchez-Sotelo, J. & Rispoli, D. M. (2008). Classification and imaging of proximal humerus fractures. *Orthopedic Clinics of North America*, 39(4), 393-403, v. <https://doi.org/10.1016/j.ocl.2008.05.002>
- Rodrigues, J. N. (2020). Different terminologies that help the interpretation of outcomes. *Journal of Hand Surgery, European Volume*, 45(1), 97-99. <https://doi.org/10.1177/1753193419870100>
- Rolstad, S., Adler, J. & Rydén, A. (2011). Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value in Health*, 14(8), 1101-1108. <https://doi.org/10.1016/j.jval.2011.06.003>

- Roy, J. S., MacDermid, J. C. & Woodhouse, L. J. (2010). A systematic review of the psychometric properties of the Constant-Murley score. *J Shoulder Elbow Surg*, *19*(1), 157-164.  
<https://doi.org/10.1016/j.jse.2009.04.008>
- Shulman, B. S., Ong, C. C., Lee, J. H., Karia, R., Zuckerman, J. D. & Egol, K. A. (2013). Outcomes after fixation of proximal humerus (OTA type 11) fractures in the elderly patients using modern techniques. *Geriatric Orthopaedic Surgery & Rehabilitation*, *4*(1), 21-25.
- Slobogean, G. P., Noonan, V. K. & O'Brien, P. J. (2010). The reliability and validity of the Disabilities of Arm, Shoulder, and Hand, EuroQol-5D, Health Utilities Index, and Short Form-6D outcome instruments in patients with proximal humeral fractures. *J Shoulder Elbow Surg*, *19*(3), 342-348. <https://doi.org/10.1016/j.jse.2009.10.021>
- Slobogean, G. P. & Slobogean, B. L. (2011). Measuring shoulder injury function: common scales and checklists. *Injury*, *42*(3), 248-252. <https://doi.org/10.1016/j.injury.2010.11.046>
- Stone, A., Shiffman, S., Atienza, A. & Nebeling, L. (2007). *The science of real-time data capture: Self-reports in health research* Oxford University Press.
- Sumrein, B. O., Huttunen, T. T., Launonen, A. P., Berg, H. E., Felländer-Tsai, L. & Mattila, V. M. (2017). Proximal humeral fractures in Sweden-a registry-based study. *Osteoporosis International*, *28*(3), 901-907. <https://doi.org/10.1007/s00198-016-3808-z>
- Südkamp, N. P., Audigé, L., Lambert, S., Hertel, R. & Konrad, G. (2011). Path analysis of factors for functional outcome at one year in 463 proximal humeral fractures. *Journal of Shoulder and Elbow Surgery*, *20*(8), 1207-1216.
- Terry, G. C. & Chopp, T. M. (2000). Functional anatomy of the shoulder. *J Athl Train*, *35*(3), 248-255.
- Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., ... de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34-42.  
<https://doi.org/10.1016/j.jclinepi.2006.03.012>
- Thorin, M. H., Wihlborg, A., Åkesson, K. & Gerdhem, P. (2016). Smoking, smoking cessation, and fracture risk in elderly women followed for 10 years. *Osteoporosis International*, *27*(1), 249-255. <https://doi.org/10.1007/s00198-015-3290-z>
- U.S. Department of Health. (2006). *Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims: Draft guidance* (1477-7525). LONDON: LONDON: Springer Nature.
- Ustün, T. B., Chatterji, S., Bickenbach, J., Kostanjsek, N. & Schneider, M. (2003). The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health. *Disabil Rehabil*, *25*(11-12), 565-571.  
<https://doi.org/10.1080/0963828031000137063>

- van de Water, A. T., Davidson, M., Shields, N., Evans, M. C. & Taylor, N. F. (2016). The Shoulder Function Index (SFInX): evaluation of its measurement properties in people recovering from a proximal humeral fracture. *BMC Musculoskeletal Disorders*, 17, 295. <https://doi.org/10.1186/s12891-016-1138-0>
- van de Water, A. T., Shields, N., Davidson, M., Evans, M. & Taylor, N. F. (2014). Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture. *Disabil Rehabil*, 36(13), 1072-1079. <https://doi.org/10.3109/09638288.2013.829529>
- van de Water, A. T., Shields, N. & Taylor, N. F. (2011). Outcome measures in the management of proximal humeral fractures: a systematic review of their use and psychometric properties. *J Shoulder Elbow Surg*, 20(2), 333-343. <https://doi.org/10.1016/j.jse.2010.10.028>
- van Oostwaard, M. (2018). Osteoporosis and the Nature of Fragility Fracture: An Overview. I K. Hertz & J. Santy-Tomlinson (Red.), *Fragility Fracture Nursing: Holistic Care and Management of the Orthogeriatric Patient* (s. 1-13). Cham (CH): Springer. Copyright 2018, The Editor(s)(if applicable) and the Author(s).
- Vittinghoff, E. & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, 165(6), 710-718. <https://doi.org/10.1093/aje/kwk052>
- Vrotsou, K., Ávila, M., Machón, M., Mateo-Abad, M., Pardo, Y., Garin, O., ... Cuéllar, R. (2018). Constant-Murley Score: systematic review and standardized evaluation in different shoulder pathologies. *Quality of Life Research*, 27(9), 2217-2226. <https://doi.org/10.1007/s11136-018-1875-7>
- Wade, D. T. & Halligan, P. W. (2017). The biopsychosocial model of illness: a model whose time has come. *Clinical Rehabilitation*, 31(8), 995-1004. <https://doi.org/10.1177/0269215517709890>
- Werner, B. C., Griffin, J. W., Yang, S., Brockmeier, S. F. & Gwathmey, F. W. (2015). Obesity is associated with increased postoperative complications after operative management of proximal humerus fractures. *Journal of Shoulder and Elbow Surgery*, 24(4), 593-600.
- Wifstad, Å. (2018). *Vitenskapsteori for helsefagene*. Oslo: Universitetsforlaget.
- Xu, B., Anderson, D. B., Park, E.-S., Chen, L. & Lee, J. H. (2021). The influence of smoking and alcohol on bone healing: Systematic review and meta-analysis of non-pathological fractures. *EClinicalMedicine*, 42, 101179-101179. <https://doi.org/10.1016/j.eclinm.2021.101179>
- Yan, T., Fricker, S. & Tsai, S. (2020). Response Burden: What Is It and What Predicts It? *Advances in Questionnaire Design, Development, Evaluation and Testing* (s. 193-212).
- Ye, T., Wang, L., Zhuang, C., Wang, Y., Zhang, W. & Qiu, S. (2013). Functional outcomes following locking plate fixation of complex proximal humeral fractures. *Orthopedics*, 36(6), e715-722. <https://doi.org/10.3928/01477447-20130523-14>

- Yian, E. H., Ramappa, A. J., Arneberg, O. & Gerber, C. (2005). The constant score in normal shoulders. *Journal of Shoulder and Elbow Surgery*, 14(2), 128-133.  
<https://doi.org/https://doi.org/10.1016/j.jse.2004.07.003>
- Zhu, Y., Lu, Y., Shen, J., Zhang, J. & Jiang, C. (2011). Locking Intramedullary Nails and Locking Plates in the Treatment of Two-Part Proximal Humeral Surgical Neck Fractures: A Prospective Randomized Trial with a Minimum of Three Years of Follow-up. *Journal of Bone and Joint Surgery (American Volume)*, 93(2), 159-168. <https://doi.org/10.2106/JBJS.J.00155>
- Aasheim, T. & Finsen, V. (2014). The DASH and the QuickDASH instruments. Normative values in the general population in Norway. *Journal of Hand Surgery, European Volume*, 39(2), 140-144. <https://doi.org/10.1177/1753193413481302>

## 8.0 Vedlegg

### Vedlegg 1: CMS spørreskjema med forklaring av utregning

OUT-PATIENT CLINIC	SHOULDER UNIT
CONSTANT SCORE	
<div style="border: 1px solid black; padding: 5px; min-height: 50px;">Patient's Details</div>	<b>Operation/Diagnosis:</b> _____ <b>Date:</b> _____ <b>Side:</b> R L
	<b>Examination:</b> Pre-op 3 months      6months 1 year        2 years        _____ years
<b>A.- Pain (/15): Average (1 + 2)</b> <input style="width: 30px; border: 1px solid black;" type="text"/> <b>A</b>	
1. Do you have pain in your shoulder (normal activities)? No = 15 pts,      Mild pain = 10 pts,      Moderate = 5 pts,      Severe or permanent = 0. _____	
2. Linear scale: If "0" means no pain and "15" is the maximum pain you can experience, please circle where is the level of pain of your shoulder. (Points given are inverse to the scale. E.g. level 5 in the scale means 10 points)	
<b>Level of pain:</b> <input style="width: 400px; height: 15px; background-color: #cccccc;" type="text"/> Points: <input style="width: 400px; height: 15px; border: 1px solid black;" type="text"/> _____	
15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0	
<b>B.- Activities of daily living (/20) Total (1 + 2 + 3 + 4)</b> <input style="width: 30px; border: 1px solid black;" type="text"/> <b>B</b>	
1. Is your occupation or daily living limited by your shoulder? No = 4,      Moderate limitation = 2,      Severe limitation = 0      _____	
2. Are your leisure and recreational activities limited by your shoulder? No = 4,      Moderate limitation = 2,      Severe limitation = 0      _____	
3. Is your night sleep disturbed by your shoulder? No = 2,      Sometimes = 1,      Yes = 0      _____	
4. State to what level you can use your arm for painless, reasonably activities. Waist = 2,      Xiphoid (sternum) = 4,      Neck = 6,      Head = 8,      Above head = 10      _____	
<b>C.- Range of movement (leave this for the doctor or physiotherapist) (/40): Total (1 + 2 + 3 + 4)</b> <input style="width: 30px; border: 1px solid black;" type="text"/> <b>C</b>	
<b>1.- FWD Flexion:</b> 0-30      0 pts _____ 31-60      2 pts 61-90      4 pts 91-120      6 pts 121-150      8 pts > 150      10 pts	<b>2.- Abduction:</b> 0-30 _____ 31-60 61-90 91-120 121-150 > 150
<b>3.- External Rotation:</b> _____ Hand behind head & elbow forward 2 Hand behind head & elbow back 4 Hand above head & elbow forward 6 Hand above head & elbow back 8 Full elevation of arm 10	<b>4.- Internal Rotation: (Dorsum hand to)</b> _____ Thigh 0 Buttock 2 SI joint 4 Waist 6 T12 8 Between shoulder blades 10
<b>D.- Power (/25): Points: average (kg) x 2 =</b> <input style="width: 30px; border: 1px solid black;" type="text"/> <b>D</b>	
First pull:      Second pull:      Third pull:      Fourth pull:      Fifth pull: Average pulls:	
<b>TOTAL (/100): A + B + C + D</b> <input style="width: 50px; border: 1px solid black;" type="text"/>	

## Vedlegg 2: DASH spørreskjema med forklaring av utregning

### Poengberegning av DASH

## Dysfunksjon i arm, skulder og hånd

### Dysfunksjon/symptom (De første 30 spørsmålene)

Svaralternativene for hvert spørsmål poengsettes fra 1 (ingen funksjonsnedsettelse/symptomer) til 5 (verste funksjonsnedsettelse/symptom).

#### Beregning av DASH score:

Legg sammen poengene, trekk fra 30, og del på 1,2

Dvs:  $\text{DASH score} = (\text{Poeng} - 30) / 1,2$

- Er opp til 3 spørsmål ubesvart, skal disse erstattes med gjennomsnittet av svarene på de øvrige spørsmålene. For eksempel: om en person har 2 ubesvarte spørsmål og har besvart alle de 28 andre spørsmålene tilsvarende 3 poeng, gis de 2 ubesvarte spørsmålene også 3 poeng.
- Om flere enn 3 spørsmål er ubesvart, går det ikke an å beregne DASH score.
- Det gjør det lettere om en teller opp hvor mange "1-ere" det er, hvor mange "2-ere", hvor mange "3-ere" osv og så ganger ut og legger sammen for å få poengsummen. Det er da også lett å se hvor mange spørsmål som er besvart.

### Arbeid (4 spørsmål, det er valgfritt om denne delen brukes)

Svaralternativer for hvert spørsmål poengsettes fra 1 (ingen) til 5 (ikke mulig).

#### Beregning av score:

Legg sammen poengene, trekk fra 4, og del på 0,16

Dvs:  $(\text{Poeng} - 4) / 0,16$

- Alle 4 spørsmålene må være besvart for å kunne beregne denne score.

### Musikk/idrett

Samme regler som for "Arbeid"

Den norske oversettelsen er vurdert og offisielt godkjent av AAOS. Vennligst referere til denne nettadressen hvis den norske utgaven av skjemaet brukes i en publikasjon.



## HELSEUNDERSØKELSE

(arm/skulder/hånd)

Dette skjemaet tar for seg dine symptomer og dine evner til å utføre visse aktiviteter.

Vær snill å svare på **alle** spørsmål, basert på hvordan det har gått **den siste uken**.

Dersom det er noen aktiviteter du ikke har utført siste uken, skal du krysse for det svaret som du mener ville stemme best om du hadde utført aktiviteten.

Det har ingen betydning hvilken arm eller hånd du bruker for å utføre aktiviteten. Basere svarene på hva du får til, uansett hvordan du utfører oppgaven.

Vennligst sett kryss for ett svaralternativ for hvert spørsmål.

Navn: \_\_\_\_\_ født: \_\_\_\_\_ Dato: \_\_\_\_\_

	Ingen vanskelig- heter	Lette vanskelig- heter	Middels vanskelig- heter	Svære vanskelig- heter	Umulig å gjøre
1. Åpne et nytt syltetøyglass	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Skrive	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Vri om en nøkkel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Forberede et måltid	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Skyve åpen en tung dør	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Legge noe på en hylle over hodehøyde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Utføre tungt husarbeide (f.eks. vaske gulv eller vegger)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Gjøre hagearbeid	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Re opp en seng	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Bære handlepose eller dokumentmappe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Bære tunge gjenstand (over 5 kilo)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Skifte en lyspære over hodehøyde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Vaske eller føne håret	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Vaske ryggen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Ta på en genser	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Skjære opp mat med kniv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Fritidsaktiviteter som krever lite anstrengelse (f.eks spille kort, strikke o.l)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Fritidsaktiviteter som krever en viss kraft eller styrke i arm, skulder eller hånd (f.eks spille golf, bruke hammer, spille tennis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Fritidsaktiviteter der du beveger armen fritt (f.eks spille badminton, svømme, gymnastikk)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Nødvendig transport (Komme deg fra ett sted til et annet)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. Seksuelle aktiviteter	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

22. I hvilken grad har dine arm-, skulder- eller håndproblemer hemmet din vanlige omgang med slektninger, venner, naboer eller andre **den siste uken?** (Sett ett kryss.)

Ikke hemmet i det hele tatt    Litt    Moderat    Ganske mye    Ekstremt

23. Var du begrenset på grunn av dine arm-, skulder- eller håndproblemer i ditt arbeide eller andre vanlige daglige aktiviteter i løpet av **den siste uken?**

Ikke begrenset i det hele tatt    Litt    Moderat begrenset    Svært begrenset    Umulig

Angi alvorlighetsgraden av de følgende symptomene i **den siste uken:**

	Ingen	Lett	Moderat	Sterk	Ekstrem
24. Smerte i arm, skulder eller hånd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Smerte i arm, skulder eller hånd i forbindelse med en spesiell aktivitet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Prikking ("mauring", "sovnet") arm, skulder eller hånd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. Svakheter i arm, skulder eller hånd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Stivhet i arm, skulder eller hånd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

29. Hvor mye vansker har du hatt **den siste uken** med å sove på grunn av smerte i arm, skulder eller hånd?

Ingen vansker    Litt vansker    Moderate vansker    Betydelige vansker    Har ikke fått sove

30. Jeg føler meg mindre handlekraftig, har mindre selvtillit eller føler meg mindre nyttig på grunn av mitt arm-, skulder- eller håndproblem.

Helt uenig    Uenig    Hverken enig eller uenig    Enig    Helt enig

De følgende spørsmålene dreier seg om hvor mye dine arm-, skulder- eller håndproblemer påvirker din evne til å arbeide (inkludert husarbeid om dette er din hovedbeskjeftigelse).

Arbeider du? Ja  Nei

**Dersom svaret er nei, kan du hoppe over de fire spørsmålene**

Hva er ditt yrke/arbeid (Hva gjør du)?

Kryss av for den påstanden som best beskriver dine fysiske prestasjoner **den siste uken**. Hadde du noen vanskeligheter med å...:

	Ingen	Litt	Moderate	Store	Ikke mulig
1. ...bruke din vanlige teknikk i ditt arbeide?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. ...utføre ditt vanlige arbeide pga smerte i arm, skulder eller hånd?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. ...utføre ditt arbeide så bra som du skulle ønske?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ...utføre arbeidet på den tid du vanligvis bruker?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

De følgende spørsmålene dreier seg om hvor mye dine arm-, skulder- eller håndproblemer har påvirket dine evner til å spille ditt musikkinstrument og/eller drive idrett.

Spiller du noe instrument eller driver noen idrett? Ja  Nei

**Dersom svaret er nei, kan du hoppe over resten av spørsmålene**

Om du spiller mer enn ett musikkinstrument eller driver mer enn en idrett, skal du svare med hensyn til den aktiviteten som er viktigst for deg.

Hvilket instrument eller idrett er viktigst for deg: \_\_\_\_\_

Kryss av for påstanden som best beskriver dine fysiske prestasjoner **den siste uken**. Hadde du noen vanskeligheter med å...:

	Ingen	Litt	Moderate	Store	Ikke mulig
1. ...bruke din vanlige teknikk for å spille instrument/drive idrett?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. ...spille instrument/drive idrett pga smerte i arm, skulder eller hånd?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. ...spille instrument/drive idrett så bra som du skulle ønske?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ...bruke like mye tid som vanlig på å spille instrument/drive idrett?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## OXFORD SKULDERSKÅR

Dato:  
 Pasientnavn:  
 Inklusjonsnummer:

Kryss av én boks (✓) på hvert av de 12 spørsmålene.

### 1. I løpet av de siste fire uker...

Hvordan vil du beskrive de **verste** smertene du hadde i skulderen?

Ingen	Svake	Moderate	Sterke	Uutholdelige
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 2. I løpet av de siste fire uker...

Har du hatt problemer med å kle på deg på grunn av skulderen?

Ingen problemer i det hele tatt	Litt problemer	Moderate problemer	Ekstreme vanskeligheter	Umulig å gjøre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 3. I løpet av de siste fire uker...

Har du hatt problemer med å komme inn og ut av bil eller bruke kollektivtransport på grunn av skulderen? (Svar ut fra det du pleier å bruke)

Ingen problemer i det hele tatt	Svært lite problemer	Moderate problemer	Ekstreme vanskeligheter	Umulig å gjøre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 4. I løpet av de siste fire uker...

Har du kunnet bruke kniv og gaffel – samtidig?

Ja, uten vansker	Med litt vansker	Med moderate vansker	Med ekstreme vansker	Nei, umulig
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 5. I løpet av de siste fire uker...

**Kunne** du gjøre innkjøpene til huset på egen hånd?

Ja, uten vansker	Med litt vansker	Med moderate vansker	Med ekstreme vansker	Nei, umulig
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 6. I løpet av de siste fire uker...

**Kunne** du bære et brett med en tallerken mat gjennom et rom?

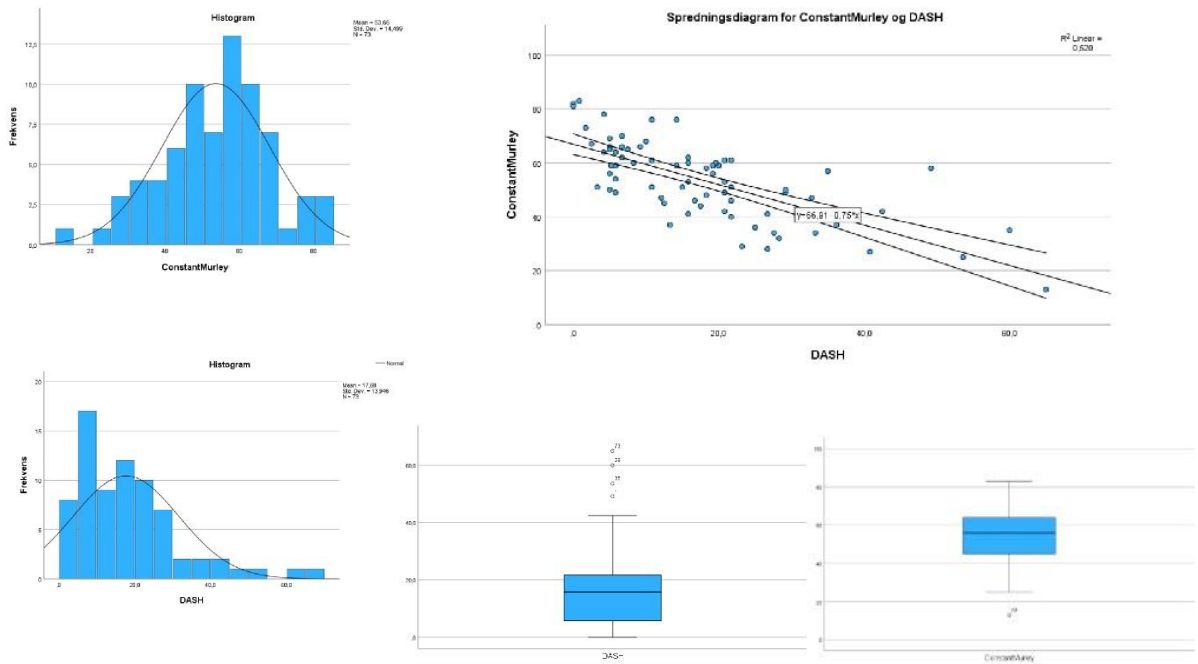
Ja, uten vansker	Med litt vansker	Med moderate vansker	Med ekstreme vansker	Nei, umulig
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

<b>7. I løpet av de siste fire uker...</b>					
Kunne du børste/gre håret <u>med den vonde armen</u> ?					
Ja, uten vansker	Med litt vansker	Med moderate vansker	Med ekstreme vansker	Nei, umulig	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>8. I løpet av de siste fire uker...</b>					
Hvordan vil du beskrive smertene du <u>vanligvis</u> hadde i skulderen?					
Ingen	Veldig svake	Svake	Moderate	Sterke	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>9. I løpet av de siste fire uker...</b>					
Kunne du henge klærne dine opp i et klesskap <u>med den vonde armen</u> ?					
Ja, uten vansker	Med litt vansker	Med moderate vansker	Med ekstreme vansker	Nei, umulig	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>10. I løpet av de siste fire uker...</b>					
Har du kunnet vaske og tørke deg under begge armene?					
Ja, uten vansker	Med litt vansker	Med moderate vansker	Med ekstreme vansker	Nei, umulig	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>11. I løpet av de siste fire uker...</b>					
Hvor mye har <u>skuldersmertene</u> forstyrret ditt vanlige arbeid (inkludert husarbeid)?					
Ikke i det hele tatt	Litt	Moderat	Mye	Fullstendig	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>12. I løpet av de siste fire uker...</b>					
Har du vært plaget med <u>skuldersmertene</u> i sengen om natten?					
Ingen netter	Bare 1 eller 2 netter	Noen netter	De fleste netter	Hver natt	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

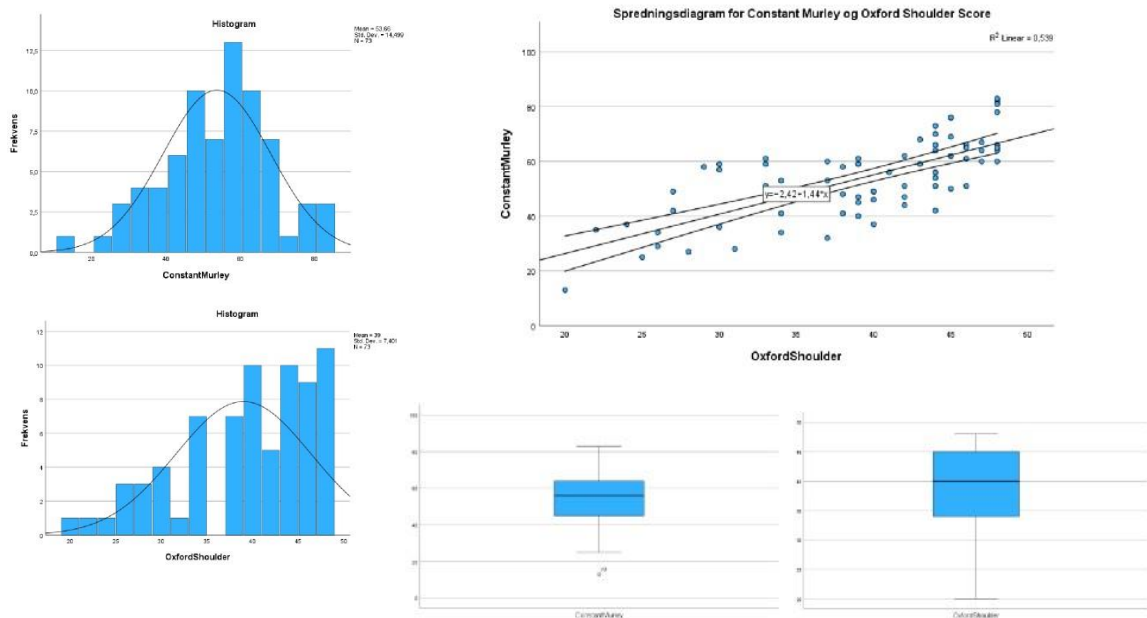
Til slutt, vennligst sjekk at du har svart på hvert spørsmål.

Tusen takk.

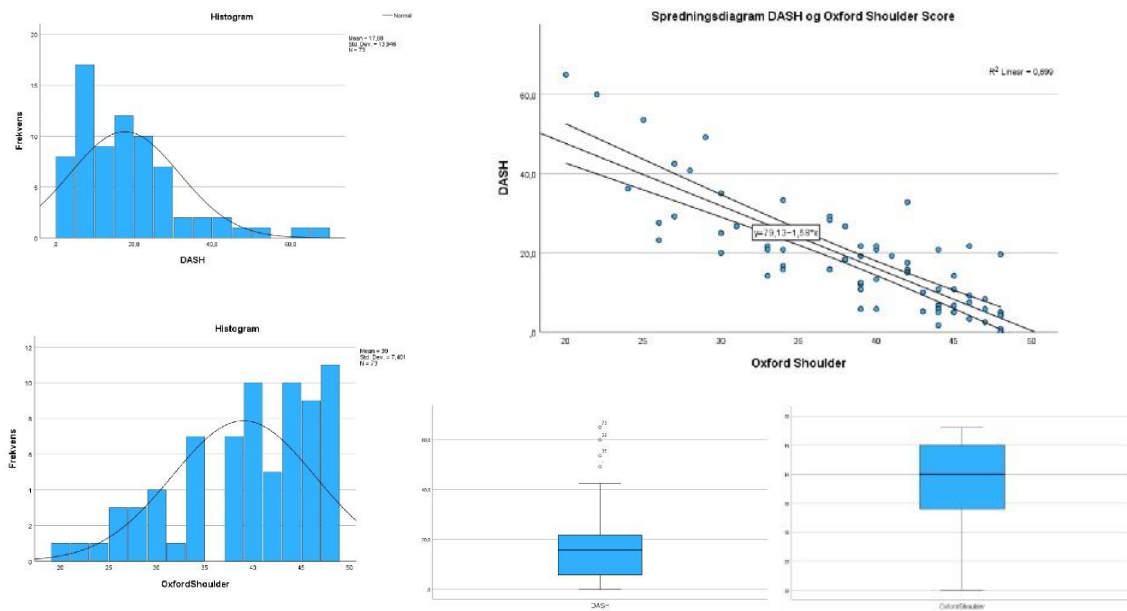
**Vedlegg 4:** Forutsetninger for Pearsons korrelasjon CMS og DASH testet ved normalfordeling i histogram, identifisere uteliggere i boksplott, og linearitet ved spredningsdiagram



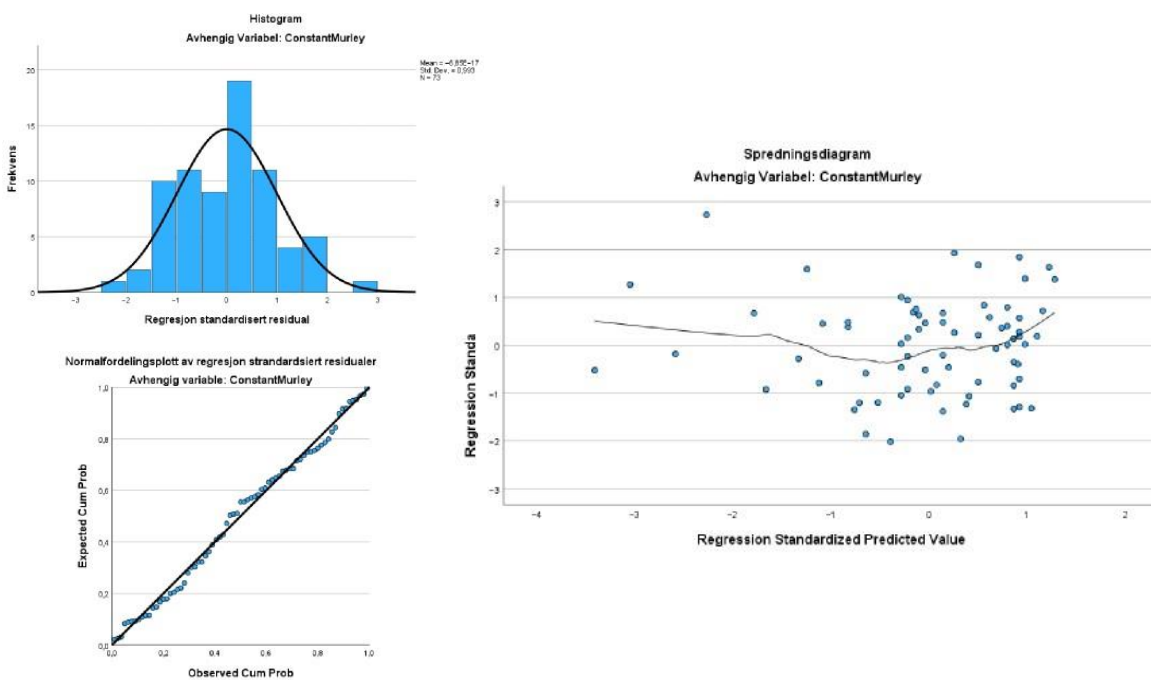
**Vedlegg 5:** Forutsetninger for Pearsons korrelasjon CMS og OSS testet ved normalfordeling i histogram, identifisere uteliggere i boksplott, og linearitet ved spredningsdiagram



**Vedlegg 6:** Forutsetninger for Pearsons korrelasjon DASH og OSS testet ved normalfordeling i histogram, identifisere uteliggere i boksploott og linearitet ved spredningsdiagram

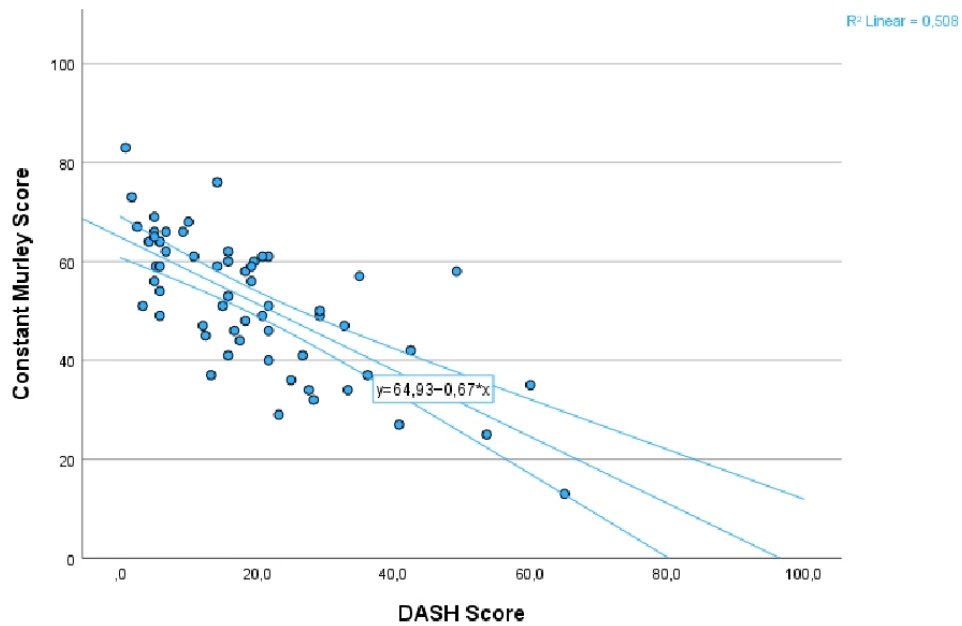


**Vedlegg 7:** Normalfordelingstest av residualer i spredningsdiagram, histogram, normalfordelingsplott og homoskedastisitet i spredningsdiagram CMS og DASH

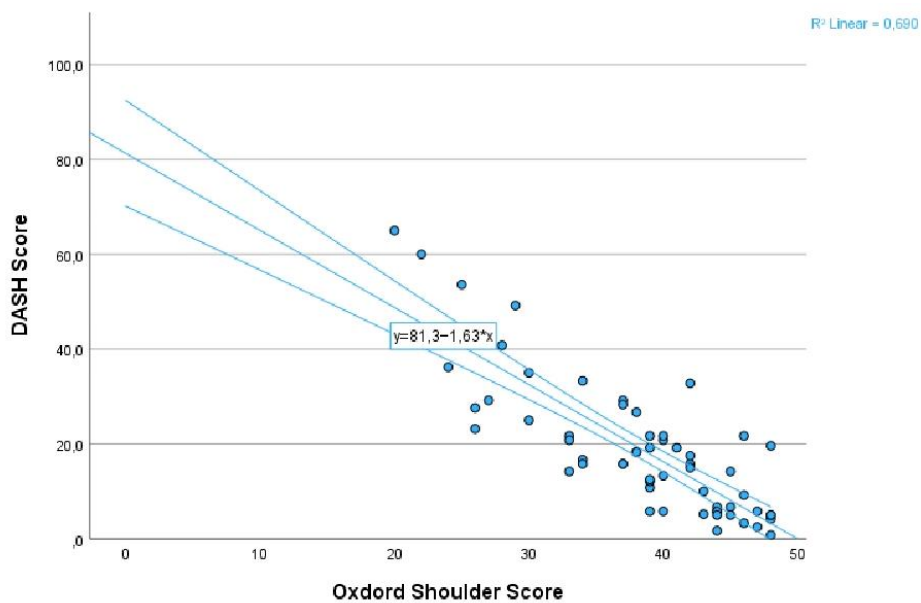




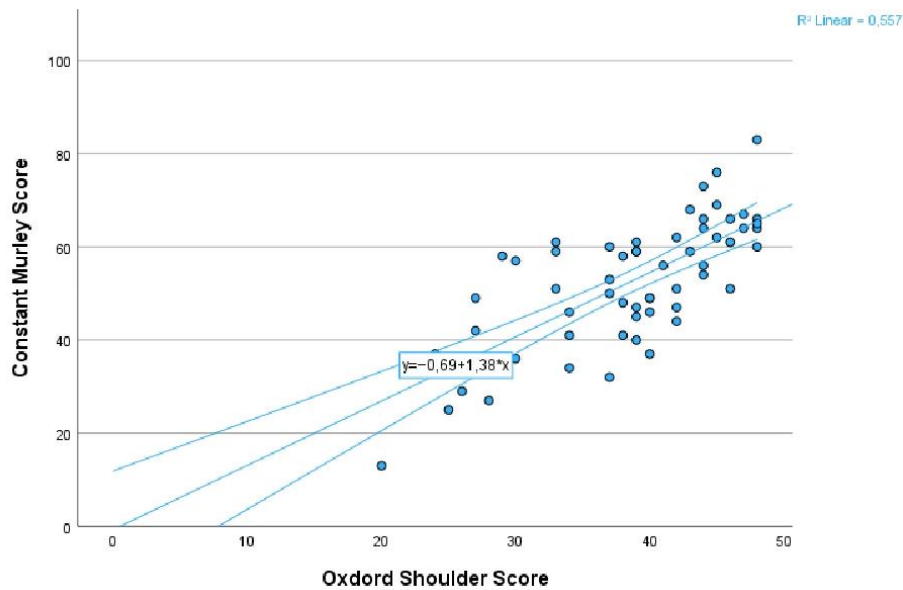
**Vedlegg 8:** Spredningsdiagram mellom funksjonsdata for DASH og CMS kun kvinner=60



**Vedlegg 9:** Spredningsdiagram mellom funksjonsdata for DASH og OSS kun kvinner=60



**Vedlegg 10:** Spredningsdiagram mellom funksjonsdata for CMS og OSS kun kvinner=60



**Vedlegg 11:** Korrelasjonstabell mellom funksjonsdata ved bruk av Pearsons korrelasjonskoeffisient kun kvinner n=60

	CMS	DASH	OSS
CMS	1		
DASH	-071**	1	
OSS	0,75**	-0,83**	1

Pearsons korrelasjonkoeffisient (r) 0,10-0,29 vurderes som svak, 0,30-0,49 vurderes som moderat, >0.5 vurderes som sterk

DASH, disabilities of arm, shoulder and hand; OSS, Oxford Shoulder Score; CMS, Constant Murley Score.

\*\* Signifikant  $p < 0,01$

**Vedlegg 12:** : Lineær regresjon mellom CMS, DASH og OSS kvinner n=60

Uavhengig variabel	Avhengig variabel	B	P	95%KI
DASH	CMS	-0,67	<0,01	-0,85, -0,45
OSS	CMS	1,38	<0,01	1,1, 1,7
OSS	DASH	1,63	<0,01	-1,91, -1,34

DASH, disabilities of arm, shoulder and hand; OSS, Oxford Shoulder Score; CMS, Constant Murley Score.

B1, regresjonskoeffisient

KI, konfidensintervall

Std.feil, standardfeil; feilmarginen

P, p-verdi

**Vedlegg 13:** Logistisk regresjon uten røyk (n=73)

Uavhengig variabel	Ujustert OR (95%KI)	p-verdi	Justert OR (95%KI)	p-verdi
Alder	1,03 (0,98-1,08)	0,27	0,96 (0,89- 1,05)	0,37
Kjønn (kvinne)	1,71 (0,50-5,83)	0,39	1,16 (0,27-4,91)	0,84
BMI	0,99 (0,90-1,09)	0,83	0,97 (0,87-1,09)	0,63
Pensjonist/ufør	2,06 (0,75-5,56)	0,16	2,04 (0,46-9,10)	0,35
ASA 1	(ref.)	0,05	(ref.)	0,07
ASA 2	2,24 (0,80-6,28)	0,13	2,68 (0,77-9,37)	0,12
ASA 3	8,00 (1,40-45,76)	0,02*	9,89 (1,39-70,33)	0,02*

Avhengig dikotom kategorisk variabel: CMS skår >56=0, <56=1

CMS, Constant Murley Score.

\* Signifikant p<0,05

^ >15% økning i estimatene uten to observasjoner med høy Cook's D verdi

OR, odds ratio

KI, konfidensintervall

Ref, referansegruppe

## Vedlegg 14: Godkjenning av endringsmelding fra REK



<b>Region:</b>	<b>Saksbehandler:</b>	<b>Telefon:</b>	<b>Vår dato:</b>	<b>Vår referanse:</b>
REK sør-øst D	Finn Skre Fjordholm	+47 22 84 58 21	14.12.2022	28707

Annette Konstanse Bordewich Wikerøy

**Prosjektsøknad:** Operativ behandling av 3 og 4 parts brudd i skulderen, valget mellom nagle eller plate  
**Søknadsnummer:** 2016/626  
**Forskningsansvarlig institusjon:** Akershus universitetssykehus HF

### Prosjektsøknad: Endring godkjennes

#### Søkers beskrivelse

*Insidensen av brudd i øvre humerus-ende (PHF) øker i Norge, spesielt hos de eldre. Nye vinkelstabile implantater gir bedre stabilitet og reduserer komplikasjonene. De vanligste komplikasjonene er at skruer penetrerer brusken eller at implantatet slipper noe av taket. Eksakt reposisjon, adekvat fiksasjon og riktig medial støtte (støtte på innsiden av halsen av skulderkula) er viktige faktorer for suksess, faktorer som vi legger spesiell vekt på når vi sammenligner vinkelstabile plater med nagler ved 3 eller 4 parts PHF i et randomisert kontrollert prospektivt studie. Vi vil se hvilket implantat som har minst komplikasjoner, som lønner seg økonomisk og som gir pasienten minst plager og best funksjon. Oppfølgingen skal skje over 2 år med standardiserte tester og røntgenbilder. Operasjonstilgangen er tilnærmet lik og implantatene viser lignende resultater når de er undersøkt alene, men til vår viten er de aldri tidligere sammenlignet systematisk for 3 og 4 parts brudd i PHF.*

Vi viser til søknad om prosjektendring mottatt 24.11. 2022 for ovennevnte forskningsprosjekt. Søknaden er behandlet av leder for REK D på delegert fullmakt fra komiteen, med hjemmel i forskningsetikkforskriften § 7, første ledd, tredje punktum. Søknaden er vurdert med hjemmel i helseforskningsloven § 11.

Prosjektleder har sendt tilleggsopplysninger til endringsøknaden pr. epost den 08.12.2022 og 12.12.2022.

#### Endringene innebærer:

1) Det skal gjennomføres en tverrsnittstudie som understudie til hovedstudien. I understudien skal man undersøke hvilken assosiasjon det er mellom Constant score, DASH og EQ-5D hos pasienter med PHFx 6 måneder postoperativt, og hvilke bakgrunnsvariabler som assosieres med lav CMS sumskår. Resultatene av underprosjektet vil kunne bidra med viktig informasjon om hvordan resultatene i hovedstudien kan

fortolkes.

2) En ny medarbeider er inkludert i prosjektet.

- Kari Krokan, masterstudent Tverrfaglig helseforskning, Universitetet i Oslo.

3) Korrigering av opplysninger om kontrollrutiner etter operasjon. Det er fremkommet at det også er utført kontroll også 6 måneder etter operasjon.

Vedlagt fulgte prosjektplan for tverrsnittstudien "Assosiasjon mellom utfallsmål hos pasienter med proksimal humerus fraktur 6 måneder postoperativt"

### **REKs vurdering**

Prosjektleder opplyser at 2 fysioterapeuter fra start av prosjektet har innhentet data fra inkluderte pasienter på 6 måneder og 2 års kontrollene for å sikre blinding. Den ene fysioterapeuten går master i Tverrfaglig helseforskning ved UiO. Hun ønsker å benytte deler av datamaterialet/ utfallsmål som hun har samlet inn fra pågående prosjekt til sin masteroppgave. Prosjektleder opplyser i avgitt tilleggsinformasjon at alle pasienter har vært til kontroll ved 6 måneder, men at informasjonsskrivet inneholder en feil/manglende opplysning når det gjelder kontroll ved 6 måneder. Det opplyses videre at det ikke skal rekrutteres flere pasienter og ingen nye data skal innhentes i forbindelse med understudien.

Komiteen har ikke innvendinger til at data fra kontroll av pasienter ved 6 måneder inkluderes i studien, da det forutsettes at pasientene har samtykket til dette ved å møte opp til kontroll ved 6 måneder. Komiteen vektlegger at det er viktige spørsmål som skal belyses.

Komiteens leder har vurdert søknaden og det fremmes ingen forskningsetiske innvendinger til endringene i prosjektet.

### **Vedtak**

REK har gjort en forskningsetisk vurdering av endringene i prosjektet og godkjenner prosjektet slik det nå foreligger, jfr. helseforskningsloven § 11 annet ledd.

Vi gjør oppmerksom på at etter ny personopplysningslov må det også foreligge et behandlingsgrunnlag etter personvernforordningen. Det må forankres i egen institusjon.

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

### **Sluttmelding**

Prosjektleder skal sende sluttmelding til REK på eget skjema via REK-portalene senest 6 måneder etter sluttdato 01.09.2024, jf. helseforskningsloven § 12. Dersom prosjektet ikke starter opp eller gjennomføres meldes dette også via skjemaet for sluttmelding.

### **Søknad om endring**

Dersom man ønsker å foreta vesentlige endringer i formål, metode, tidsløp eller organisering må prosjektleder sende søknad om endring via portalen på eget skjema til REK, jf. helseforskningsloven § 11.

**Klageadgang**

Du kan klage på REKs vedtak, jf. forvaltningsloven § 28 flg. Klagen sendes på eget skjema via REK portalen. Klagefristen er tre uker fra du mottar dette brevet. Dersom REK opprettholder vedtaket, sender REK klagen videre til Den nasjonale forskningsetiske komité for medisin og helsefag (NEM) for endelig vurdering, jf. forskningsetikkloven § 10 og helseforskningsloven § 10.

Med vennlig hilsen

Pål Aukrust  
Prof. em. dr. med.  
Komitéleder REK sør-øst D

Anne Åbyholm-Brodal  
seniorkonsulent

*Kopi til:*

Akershus universitetssykehus HF  
Kari Finnbråten Krokan