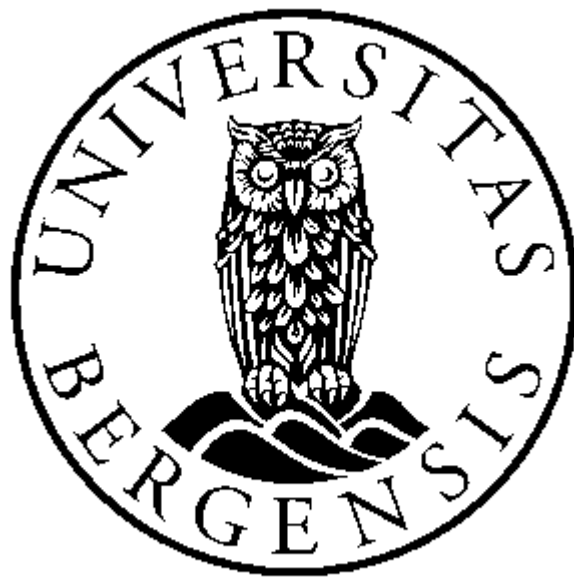


**Intra –og intertester reliabilitet for ultralydsmålinger af
patellarsenens tykkelse hos individer med symptomatiske
og asymptomatiske patellarsener**



Niels Bjerre Thygesen

Masteropgave i helsefag – studieretning fysioterapividenskab

Sektion for fysioterapividenskab

Institut for global helse og samfundsmedicin

Universitet i Bergen

4. semester

Foråret 2017

Forord

I mit kliniske arbejde har muskuloskeletale lidelser haft stor bevågenhed, og interessen har specielt været førende indenfor tendinopati. Derfor var udgangspunktet for mit masterprojekt tendinopati. Processen har været krævende, men ekstrem lærerigt. Jeg har fået en masse viden omkring forskningsmetode samt erfaringen med strukturering og projekteringen af et forskningsprojekt. Ikke mindst har arbejdet med dette masterprojekt givet mig et meget bredere teoretisk perspektiv til at sammenfatte evidensbaseret kundskab og klinisk praksis.

Jeg skylder en stor tak til:

Min familie som har vist stor støtte og videt stor plads til færdiggørelse af denne opgave.

Min hovedvejleder Jon Joensen som gjorde det mulig at gennemføre dette projekt. Tak for hjælp under oplæringen i RTUS samt indsamlingen af data. Tak for at du stillede op som tester. Ikke mindst tak for gode diskussioner og tilbagemeldinger.

Ida Ersland Kvinge som stillede op som tester og hjalp med blindingen.

Rolf Moe-Nilssen for konstruktive tilbagemeldinger og hjælp med indsigt i det teoretisk perspektiv rund reliabilitet.

Jan Magnus Bjordal og Ingvill Fjell Naterstad for inspiration til ideen bag projektet.

Niels Bjerre Thygesen

Bergen, foråret 2017

Intra –og intertester reliabilitet for ultralydsmålinger af patellarsenens tykkelse hos individer med symptomatiske og asymptomatiske patellarsener

Resumé

Baggrund: RTUS er på fremmarch som et tilgængeligt måleinstrument blandt fysioterapeuter specielt inden diagnostik af superficielle strukturer som sener og anvendes hyppigt i studier omhandlende patellar tendinopati. Alligevel har få studier undersøgt reliabilitet for målinger tilknyttet RTUS.

Formål: At beskrive intra -og intertester reliabiliteten for målinger af patellarsenens tykkelse med RTUS som målemetode på en gruppe individer med symptomatiske sener og en gruppe individer med asymptomatiske sener.

Design: Tværsnitsdesign – En beskrivende undersøgelse af intra -og intertester reliabiliteten.

Metode: Sonografiske billeder af 30 symptomatiske og 33 asymptomatiske patellarsener blev inkluderet. Tre testere med forskellig grad af erfaring foretog målinger af senerne på longitudinale og transversale billeder. Alle tre testere indgik i intertester reliabilitetsanalysen og to tester i intratester reliabilitetsanalysen.

Resultat: Intratester reliabilitetsanalysen viste en højere reliabilitet for posterior-anterior målinger af symptomatiske og asymptomatiske sener longitudinal (ICC: 0.88-0.96, SDC: 0.53-1.08mm) end transversal (ICC: 0.87-0.96, SDC: 0.55-2.00mm). Intertester reliabilitetsanalysen viste ligeledes en højere reliabilitet for målinger longitudinal (ICC: 0.87-0.95, SDC: 0.69-1.03mm) end transversal (ICC: 0.84-0.87, SDC: 1.08-2.22mm).

Konklusion: Intra -og intertester reliabiliteten var meget god og bedre for målinger af den posterior-anterior senetykkelse longitudinal end transversal på symptomatiske og asymptomatiske sener. Reliabiliteten var også høj for målinger på tværsnitsbilleder, men befæstet med systematisk shift i måleresultaterne mellem testerne samt større målefejl. Patellarsenens tykkelse for symptomatisk og asymptomatiske sener bør måles i et longitudinal snit hvor målingerne forekommer mest reliabel. Gentagende målinger af senetykkelse på tværsnitsbilleder bør udføres af samme tester. Alligevel grundet få testere er resultat kun overførbart til indeværende testere.

Nøgleord: Tendinopati, Patellar tendinopati, senetykkelse, diagnostisk ultralyd, reliabilitet

Intra and inter-rater reliability of real time ultrasonography for the patellar tendon thickness in individuals with symptomatic and asymptomatic patellar tendons

Abstract

Background: RTUS is increasingly used as an available measuring instrument among physiotherapists for diagnosing superficial structures like tendons. Although RTUS is frequently used in scientific studies of patellar tendinopathy, few studies have reported on the reliability of RTUS measurements.

Purpose: To examine intra and inter-rater reliability for measurements of the patellar tendon thickness with ultrasonography, on individuals with symptomatic tendons and individuals with asymptomatic tendons.

Design: Cross-sectional design - a descriptive study of intra and inter-tester reliability.

Method: Sonographic images of 30 symptomatic and 33 non-symptomatic patellar tendons were included. Three raters with varying degrees of experience measured the tendons on longitudinal and transversal images. All three examiners participated in the inter-rater reliability analysis. Two examiners were included in the intra-rater reliability analysis.

Results: Intrarater reliability showed a higher reliability for measurements of symptomatic and asymptomatic tendons longitudinally (ICC: 0.88-0.96, SDC: 0.53-1.08mm), than transversally (ICC: 0.87-0.96, SDC: 0.55-2.00mm). The inter-rater reliability displayed a higher reliability for measurements on longitudinal images (ICC: 0.87-0.95, SDC: 0.69-1.03mm) than for transversal images (ICC: 0.84-0.87, SDC: 1.08-2.22mm).

Conclusion: Intra and inter-rater reliability was very high, and better for measurements of posterior-anterior tendon thickness longitudinally than transversally on symptomatic and asymptomatic tendons. Reliability was also high for measurements on cross-sectional images, but the measurements were related with a systematic shift between raters and generally subject to a greater margin of error. Measurements of the patellar tendon thickness should be made on longitudinal images where reliable measures are easier to obtain. Repeated measures of the tendon thickness in cross-sectional images should be performed by the same rater. However, because of the low number of raters in this study, the result can only be generalized to the raters included in this study.

Keywords: Tendinopathy, patellar tendinopathy, tendon thickness, reliability, ultrasonography

Indholdsfortegnelse

1. Introduktion.....	8
1.1. Baggrund.....	8
1.2. Tidligere forskning.....	11
1.3. Hensigt.....	14
1.4. Problemformulering.....	15
1.5. Opgavens afgrænsning.....	15
2. Den teoretiske baggrund.....	15
2.1 Studiets videnskabsteoretiske forankring.....	15
2.2. Patellarsenen.....	16
2.3. Patellar tendinopati.....	16
2.4. Real time ultrasonography (RTUS).....	19
2.5. RTUS som måleinstrument.....	23
2.5.1. Måleegenskaber ved målinger af senetykkelsen med RTUS.....	24
2.6. Reliabilitet.....	25
2.6.1 Variation og reliabilitet.....	26
2.6.2. Målefejl.....	27
2.6.3 Relativ og absolut reliabilitet.....	28
2.6.4 Relativ reliabilitet.....	29
2.6.5. Absolut reliabilitet.....	35
2.7. Opsummering.....	38
2.8. Kilder til variation.....	39
3. Metode.....	39
3.1. Design.....	39
3.2. Begrebsdefinitioner.....	41
3.2.1. Intratester reliabilitet:.....	41
3.2.2. Intertester reliabilitet.....	41
3.3. Udvalg.....	41
3.3.1. Rekruttering.....	42
3.3.2. Beskrivelse af udvalget.....	43
3.3.3. Variabler.....	45
3.4. Testere.....	45
3.4.1. Operationalisering af begreber og beskrivelse af testerne.....	46
3.4.2. Beskrivelse af testerne i studiet.....	47
3.5. Ethiske betragtninger.....	48
3.6. Dataindsamling.....	49

3.6.1. Dataindsamlingsfaser	49
3.6.2. Udstyr.....	52
3.6.3. Procedure	53
3.6.4. Procedure for ultralydsskanning	53
3.6.5. Procedure for måling af senerne.....	56
3.7. Databearbejdning	57
3.7.1. Relativ reliabilitet.....	58
3.7.2. Absolut reliabilitet.....	59
3.7.3. Normalfordeling.....	60
4. Resultat.....	62
4.1 Normalfordeling og variations homogenitet	63
4.2. Deskriptiv statistik og forskel mellem grupperne.	64
4.2.1. Demografiske variabler.....	64
4.2.2. Senetykkelsen.....	65
4.3. Intratester reliabilitet	65
4.3.1. Relativ reliabilitet.....	66
4.3.2. Absolut reliabilitet.....	67
4.4. Intertester reliabiliteten	68
4.4.1. Intertester reliabiliteten	68
4.4.2. Absolut reliabilitet.....	69
5. Diskussion.....	69
5.1. Resultat diskussion.....	70
5.1.1. Forskel i Senetykkelse.....	70
5.1.2. Intratester reliabilitet	71
5.1.3. Intertester reliabilitet	72
5.1.4. Diskussion af eksisterende litteratur	73
5.1.5. Opsummering.....	76
5.2. Valg af statistiske parametre	77
5.2.1. Relativ reliabilitet.....	77
5.2.2. Absolut reliabilitet.....	78
5.3. Metode diskussion.....	78
5.3.1. Udvalget	78
5.3.2. Testerne	80
5.3.3. Indsamling af skanningsbilleder.....	81
5.3.4. Målinger af patellarsenen.....	82
5.3.5 Opsummering.....	83

5.4. Kliniske implikationer.....	84
6. Konklusion	86
7. Litteraturliste	88

Figurindeks

Figur 1: Ultralydsbilleder af patellarsenen.....	20
Figur 2: Ultralydsbillede af patellarsenen.....	21
Figur 3: Ultralydsbillede med Power Doppler.....	22
Figur 4: Illustration af de måletekniske egenskaber for reliabilitet.....	29
Figur 5: Tidslinje for dataindsamlingsprocessen.....	49
Figur 6: Blinding af Tester A.....	51
Figur 7: Reliabilitetsprocessen.....	52
Figur 8: Ultralydsapparatet.....	53
Figur 9: Positioneringen af deltager for skanning.....	54
Figur 10: Ultralydsbilleder.....	55
Figur 11: Ultralydsbilleder.....	55
Figur 12: Måling af patellarsene tykkelsen på et longitudinal billede.....	56
Figur 13: Måling af patellarsene tykkelsen på et tværsnitsbillede.....	57
Figur 14: Flowdiagram for reliabilitetsundersøgelsen.....	62
Figur 15: Normal Q-Q plots.....	63

Tabel indeks

Tabel I: Måle karakteristika for RTUS.....	24
Tabel II: Inklusion og eksklusionskriterier.....	43
Tabel III: Den kliniske undersøgelse.....	45
Tabel IV: Beskrivelse af testerne.....	47
Tabel V: Frekvens af alder, demografisk statistik.....	64
Tabel VI: Demografisk statistik.....	64
Tabel VII: Forskel i senetykkelse mellem grupperne.....	65
Tabel VIII: Intratester reliabilitet symptomatiske sener.....	65
Tabel IX: Intratester reliabilitet asymptomatiske sener.....	66

Tabel X: Intertester reliabilitet symptomatiske sener.....	68
Tabel: XI: Intertester reliabilitet asymptomatiske sener	68

Bilag

Bilag 1: REK godkendelse	
Bilag 2: Informeret samtykke for deltagelse i projektet	
Bilag 3: Dataskema for den kliniske undersøgelse	
Bilag 4: Dataskema for registrering af målinger	
Bilag 5: Histogrammer, fordeling af data.	
Bilag 6: Scatterplots for illustration af forholdet mellem målingerne	

Forkortelser

PT	- Patellar tendinopati
RTUS	- Real timed ultrasonography
MR	- Magnetisk Resonans Skanning
CTT	- Classical test theory
ICC	- Intraclass correlation coefficient
SEM	- Standard error of measurement
SDC	- Smallest detectable change
SD	- Standard deviationen
CI	- Konfidensinterval
MS	- Mean square
ANOVA	- Analysis of Variance
Gns.	- Gennemsnit
LOA	- Limits of agreement

Intra –og intertester reliabilitet for ultralydsmålinger af patellarsenens tykkelse hos individer symptomatiske og asymptomatiske patellarsener

1. Introduktion

1.1. Baggrund

Tendinopati er en muskuloskeletal lidelse, der klinisk manifesterer sig ved belastningsrelateret smerte i sener ofte med ømhed ved palpation, lokal hævelse og nedsat funktions –og aktivitetsniveau (J L Cook & Purdam, 2009). Tendinopati er en hyppigt forekommende muskuloskeletal lidelse både blandt idrætsudøvere og ikke-idrætsudøvere. Tilstanden synes især at virke ind på rotatorcuffen, extensorsenerne tilhørende underarmen, achillessenen og patellarsenen. (Albers, Zwerver, Diercks, Dekker, & Van den Akker-Scheek, 2016; Clarsen et al., 2015; McCreesh & Lewis, 2013). Belastning har vist sig både at have en anabolsk og katabolsk effekt på sener, og overbelastning og/eller kompression af senestrukturen fremstår som en betydelig faktor i udvikling af tendinopati (Dean, Gettings, Dakin, & Carr, 2016; Drew, Smith, Littlewood, & Sturrock, 2014). Repeterende belastning der stiller store krav til seneapparatet synes at provokere en symptomatisk respons. Eksempelvis hos volleyballspilleren som gentagende gange udsætter patellarsenen for store kræfter igennem hop og landinger (Malliaras, Cook, Purdam, & Rio, 2015; Visnes & Bahr, 2013).

Patellarsenen er således udsat for tendinopati. Dette er speciel prævalent, i området ved senens proximale udspring fra den dybe pol af patella og betegnes typisk i litteraturen som Patellar tendinopati (PT) eller Jumpers knee (J. L. Cook, Khan, Kiss, Purdam, & Griffiths, 2000; Visnes, Tegnander, & Bahr, 2015).

PT er vist at forekomme både hos idrætsudøvere og ikke-idrætsudøvere (Albers et al., 2016; van Middelkoop, van Linschoten, Berger, Koes, & Bierma-Zeinstra, 2008). Alligevel fremstår PT specielt frekvent blandt idrætsudøvere. Især idrætsdiscipliner involverende eksplosive bevægelser med betydelig energiophobning i patellarsenen forekommer at påvirker prævalensen. Dette er dokumenteret i et tværsnitstudie (n = 613) blandt norske idrætsudøvere på eliteplan fra forskellige idrætsdiscipliner. Her fandt de en samlet prævalens af PT på 14, 2 %, hvor især disciplinerne volleyball og

basketball var iøjefaldende med en prævalens på henholdsvis 44,6% og 31,9% (Ø. B. Lian, Engebretsen, & Bahr, 2005). I et andet tværnsnitsstudie (n = 891) fandt Zwerver et al (2012) en samlet prævalens på 8,5% for ikke-elite idrætsudøvere i Holland. Ligeledes var der her en højere forekomst blandt idrætsdiscipliner implicerende repeterende energiophobning i patellarsenen. For volleyballspillere og håndboldspillere var forekomsten respektiv 14,4% og 13,3% (Zwerver, Bredeweg, & Akker-Scheek, 2011). Der er således rapporteret en varierende forekomst af PT blandt idrætsudøver med en prævalens op til 44,6% hos volleyballspillere på eliteplan. Dette støttes af Clarsen et al (2015). De fandt i tilknytning hertil, at overbelastningsskader i knæet hos volleyballspillere i Norge var den mest frekvente lokalisering for overbelastning med en forekomst på 36% (Clarsen et al., 2015). Udover at PT især blandt idrætsudøvere fremstår prævalent, kan tilstanden potentielt også medføre mere vidtgående begrænsninger. Det er rapporteret at op mod en tredjedel af idrætsudøvere under behandling for PT ikke kan returnere til idrætten på samme niveau. I tillæg er det berettet, at op mod 53% med PT er nødsaget til at lægge idrætten på hyllen (Kettunen, Kvist, Alanen, & Kujala, 2002; Malliaras et al., 2015). Således kan PT have mulige skadevirkninger udover smerte ved fysisk aktivitet.

Selvom belastning fremstår som en betydelig pato-ætiologisk komponent, er der en generel konsensus rundt udviklingen af tendinopati som multifaktoriel. Både ydre faktorer som belastning og indre faktorer som blandt andet køn, gener og fysiologiske parametre spiller ind (Abate et al., 2009; McAuliffe, McCreesh, Culloty, Purtill, & O'Sullivan, 2016; Visnes & Bahr, 2013). Alligevel kan den eksakte ætiologi siges fortsat at være relativ ukendt. Histologisk er der fortsat manglende konsensus rundt patogenesen, og flere teorier eksisterer og har været emne for megen diskussion. Termen tendinopati udspringer netop fra en ikke entydig sammenhæng vedrørende symptomer fra senen, som de klinisk manifesterer sig og opleves kontra de strukturelle og histologiske fund. Dermed er forholdet mellem tendinopati versus smerte og funktion endnu ikke helt stadfæstet (Abate et al., 2009; J L Cook, Rio, Purdam, & Docking, 2016; Dean et al., 2016; Drew et al., 2014).

Der eksisterer heller ingen gylden standard for diagnosticering af PT. Standarden for diagnosen PT er derfor en grundig anamnese og klinisk undersøgelse. Det kliniske billede er kendetegnet ved lokal aktivitetsrelateret smerte og palpationsømhed ved

patellarsenens proksimale udspring fra apex patella, og fremstår ofte som det diagnostiske kriterium i flere studier (Comin et al., 2013; J. L. Cook et al., 2000; Malliaras et al., 2015; Visnes et al., 2015). PT angives hyppigst at forekomme netop i den proksimale del af patellarsenen ved apex patella. Det støttes af flere kliniske studier, der i tillæg viser, at de strukturelle forandringer af patellarsenen på parakliniske undersøgelser ligeledes typisk forekommer i den proksimale del af patellarsenen (J. L. Cook et al., 2000; J L Cook, Khan, Kiss, Purdam, & Griffiths, 2001; Visnes et al., 2015). Derfor er palpationsømhed af patellarsenen lokalt ved apex patella en undersøgelsesteknik, der i flere muskuloskeletale lærebøger og studier indgår som et diagnostisk kriterium (Brukner & Khan, 2012, pp. 700-706; Comin et al., 2013; Visnes et al., 2015). Diagnosen PT er altså ikke udelukkende defineret ud fra den kliniske undersøgelse, men ud fra det samlede kliniske billede, hvorfor denne betragtes som standarden for identificeringen af PT (Malliaras et al., 2015).

Skønt den kliniske undersøgelse betragtes som standarden for diagnosticeringen af PT, har real timed ultrasonography (RTUS) etableret sig som et supplement i vurdering og identificering af strukturelle og patologiske parametre i patellarsenen. RTUS er en omkostningseffektiv, portabel og smertefri metode. Modsat er Magnetisk Resonans Skanning (MR) der er væsentlig dyrere at foretage og stationær. RTUS fremviser et godt anatomisk billede af patellarsenen uden fare for stråling af patienten. Samtidig er det en enkel og tidsbesparende metode i klinisk praksis. Typiske parametre som vurderes ved ultralydsskanning af patellarsenen er måling af strukturelle forandringer og hyperæmi. Til eksempel patellarsene tykkelsen (Docking, Ooi, & Connell, 2015; McAuliffe et al., 2016). Den teknologiske udvikling har forbedret billedkvaliteten betragtelig samt gjort RTUS mere anvendelig. Dette sammen med metodens praktiske egenskaber er muligvis årsagen til, at RTUS er begyndt at etablere sig hos fysioterapeuter i primærhelsetjenesten. Endelig er Ultralydsmålinger ikke mindst bredt anvendt indenfor seneforskning, når det kommer til diagnostik, klassificering og evaluering af intratendinøse forandringer (Bolvig, Fredberg, & Schifter Rasmussen, 2011, pp. 9-20; Ulrich Fredberg, Bolvig, & Andersen, 2008; McAuliffe et al., 2016).

Flere har taget til orde for, at RTUS i høj grad er afhængig af operatøren og kvaliteten på udstyret. Samtidig som RTUS fremlægges at have en lang læringskurve (Docking et al., 2015; U. Fredberg, Bolvig, Andersen, & Stengaard-Pedersen, 2008). RTUS

fremtræder dermed, som en målemetode der indsamler information, men som er under indflydelse af variation. Information om patellarsenens tykkelse er således under indflydelse af kilder til variation. Typisk systematisk og tilfældig målefejl der i forskellig grad kan influere på informationen. Eksempelvis systematisk variation mellem forskere eller fysioterapeuter. Siden information om patellarsenens tykkelse indgår som et element i informationen, der danner det kliniske billede tilknyttet tendinopati og hvorfra bestemmelser og vurderinger udgår, er det vigtigt, at målingerne er pålidelige. Målemetoder er et vigtigt element i systematiske vurderinger, beskrivelser og kommunikation af relevante fund hos patienter. For RTUS i et diagnostisk og et differentialdiagnostisk perspektiv er det blandt andet vigtigt, at kunne skelne mellem symptomatiske sener og asymptomatisk i relation til symptomer og differentialdiagnoser. Således er det essentielt at målinger med RTUS er reliabel og valide (Beyer, Magnusson, & Thorborg, 2012, pp. 11-21; Docking et al., 2015).

Til trods for at RTUS tillægges at være operatør og udstyrsafhængighed kombineret med en stadig angivelse af strukturelle parametre i forskning tilknyttet PT, er der få studier, der har undersøgt reliabiliteten af RTUS tilknyttet måling af patellarsenens tykkelse (Black, Cook, Kiss, & Smith, 2004; Sunding, Fahlström, Werner, Forssblad, & Willberg, 2016).

1.2. Tidligere forskning.

Fra et diagnostisk synspunkt kan det i tvivlstilfælde forekomme formålstjenlig at vurdere de intratendinøse forhold med RTUS. Dette med henblik på at skelne mellem symptomatiske og asymptomatiske sener i relation til mulige differentialdiagnoser. Samtidig er intratendinøse forandringer vist at forekomme i asymptomatiske sener. Ligeså er ingen forandringer vist at forekomme i symptomatiske sener (J. L. Cook et al., 2000; Docking et al., 2015) Det er således vigtig at kende reliabiliteten tilknyttet diagnosen PT, men også tilknyttet asymptomatiske sener.

RTUS-studier vedrørende forskning af PT refererer bestandig til senetykkelse, hypoekkoide områder og graden af hyperæmi. I tillæg tillægges RTUS at være betydelig afhængig af operatør-variation og kvaliteten på ultralydsapparatet (Bolvig et al., 2011; U. Fredberg et al., 2008). Således vil intra -og intertester reliabiliteten, ved måling af patellarsenens tykkelse formentlig kunne variere i henhold til oplæring og erfaring med

metoden samt kyndighed tilknyttet måleproceduren. Alligevel rapporteres mål for reliabilitet sjældent. Kun få studier har undersøgt reliabiliteten for måling af patellarsenens tykkelse tilknyttet individer med PT. Til indeværende forfatters kendskab er der kun to studier som har rapporteret reliabilitetsparametre tilknyttet klinisk diagnosticeret PT (Black et al., 2004; Sunding et al., 2016). Flere studier har undersøgt reliabiliteten tilknyttet friske individer. Disse er imidlertid baseret på forskellige måleprocedure, måledimensioner af patellarsenen og forskellige reliabilitetsparametre (U. Fredberg et al., 2008; Mc Auliffe, Mc Creesh, Purtill, & O'Sullivan, 2016; Skou & Aalkjaer, 2013).

Black et al (2004) så på intertester reliabiliteten for to radiologer med komparabel specialist kompetence (>20 år) inden RTUS. I studiet blev der udført målinger af patellarsenens tykkelse i et longitudinal og transversal plan. I alt indgik der 12 personer hvoraf otte personer havde kliniske diagnosticeret PT. De fandt en god reliabilitet for både det longitudinale plan (r 0,93, 95% CI 0,78 – 0,98) og det transversale plan (r 0,96 95% CI 0,88 – 0,99) kvantificeret i Pearsons korrelationskoefficient. Den gennemsnitlige forskel mellem testerne var ikke signifikant (0.8 ± 1.4 (SD), $p=0.07$) og 95% Limits of agreement (LOA) for longitudinale målinger var på -2.0mm – 3.6mm. For transversale målinger var den gennemsnitlige forskel ligeledes ikke signifikant (-0.4 ± 1.2 (SD), $p=0.27$) og LOA var på -2.8mm - 2.0mm. Selvom korrelationskoefficient angav en god relativ reliabilitet synes målefejlen betydelig. En øvre grænse på 3.6mm for at være 95% sikker på at målingen ikke skyldes målefejl forekom betragtelig (Black et al., 2004).

Et nyere studie undersøgte både intra -og intertester reliabiliteten for to biomedicinske teknikere og to læger med forskellig grad af kliniske erfaring og erfaring med RTUS. Der indgik i alt 26 sener, hvoraf 13 var med klinisk diagnosticeret PT. Reliabiliteten for hele undersøgelsen inklusiv skanning og måling blev udforsket tilknyttet de to mest erfarnde testere. Videre blev reliabiliteten undersøgt for alle de fire testere på sonografiske billeder indsamlet af de to mest erfarnde testere. De fandt gennemgående en meget god intra -og intertester reliabilitet ved måling af patellarsenens tykkelse i det longitudinale plan for begge procedure. Dette var gældende uafhængigt af erfaringsniveauet (ICC 0,983 95% CI 0.961-0.992 – ICC 0,996 95% CI 0.993-0.998). Ligeledes var coefficient of variation (CV) gennemgående lav (CV 4.0% – 1.6%),

hvilket indikerede en god reproducerbarhed for målemetoden i denne kontekst (Sunding et al., 2016). Imidlertid fandt Fredberg et al (2008) en større intratester variation (n=10 sener) hos individer med sene-forandringer indikerende en større målefejl. De rapporterede en CV mellem 7.7% og 14.1%, hvor intratester variation var mindst ved longitudinale målinger (7.7%).

For asymptomatiske individer fandt Fredberg et al (2008) i samme studie som ovenfor nævnte en acceptabel intratester variation hos en tester for målinger på longitudinale og transversale billeder (CV 16.7% - 21.2%). Intratester variationen var i dette studie mindre ved målinger longitudinal (CV 16.7-17.8%) end transversal (CV 16.8%-21.2%). En ændring af patellarsene tykkelsen under dette vil med stor sandsynlighed kunne betragtes som målefejl (U. Fredberg et al., 2008).

Skou & Aalkjaer (2013) fandt en varierende intra -og intertester reliabilitet for to tester. Intratester reliabiliteten viste en varierede ICC med en lav nedre grænse for relativ reliabilitet. Intertester reliabiliteten var ligeledes præget af stor variation og en lav nedre grænse (ICC 0.70, 95% CI 0.18-0.89). Den relative reliabilitet var derimod god for de to testere, når gennemsnittet af to målinger blev anvendt (ICC 0.94 95% CI 0.84-0.98 - ICC 0.89 95% 0.71-0.96) (Skou & Aalkjaer, 2013).

Litteraturen synes ofte at stadfæste RTUS som en operatør-afhængig metode med høj og lang læringskurve (Bianchi & Martinoli, 2007, pp. 71-75; U. Fredberg et al., 2008). I tillæg er flere reliabilitetsstudier omhandlende forskellige RTUS-målinger af patellarsenen udført af specialiserede operatører med lang oplæring og flere års erfaring (Black et al., 2004; U. Fredberg et al., 2008; Skou & Aalkjaer, 2013). I kontrast hertil har andre studier fundet god reliabilitet ved RTUS-målinger af velafgrænsede og veldefinerede anatomiske strukturer. Såsom sener og led udført af novicer i brugen af RTUS. Dette antyder, at udførelse af målinger med RTUS er mulig for novicer inden for en velafgrænset anatomisk struktur (Bjordal, Demmink, & Ljunggren, 2003; Ohrndorf et al., 2010; Sunding et al., 2016).

Anvendelsen af RTUS er ikke længere kun forbeholdt specialister, men ses nå også anvendt af fysioterapeuter i primærhelsetjenesten. RTUS er en relevant målemetode som et supplement til den kliniske undersøgelse. Metoden giver indsigt i vævstrukturen

samt graden af seneforandring. Samtidig kan RTUS vise sig nyttig diagnostisk ved verificering af PT samt skelnen mellem seneforandringer og eventuelle differentialdiagnoser (Docking et al., 2015; Malliaras et al., 2015). Brugen af RTUS kan også have sin relevans som en komponent blandt flere i en risikovurdering for udvikling af PT (McAuliffe et al., 2016). I tillæg synes RTUS at have en god sensitivitet og specificitet for verificering af klinisk diagnosticeret PT (Warden et al., 2007).

Studier omhandlende forskning af PT anvender i stor stil RTUS. Dette gælder både tilknyttet klassificering, diagnosticering samt relateret effektevaluering. Sideløbende er ultralydsmålinger sandsynligvis under indflydelse af operatør-variation og tilfældigvariation. Alligevel er det kun få studier, der har søgt at belyse intra -og intertester reliabiliteten for måling af patellarsenens tykkelse med RTUS relateret til PT. Begge med et begrænset udvalg (Black et al., 2004; Sunding et al., 2016). Ligeledes gør det sig gældende for patellarsenens tykkelse tilknyttet friske individer. Således er der behov for at producere mere viden omhandlende intra -og intertester reliabiliteten for ultralydsmålinger af patellarsenens tykkelse hos individer med symptomatiske sener og friske individer med asymptomatiske sener.

1.3. Hensigt

RTUS anvendes i stor grad indenfor forskning, diagnosticering og behandling og synes at være på fremmarch i den kliniske praksis blandt fysioterapeuter. Det er et nyttigt og værdifuldt instrument og skaber indsigt i de intratendinøse forhold specielt i et diagnostisk og klassificeringsperspektiv.

Patellarsenens tykkelse målt med RTUS vil være under indflydelse af variation, blandt andet tilknyttet den eller de testere som udføre undersøgelsen. Det er derfor af stor betydning at kende intra -og intertester reliabiliteten for ultralydsmålinger af patellarsenens tykkelse. Hensigten med dette studie er, at bestemme intra -og intertester reliabiliteten ved målinger af patellarsenens tykkelse på sonografiske billeder hos en gruppe individer med symptomatiske patellarsener og en gruppe friske individer med asymptomatiske patellarsener. Med udgangspunkt i dette tilrettelægges dette projekt med henblik på at skabe mere viden herom.

1.4. Problemformulering

På baggrund af formålet med dette projekt, opstilles følgende problemformulering:

1. Hvordan er intratester reliabiliteten ved ultralydsmålinger af tykkelsen til symptomatisk og asymptomatiske patellarsener
2. Hvordan er intertester reliabiliteten ved ultralydsmålinger af tykkelsen til symptomatiske og asymptomatiske patellarsener.

1.5. Opgavens afgrænsning

RTUS er en specifik paraklinisk målemetode, der kan indgå som et delelement i diagnosticering, klassificering og evaluering. Denne opgave er afgrænset til intra -og intertester reliabilitet for målinger af patellarsener på sonografiske billede. Intra -og intertester reliabilitet indbefatter således ikke hele undersøgelsesprocessen og dermed variation tilknyttet selve ultralydsskanningen inkluderet placering og brugen af transduceren. Hovedfokus i opgaven er intra -og intertester reliabilitet.

2. Den teoretiske baggrund

2.1 Studiets videnskabsteoretiske forankring

I problemstillingen i dette studie indgår UL som måleinstrument, hvor målingerne kvantificeres i objektive tal. Disse målinger analyseres for variation, hvilket igen kvantificeres i objektive tal. Fremgangsmåden bygger altså på den kvantitative metode. Den kvantitative metode favoriserer generaliserbar målinger af en objektiv virkelighed i søgen efter sammenhæng og årsagssammenhæng ofte med en vis grad af kontrol, for eliminering af subjektiv indflydelse på resultatet (Carter & Lubinsky, 2016, pp. 56-62) Den kvantitative forskning har en forankring i naturvidenskaben. Mere specifikt i relation til dette studiedesign en forankring i den medicinske videnskabstradition (Birkler, 2005).

RTUS vil i diagnosticering af PT kun fungere som en del af en helhed, nemlig som analyse på kropsniveau. Subjektive parametre som anamnese, smerte og vurdering af funktionsniveau er med til at sætte rammerne for inddragelsen af RTUS. Ligeledes er det ikke muligt at være helt objektiv i denne proces. Tolkning af billederne samt måling af patellarsenens tykkelse vil være baseret på den enkeltes baggrund og forforståelse.

Dermed kan vurderingen ikke forstås uden nogen form for subjektivitet, selvom dette ofte er forsøgt kontrolleret i et kvantitativ design. På trods af den kvantitative tilgang er det herfor vigtigt at erkende, at forskning altid er et resultat af interaktion mellem det erkendende subjekt og det erkendte objekt (Thornquist, 2003, p. 197).

2.2. Patellarsenen

Patellarsenen er også benævnt ligamentum patellae. Senen udspringer dels proksimalt fra den nedre patella pol, og dels er den en forsættelse af quadricepsenen. Patellarsenen hæfter til tuberositas tibiae. I et longitudinal snit er senen kegleformet med den tykkeste del proksimal. I et transversal snit er den fusiform med den midterste del som den tykkeste. Selve senen består af senefibre som er beklædt med en hinde kaldt peritendineum (Bojsen-Møller, Tranum-Jensen, & Simonsen, 2014, pp. 334-335).

2.3. Patellar tendinopati

Tendinopati indbefatter et bredt spekter af overbelastningsinducerede seneskader, der klinisk fremtræder med smerte og funktionsbegrænsninger uden senen er rumperet (J L Cook et al., 2016). Med baggrund i epidemiologiske studier er flere risikofaktorer identificeret tilhørende to store kategorier, ydre faktorer og indre faktorer. For PT er smerter i relation til mekanisk overbelastning ved idræt eller anden fysisk aktivitet dokumenteret som en grundlæggende ydre faktor. Hvilket bekræftes af den høje prævalens blandt idrætsudøvere. Typisk er der snak om aktiviteter med eksplosive bevægelser, hvilket indbefatter stor belastning i tillæg til høj hastighed. Kombinationen gør at der ophobes store mængder mekanisk energi i senen. Andre ydre faktorer som forkert fodtøj og vertikal hopevne kan muligvis medføre en forhøjet risiko (Ø. B. Lian et al., 2005; van der Worp et al., 2011; Visnes, Aandahl, & Bahr, 2013).

De indre faktorer er tilknyttet den intratendinøse patogenese som fortsat ikke er fuldt belyst. Samtidig er der i litteraturen ingen entydig sammenhæng mellem seneforandringer fremvist på RTUS og kliniske symptomer. Derfor skal begrebet tendinopati forstås som det kliniske billede, der beskriver smerte og dysfunktion uafhængig af den intratendinøse patogenese. Således er begrebet tendinopati en klinisk diagnose defineret ud fra symptomer og ikke histopatologiske eller parakliniske forandringer (Drew et al., 2014; McAuliffe et al., 2016).

De kliniske kardinalsymptomer for PT er lokale belastningsrelaterede smerter, palpationsømhed og funktionsbegrænsninger. Dette ses eksempelvis hos basketball -og volleyballspilleren med PT. Disse angiver typisk hop som en symptomprovokerende faktor netop med lokal smerte rundt apex patella og ofte palpationsømhed på den proksimale del af patellarsenen. Både klinisk, paraklinisk og histologisk er PT vist at forekomme hyppigst i den proksimale del af patellarsenen. Tilstanden kan dog også forekomme i quadricepsenen og distalt i patellarsenen (Malliaras et al., 2015; Visnes et al., 2015). Kendetegnende berettes symptomerne at starte gradvis under fysisk aktivitet eller i forbindelse med en specifik bevægelse. Smerterne kan forsvinde under aktiviteten for så at vende tilbage i slutningen eller efterfølgende (Malliaras et al., 2015).

Indre mekanismer relateret til seneforandringer er emne for megen diskussion. Der eksisterer mange teoretiske modeller som søger at forklare patoætiologien og patogenesen. Elementer i teorierne overlapper hinanden, og der findes argumenter for og i mod de forskellige teorier. Dette illustrerer blot kompleksiteten og den multifaktorielle natur ved seners respons på skade (Abate et al., 2009; J L Cook et al., 2016).

Når energi ophobes i patellarsenen, indgår det i et energieffektivitetssystem ved bevægelse, idet energi frigøres fra patellarsenen som gratis energi. Gentagende kraftig energioplagering i senen kan medføre en akut katabolsk adaptation som typisk resulterer i en forøget senetykkelse. Manglende restitution og tilpasset belastning kan føre til en vedvarende katabolsk effekt og senedeformation. Herved opstår mulig patologi i senen samt forandringer i senens mekaniske egenskaber (Abate et al., 2009). Dette vil på RTUS typisk fremstilles ved en fortykket sene, uorganiseret kollagen struktur og hyperæmi. Yderligere kan der forekomme degenerative seneforandringer. Degeneration er histopatologisk kendetegnet ved apoptose, en fortykket sene, betydelig heterogenitet af kollagen og den ekstra cellulære matrix samt neurovaskularisering (Abate et al., 2009; Docking et al., 2015).

Fremkomsten af patologiske forandringer i patellarsenen betegnes ofte ”tendinitis eller tendinose”. Grundet manglende konsensus om inflammationens og degenerationens rolle forekommer betegnelsen ”intratendinøse forandringer” eller ”seneforandringer” mere retvisende for den histopatologiske og parakliniske tilstand (Docking et al., 2015;

Drew et al., 2014). De fleste histopatologiske fund ved PT er vist at sammensvare med strukturelle og metaboliske seneforandringer på RTUS. Selvom der bestandig observeres seneforandring hos individer med PT, er der ikke et direkte forhold mellem seneforandringer og symptomer som smerter og dysfunktion (Dean et al., 2016; Docking et al., 2015; Khan et al., 1996).

Drew et al (2014) har i et systematisk review (n = 20 studier) netop set på årsagssammenhæng mellem seneforandringer og kliniske symptomer hos individer med PT. Mere konkret hvorvidt en symptombedring ved forskellige træningsmodaliteter kan forklares af seneforandringer på blandt andet RTUS. Udfra studiet fremtræder der ingen årsagssammenhæng mellem strukturelle seneforandringer og bedring af symptomerne tilknyttet excentrisk styrketræning og medicinsk træningsterapi. Derimod forekommer der moderat evidens (baseret på et RCT-studie) for en association mellem en reduktion af patellarsenens tykkelse på den ene side og en smertereduktion og funktionsforbedring på den anden side efterfulgt af tung-langsom styrketræning (Drew et al., 2014).

I tillæg har et prospektiv kohorte studie (n = 63) søgt at belyse den diagnostiske værdi af RTUS. I studiet er RTUS fundet at have god og større præcision end MR ved verificering af klinisk diagnosticeret PT. Seneforandringer på RTUS er i studiet fundet at have en betydelig sensitiv evne, til at identificere individer med klinisk diagnosticeret PT (sensitivitet 87%). Ligeledes er normale fund på RTUS vist at have en acceptabel evne til identificering af friske individer (specificitet 82%). Dertil kommer en fem til elleve gang større sandsynlighed for klinisk diagnosticeret PT ved fund af både seneforandringer og hyperæmi med RTUS (Warden et al., 2007). Dette studie som flere longitudinale -og tværsnitstudier har kontinuerlig fundet seneforandringer på RTUS hos individer med PT. Imidlertid viser studierne samtidig kontinuerligt seneforandringer hos asymptomatiske individer (Comin et al., 2013; Visnes et al., 2015; Warden et al., 2007). Cook et al (2000) rapporterede at 19% ud af 52 asymptomatiske patellarsener havde strukturelle forandringer på RTUS (J. L. Cook et al., 2000). Hertil er der i studierne fundet individer med PT uden seneforandringer på RTUS, dog i et mindre omfang.

McAuliffe et al (2016) har i systematisk review og meta-analyse set på hvorvidt RTUS kan prædikere udvikling af PT. Metaanalysen baseret på ni studier omhandlende PT demonstrerede at asymptomatiske seneforandringer ved baseline er tilknyttet en relativ

risiko for senere udvikling af PT (RR = 4.35, 95% CI 2.62 - 7.23). Således er der en mulighed for at asymptomatiske seneforandringer kan være en markør for tidlig asymptomatisk senepatologi, der med tiden udvikler smerter og dysfunktion (McAuliffe et al., 2016). Seneforandringer på RTUS kan herved være en mulig risikofaktor for udviklingen PT, men alligevel kun moderat set i lyset af de absolutte tal. 33 ud af 159 patellarsener i metaanalysen med asymptomatiske seneforandringer på RTUS udviklede symptomer, hvilket svare til 21%. Således udviklede en betragtelig del med asymptomatiske seneforandringer ikke PT. Andelen uden seneforandringer som udviklede PT er dog meget lav (McAuliffe et al., 2016). Alligevel kan seneforandringer på RTUS ikke stå alene og bør kun betragtes som en prædikator blandt flere i prædiktionen af PT (Docking et al., 2015; McAuliffe et al., 2016).

Med udgangspunkt i ovenforstående er der således associationer og dissociationer mellem seneforandringer på RTUS og symptomer som smerte og dysfunktion. Dette fremhæver den komplekse interaktion mellem seneforandringer, smerte og funktion. Hvorfor nogen individer med seneforandringer får smerte og nedsat funktion og andre ikke gør, forbliver et uafklaret spørgsmål (J L Cook et al., 2016).

Seneforandringer må som nævnt ikke misforstås med tendinopati, da tendinopati er afgrænset ud fra kliniske symptomer. Således har begrebsdifferentieringen mellem tendinopati og intratendinøse forandringer sin begrundelse. Ikke desto mindre repræsenterer de intratendinøse forhold på RTUS indsigt i vævsstrukturen og graden af patologi ved tendinopati samt som en indikator for mulig udviklingen af tendinopati. Dermed er tilstandene relevante og interessante at sammenstille (Docking et al., 2015; McAuliffe et al., 2016).

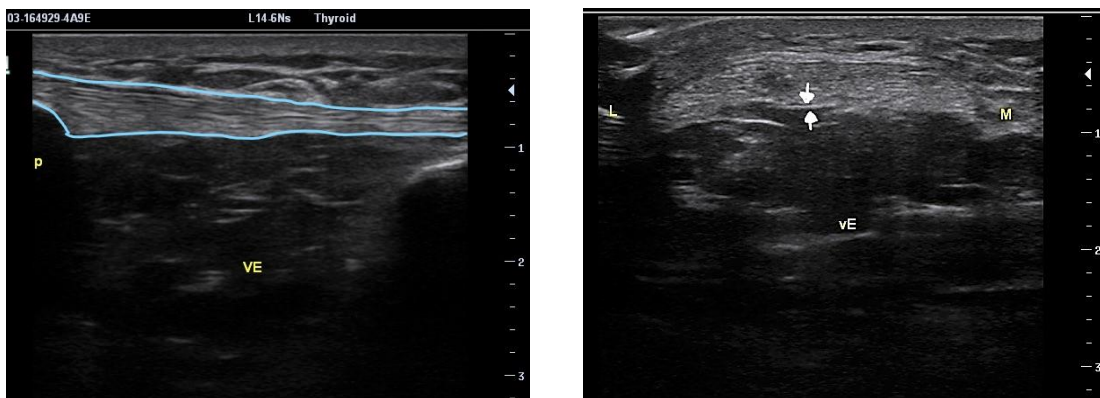
2.4. Real time ultrasonography (RTUS)

RTUS er en relevant og effektiv måde at fremstille vævsstrukturen og eventuel patologi, ved måling af sene tykkelsen og visualisering af vævstrukturen (Docking et al., 2015).

Ved RTUS omdannes elektriske signaler til lydbølger via transduceren, der videresender disse til vævet. Lydbølger reflekteres tilbage fra vævet til transduceren og ud fra pulsekkomålinger dannes der et elektrisk signal som resulterer i et billede. Frekvensen af lydbølgerne er direkte proportionel med billedkvaliteten. Dette medfører

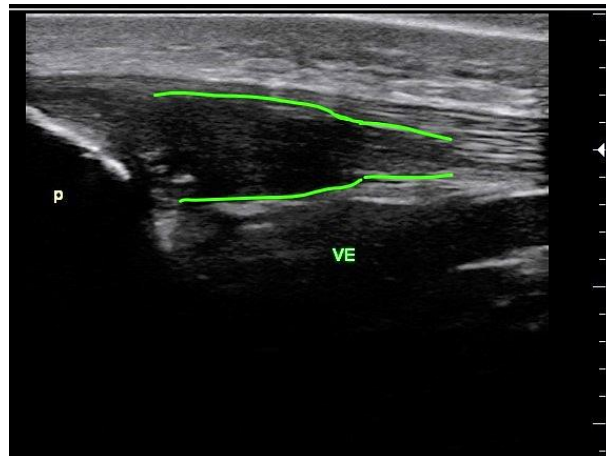
brug af høje frekvenser ved overfladiske strukturer og lave ved dybe strukturer. Patellarsenen er en relativ overfladisk struktur, hvorfor en høj frekvens på mellem 10-20 MHz er relevant (Bianchi & Martinoli, 2007, pp. 3-11). Senen er i et studie med 112 patellarsener vist at ligge superficiel. Gennemsnitlig lå patellarsenen 3.1(0.8) mm fra hudoverfladen. Studiet viste ingen signifikant korrelation mellem senedybde og demografiske variabler for patellarsenen (Bjordal et al., 2003). En frekvens på 10-12 MHz som anvendt i andre studier kan derfor antages som optimal til fremstilling af patellarsenen (U. Fredberg et al., 2008; Sunding et al., 2016).

Forskellige gråtoner på ultralydbillederne viser graden af ekko, hvilket illustreres ved de forskellige vævsstruktureres evne til at bryde lydbølger. Lyse gråtoner betegnes hyperekkoid og mørke gråtoner hypoekkoide. En normal patellarsene består primært af type I kollagen som parallelt anlagt og hierarkisk arrangeret i fibriller, fibre og fascikler. Disse kan visualiseres som en homogen og hyperekkoid struktur med fremstilling af parallelle kollagene fibre på fascikel niveau. Hele senen er beklædt med et tyndt bindevævslag, peritendineum, som afgrænser senen. Denne fremtræder på RTUS som et tynd hyperekkoid omslag kontinuerligt langs senen (Bojsen-Møller et al., 2014, p. 59; Bolvig et al., 2011, pp. 20-25). Til venstre på figur 1. ses et longitudinal snit af patellarsenen, og illustrerer de parallelle kollagene fibre samt patellarsenens kegleformede struktur med den tykkeste del ved patella. På figur 1. til højre vises et transversalsnit (tværsnit) af patellarsenen og de hvide pile markerer peritendineum.



Figur 1. Til venstre ses en normal patellarsene i et longitudinal plan. P markerer patella. Her fremstår senen kegleformet med hyperekkoid parallelle fibre. Det blå omrids markerer senens afgrænsning. Til højre ses et tværsnit af senen. Denne fremstår fusiform og hyperekkoid. Pile markerer peritendineum.

En fusiform eller diffus fortykket sene samt hypoekkoide områder er associeret med patologi og reflekteres som områder med mindre refleksion, der fremstår mørkere (U. Fredberg et al., 2008; McAuliffe et al., 2016). På figur 2. ses et longitudinal snit af en fortykket sene med hypoekkoide områder (mørke område) og vilkårlige arrangerede fibre.



Figur 2. Patellarsenen i et longitudinal snit. I den proximale del fremstår senen fortykket og hypoekkoide. Det grønne omrids afgrænser dette område.

På RTUS vil en fortykket sene synliggøres ved en måling af senens tykkelse. Senediameteren kan måles i en posterior til anterior retning. Intratendinøse forandringer manifesterer sig histopatologisk blandt andet ved ændret udseende og funktion af tenocytter samt en ændring af mindre proteoglykaner (proteiner) til større hydrofile proteoglykaner. Disse hydrofile proteiner binder væske bedre, hvorfor en fortykkelse af senen opstår og kan observeres på RTUS. Videre fremstår den kollagene struktur uorganiseret ved seneforandringer dels grundet et skift fra type I kollagen til type II og III kollagen. Dette udtrykkes på RTUS ved hypoekkoide områder. Hyperæmi eller neurovaskularisering kan også observeres med Color eller Power Doppler og er illustreret på figur 3. (Dean et al., 2016; Docking et al., 2015; Khan et al., 1996) Ovenfor nævnte seneforandringer ved RTUS viser sig ofte hos individer med symptomatiske sener. RTUS bidrager derfor med relevant diagnostisk indsigt i vævsstrukturen, og patologi ved PT.



Figur 3. Ultralydsskanning med Power Doppler. Det orangerøde område indikerer forhøjet vaskularitet.

Ved at udfører skanninger af patellarsenen i et longitudinal og transversal plan opnås et tredimensionelt indtryk af senen. Derved opnås et bedre overblik over den specifikke lokalisation til seneforandringerne end ved skanning i kun et plan. Teoretiske set bør målinger af patellarsenens tykkelse på RTUS i det longitudinale og transversale plan være ens. Patellarsenen er i midlertidig kegleformet og signifikant tykkere proksimal kontra distalt (U. Fredberg et al., 2008). Dette vil formentlig medføre en variation i målingerne alt efter hvor transduceren placeres. Siden symptomer oftest angives proksimalt i patellarsenen, og seneforandringer ligeledes er fundet hyppigst at forekommer her, refererer de fleste studier til målinger fra den proksimale del af patellarsenen (J L Cook et al., 2001; Visnes et al., 2015).

Grundet patellarsenens kegleform, forskellig standardisering og målepræferencer i litteraturen fremkommer der en stor difference mellem studier for den proksimale posterior-anterior patellarsene tykkelse. Gennemsnitsværdier for målinger af den proksimale tykkelse på longitudinale billeder hos individer med asymptomatiske sener fra forskellige populationer er rapporteret (3.6 ± 0.7 (SD) mm til 5.8 ± 0.8 (SD) mm) (Bjordal et al., 2003; Comin et al., 2013). Hos individer med symptomatiske sener er der rapporteret værdier mellem 6.2 ± 2.1 (SD) mm til 8.3 ± 2.2 (SD) mm (Kongsgaard et al., 2009; Sunding et al., 2016). Hertil er patellarsenen i flere studier fundet signifikant tykkere hos tendinopatiske individer sammenholdt med friske individer (O. Lian, Holen, Engebretsen, & Bahr, 1996; Sunding et al., 2016; Visnes et al., 2015).

Som følge af den højt strukturerede arkitektur af tætliggende og ovenpå liggende rækker af kollagen er patellarsenen stærkt ansiotropisk ved RTUS undersøgelser. Derfor er det essentielt, at transduceren er placeret med i en 90 graders vinkel til senen. Selv en mindre skævhed af vinklen kan medføre, at lydbølgerne bliver reflekteret delvist væk fra transduceren og dermed ikke registreres af transduceren. Dette kan medføre, at graden af ekko reduceres og strukturelle detaljer, som senens afgrænsning kan tilsløres. Faren for at senen fremstår isoekkoid (ligner omkringliggende struktur) med omkringliggende strukturer er muligvis større. En normal sene kan tilmed fejlagtig fremstå hypoekkoid og dermed patologisk. Dette kaldes ansiotropi, og er et så kaldt artefakt ved brugen af RTUS (Bianchi & Martinoli, 2007, pp. 73-74; U. Fredberg et al., 2008). Foreliggende er uønsket information, som kan lede til målefejl og tolkningsfejl. Således er det essentielt at lydbølgerne møder senen i 90 grader. Siden senefibrene har et let bølgeformet forløb i afslappet tilstand, bør senen positioneres let forlænget for at opnå en ret vinkel til senen (U. Fredberg et al., 2008).

Herved kan en hypoekkoid fortykket sene med desorganiseret fiberstruktur, i større grad antages at tilsløre senens afgrænsning sammenlignet med en normal hyperekkoid sene, hvilket vil stille større krav til skanning og tolkningen af sonografiske billeder. Dette indikerer, at processen er afhængig af operatørens håndtering af transduceren, anatomiske kundskaber og erfaring med tolkning af sonografiske billeder (Docking et al., 2015; U. Fredberg et al., 2008).

2.5. RTUS som måleinstrument

RTUS er en dimensionsspecifik/unidimensional målemetode, der i dette studie overordnet begribes som dimensionen senetykkelse målt i på longitudinale -og tværsnitsbilleder i en posterior-anterior dimension. Ud fra *International Classification of Function and Disabilities (ICF)* er RTUS en observerende og beskrivende målemetode af kropsfunktion og anatomi, der kommer til udtryk ved en digital kvantificering af patellarsenenstykkelse i millimeter på en ratioskala (Cieza & Stucki, 2008). RTUS som måleinstrument kan fejlagtig opfattes som objektiv, men målingerne og skanningerne er afhængige af øjnene som ser og målingerne anses derfor som subjektive. Målemetodens overordnede anvendelsesområde er som delelement i diagnosticering og klassificering samt vurdering af ændring over tid (McAuliffe et al., 2016). Dette principielt bestemt om behovet for viden er diagnose eller effekt relateret.

Imidlertid er dokumentation for dens anvendelse som evaluerende målemetoder begrænset. Alligevel anvendes RTUS som en del af flere evaluerende målemetoder i flere studie (Beyer et al., 2012, pp. 13-20; de Vet, Terwee, Mokkink, & Knol, 2011, pp. 7-13; Drew et al., 2014). I tabel I ses et overblik over karakteristika ved målingerne udført med målemetoden UL.

Tabel I: Karakteristika ved målinger med RTUS som målemetode

Måle karakteristika	RTUS Karakteristika
Unidimensional (<i>specifik</i>) eller multidimensional (<i>generel</i>)	Unidimensional (<i>specifik</i>)
For observerbare eller ikke-observerbare karakteristika	Observerbar
ICF	kropsfunktion og anatomi
Objektiv/subjektiv	Subjektiv
Anvendelsesområde	Diagnosticering og klassificering

Troværdige målemetoder er vigtige i systematiske vurderinger og beskrivelser af patienter. Måleværktøjet som anvendes må være præcis indenfor det, som skal måles. Videre hvordan målemetoden kan sammenlignes med andre målinger og hvordan resultaterne skal tolkes (de Vet et al., 2011, pp. 1-12). RTUS som en komponent i diagnosticering og klassificering af PT baseres som nævnt på målinger af patellarsenens dimensioner (Drew et al., 2014; McAuliffe et al., 2016). Som følge heraf er det essentielt, at målingerne af patellarsenens tykkelse er reliabel, hvis f.eks. tykkelsen af patellarsenen indgår som en del af et beslutningsgrundlag. Dette kan betegnes som en egenskab ved målingerne. Hvor målekarakteristika mere fortæller nået om målingernes væsen og formål, er måleegenskaber faktorer som informerer om hvordan en målemetode og tilhørende målinger fungerer (Carter & Lubinsky, 2016, pp. 240-241; Mokkink et al., 2010b).

2.5.1. Måleegenskaber ved målinger af senetykkelsen med RTUS

Reliabilitet har mange navne. *Consensus-based Standards for the selection of health Measurement Instruments (COSMIN)* har netop haft til formål at opnå konsensus omkring måleegenskaber, terminologi, definitioner og studiedesign samt hertil fortrukne statistiske metoder (Mokkink et al., 2010b). COSMIN er forankret i et Delphi studie

omhandlende en international konsensusproces. Studiet resulterede i konsensus vedrørende taksonomi, terminologi og definitioner rundt måleegenskaber særligt rettet mod patient -og helbredsrapporterende udfaldsmålinger. Herunder måleegenskaben reliabilitet. Reliabilitet som et overordnet begreb indeholder flere komponenter, der træder frem i relation til problemstillingen og konteksten målemetoden skal anvendes i (Mokkink et al., 2010a, 2010b). Er det til eksempel reliabiliteten blandt forskellige tester som ønskes belyst, træder intertester reliabiliteten frem.

2.6. Reliabilitet

COSMIN har defineret reliabilitet som et overordnet begreb: *"The degree to which the measurement is free from measurement error"* (Mokkink et al., 2010b). Videre har COSMIN udarbejdet en udvidet definition, relateret til kontekstualiseringen reliabiliteten beskriver og derved i tilknytning til intra -og intertester reliabilitet.

The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. [...] by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater). (Mokkink et al., 2010b).

Med dette udgangspunkt begriber intertester reliabilitet, hvorvidt fysioterapeuter der arbejder med RTUS vil opnå samme måleresultat ved to eller flere målinger. For eksempel ved målinger af patellarsenens tykkelse på et eller flere sonografiske billeder. Således om målingerne kan benyttes omvekslende blandt fysioterapeuter på en klinik, i et studie eller generelt blandt fysioterapeuter. Her repræsenterer forskellige fysioterapeuter, *different persons*, og ultralydsbillederne, *same occasion*. Falder interessen på intratester reliabilitet, går det således på om den enkelte fysioterapeut opnår samme måleresultat, hvis ultralydsbillederne måles ved to eller flere anledninger. Den enkelte fysioterapeut agerer her, *same persons*, og målinger ved forskellige anledninger, *different occasions*. Uanset er det grundlæggende at der søges viden om hvor træfsikker målingerne er og således i hvilken grad de er fri for målefejl (Beyer et al., 2012, pp. 43-45; de Vet et al., 2011, pp. 96-97).

I tillæg til at reliabilitet træder frem som forskellige komponenter, består reliabilitet som et overordnet begreb af to måletekniske egenskaber. Disse er af COSMIN benævnt reliabilitet og målefejl (Mokkink et al., 2010b).

2.6.1 Variation og reliabilitet

Målinger med RTUS udføres oftest af -og på mennesker som individer. Imidlertid synes enkeltindivider aldrig at være helt ens. Hverken testpersonerne som der måles på, eller testerne der udfører målingerne er fuldstændig homogene. De udviser en vis variation fra hinanden. Denne variation træder blandt andet frem i lyset, når flere testere eller samme tester gennemfører gentagende målinger af patellarsenens tykkelse. Derfor vil det ved gentagende ultralydsmålinger af patellarsenen, formodentlig ikke være realistisk at få eksakt samme måleresultat, der vil sandsynligvis være en variation. Der vil antageligvis være flere kilder til variation. Både deltager, testere og selve målemetoden kan agere flere kilder til variation og er blot eksempler blandt flere. Når det anføres at målinger i en bestemt kontekst er reliabel og anvendelig, accepteres der således en vis variation. Således træder variation frem i målinger og udfordrer graden af reliabilitet. Det er denne variation som søges belyst med statistik (Beyer et al., 2012, pp. 43-48; Lund & Røgind, 2016, pp. 9-11).

Reliabilitet og variation i tilknytning til kvantitative unidimensionale målinger er grundlæggende baseret på en af to teorier. Den ene kaldet classical test theory (CTT) og den anden generalizability theory. CTT forudsætter, at hver eneste måling (Y) består af to komponenter. En sand score (η) og en fejl komponent (ε), hvilket kan beskrives med følgende formel:

$$Y = \eta + \varepsilon$$

Den sande score (η) + Målefejl(ε) = observert måling (Y).

Dermed antager CTT, at hver enkelt person har en sand score for patellarsenens tykkelse. Da det ikke er mulig å kende en given persons sande score, må dette løses ved at indsamle gentagende målinger. Samtidig er en vigtig antagelse ved CTT at målefejl ikke korrelerer med den sande måling og heller ikke med hinanden. Dermed vil den totale variation (σ_y^2) ved gentagende målinger bestå af den sande variation (σ_s^2) og en variation relateret til målefejl (σ_e^2) (de Vet et al., 2011, pp. 98-101). Dette kan skrives som:

$$\sigma_y^2 = \sigma_s^2 + \sigma_e^2$$

Alt den variation som ikke kan tilskrives den sande variation mellem deltagerne må derfor i perspektiv af CTT skyldes målefejl (de Vet et al., 2011, pp. 19-20, 98-101). Måleegenskaben reliabilitet belyser således forholdet mellem den sande variation og den totale variation. COSMIN har defineret reliabilitet som en måleegenskab som følger: ”*The proportion of the total variance in the measurement which is because of “true” difference among patients*” (Mokkink et al., 2010b).

2.6.2. Målefejl

Målefejl er en anden, men beslægtet egenskab ved reliabilitet. Hvis man med udgangspunkt i CTT ser nærmere på målefejl, er målefejl jo et udtryk for variation. Her er der typisk to typer af variation som skal falde en for øje, systematisk variation (*systematisk målefejl/bias*) og tilfældig variation (*tilfældig fejl*). Systematisk fejl implicerer at målinger systematisk afviger fra hinanden i en positiv eller negativ retning (Atkinson & Nevill, 1998). Dette kan eksempelvis udforme sig som en systematisk forskel mellem to testere. Det kan være at tester A systematisk scorer patellarsenens tykkelse på ultralydsbilleder højere en tester B, hvilket til eksempel kan skyldes testernes udførelse af måleproceduren, erfaring, indlæring eller tolkning af sonografiske billeder (Beyer et al., 2012, pp. 48-51).

Tilfældige fejl kan skyldes, en iboende biologisk variation hos testdeltagerne, en mekanisk variation relateret måleinstrumentet i dette tilfælde ultralydsscanneren eller inhærent variation i testprotokollen. Dette kan føre til ustadige målinger. Kilder til variation i testprotokollen kan til en vis grad søges kontrolleret ved en mere stringent testprotokol, men generelt er det vanskelig at kontrollere tilfældig fejl grundet biologisk variation eller instrument variation (Atkinson & Nevill, 1998; Beyer et al., 2012, p. 48).

En målemetode vil som nævnt aldrig være helt fri for målefejl. Tværtimod vil testpersoner udvise variation. Således vil fysioterapeuter udvise variation uanset hvor homogene de måtte være, og specielt hvis en målemetode eller en testprotokol kræver at de i større eller mindre grad må analysere, vurdere og fatte en bestemmelse. Derfor når to eller flere fysioterapeuter udfører målinger på sonografiske billeder, vil en variation repræsenteret som en difference mellem de altså opstå (de Vet, Terwee, Knol, & Bouter, 2006). Målinger vil dermed i suppen af variation indeholde en kombination af variation. En del grundet i en systematisk forskel imellem deltagernes ”sande” målinger,

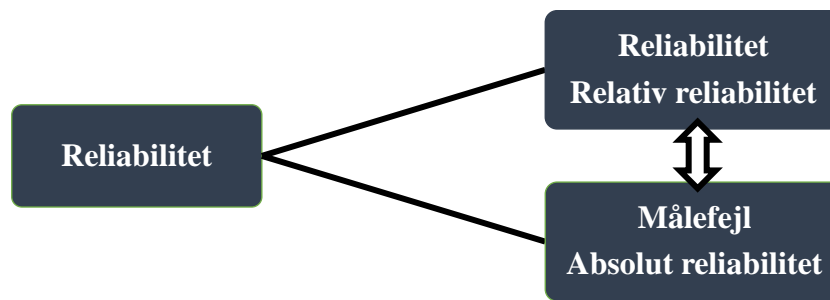
systematisk fejl og tilfældig fejl. Summen af de to fejlvariationskomponenter kan vi udtrykke som den totale fejlvariation. Da systematisk fejl ofte forsøges kontrolleret i en eller anden grad, vil tilfældig fejl typisk være større (Atkinson & Nevill, 1998).

COSMIN har defineret målefejl som: *"the systematic and random error of a patient's score that is not attributed to true change in the construct to be measured"* (Mokkink et al., 2010b).

2.6.3 Relativ og absolut reliabilitet

I det foregående ses det at reliabilitet overordnet omhandler i hvilken grad målinger er fri for målefejl. Hertil kommer som nævnt tidligere de måletekniske egenskaber reliabilitet og målefejl. Hvor egenskaben reliabilitet er relateret til, hvor godt testdeltager kan differentieres fra hinanden selvom målefejl er tilstede. Egenskaben målefejl er relateret til den absolutte målefejl af en persons score, der ikke kan tilskrives en reel ændring i den målte egenskab såsom senetykkelse (de Vet et al., 2006). Det forekommer derfor vigtig at supplere disse med hinanden. Dette netop fordi reliabilitet er relateret til variationen mellem testpersoner og målefejl til variation inden for den enkelte person. Således hvis variationen mellem personer er stor, vil målefejl i mindre grad påvirke evnen til at skelne testpersonerne fra hinanden. Derfor er det hensigtsmæssig at kende målefejlen (de Vet et al., 2006; Mokkink et al., 2010b).

Begge egenskaber bidrager herved måleteknisk med forskellig information om den samlede reliabilitet. I tilknytningen til statistikken bliver reliabilitet i association til dens måletekniske egenskaber typisk operationaliseret som relativ reliabilitet og absolut reliabilitet (Beyer et al., 2012, pp. 41-42; Mokkink et al., 2010b; Weir, 2005). Figur 4 illustrerer reliabiliteten og dets bestanddele.



Figur 4. Viser reliabilitet bestående af de beslægtede måletekniske egenskaber, reliabilitet og målefejl, der kan operationaliseres i relativ reliabilitet og absolut reliabilitet, der igen kan kvantificeres i statistiske parameter.

2.6.4 Relativ reliabilitet

I perspektiv af komponenterne intra -og intertester reliabilitet undersøger relativ reliabilitet, i hvilken grad forskellige testpersoner opretholder deres måleresultat for patellarsenens tykkelse ved gentagende målinger. Altså hvilken andel af den totale variation ved gentagende målinger som skyldes en sand forskel mellem testpersoner (de Vet et al., 2006; Mookink et al., 2010b). En måde at gøre dette på er at henvise til et ratioforhold for variationen mellem testpersoner og den totale variation typisk udtrykt i en korrelationskoefficient. Relativ reliabilitet kan derfor med baggrund i CTT beskrives som:

$$Reliabilitet = \frac{\text{Den sande variation mellem deltagerne}}{\text{Totale variation}}$$

Da målinger med RTUS forekommer i millimeter på en ratioskala kan relativ reliabilitet kvantificeres som en reliabilitetskoefficient mellem 0 og 1 også kaldet *Intraclass correlation coefficient (ICC)*, hvor 0 betyder ingen sammenhæng. En ICC på 1 indikerer en perfekt korrelation og jo mere ICC undviger fra 1 desto dårligere reliabilitet. Da den totale variation består af den sande variation mellem testpersoner og fejlvariation kan ICC formuleres på følgende måde:

$$ICC = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_\epsilon^2}$$

Imidlertid kan den totale fejlvariation (σ_{ε}^2) inddeles i variation mellem tester (σ_o^2) og random error ($\sigma_{residual}^2$) og her opstår et vigtigt spørgsmål. Er formålet kun at kvantificere sammenhæng mellem gentagende målinger eller at kvantificere sammenhæng og enighed mellem gentagende målinger. Således falder interessen på rangeringen af testpersoner, vil den systematiske forskel mellem testere ikke være relevant (de Vet et al., 2011, pp. 104-107; Streiner, Norman, & Cairney, 2015, pp. 167-168).

Da RTUS typisk er et diagnostisk måleinstrument, og interessen ligger i, hvorvidt testerne når til den samme konklusion og ikke rangeringen af testpersonerne, vil det typisk være relevant at inkludere enighed. I tillæg tillægges RTUS at være operatørafhængig, hvorfor det er nærliggende at ønske viden om testervariation (de Vet et al., 2011, p. 105; U. Fredberg et al., 2008).

Således kan ICC baseres på sammenhæng eller absolut samsvar. Dette vil videre blive omtalt med de engelske termer '*consistency*' og '*absolute agreement*' i relation til litteraturen og forståelsen af ICC formler (de Vet et al., 2006; McGraw, Wong, & Appelbaum, 1996). ICC-consistency og ICC- absolute agreement kan defineres ud fra følgende formler:

$$ICC_{consistency} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{residual}^2}$$

Ved ICC-consistency indeholder den totale fejlvariation kun random error. $\sigma_{\varepsilon}^2 = \sigma_{residual}^2$.

$$ICC_{absolute\ agreement} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_o^2 + \sigma_{residual}^2}$$

For ICC- absolute agreement indbefatter den totale fejlvariation både den systematiske forskel mellem tester og random error. $\sigma_{\varepsilon}^2 = \sigma_o^2 + \sigma_{residual}^2$.

Foregående skal ikke forstås som ICC i sin eneste form, men som en af flere former. Der er således forskellige versioner af ICC alt efter hvilke spørgsmål

problemformulering stiller, studie design og forudsætningerne som stilles. Shrout og Fleiss (1979) har præsenteret system med seks former af ICC og efterfølgende har McGraw og Wong (1996) tilføjet yderligere fire former, hvorfor der ved brug af deres terminologi kan siges at være 10 former af ICC. McGraw og Wong (1996) har tilføjet systemet elementerne consistency og absolute agreement og en del statistisk software bygger på dette. Fundamentalt set kan der overordnet fremholdes tre modeller. Dette ICC1.1, ICC2.1 og ICC3.1 som igen kan være baseret på gennemsnittet af k antal målinger. I tillæg kan ICC2.1 og ICC3.1 inkorporere absolute agreement eller kun consistency og således adskille effekten af testervariationen fra den totale fejlvariation (McGraw et al., 1996; Shrout & Fleiss, 1979; Weir, 2005).

Videre i dette studie vil terminologien fra disse forfattere kombineret referere til de forskellige ICC formler jævnfør Streiner et al (2015). Ovenforstående stående definition af ICC vil dermed refereres til som ICC2(A,1), hvor 2 står for ICC 'model 2', A for 'absolute agreement' og 1 fordi der indgår måling ($1 = k$) fra en tester og ikke er gennemsnit af k målinger. ICC2(A,1) og ICC3(A,1) er udregningsmæssigt ens, men tolkningsmæssigt forskellig, hvilket vil blive belyst senere (McGraw et al., 1996; Shrout & Fleiss, 1979; Streiner et al., 2015, p. 167).

Før de statistiske parametre udvælges er det nødvendig at gøre sig nogle tanker om studiets hensigt, design og forudsætninger relateret til datatype. Derefter er der tilknyttet ICC, grundlæggende tre bestemmelser som må fattes. Skal model et, to eller tre benyttes, skal consistency eller absolute agreement kvantificeres, er testerne en tilfældig eller kendt faktor og er variationen baseret på et sæt af målinger eller gennemsnittet af k målinger. Modellerne er tilknyttet mean square (MS) værdier fra Analysis of varians (ANOVA) (McGraw et al., 1996).

Går man dybere kan følgende spørgsmål stilles relateret til Shrout og Fleiss (1979) og McGraw og Wong's (1996) artikler om ICC:

1. Er en envejs eller tovejs ANOVA passende? Dette baserer sig grundlæggende på studiedesignet. Er alle testerne krydset med alle testdeltagerne, hvis så kan systematisk fejl separeres fra fejlkomponenten og en tovejs-model er at foretrække. Hvis studiedesignet ikke tillader at testerne krydses med alle

testdeltagerne kan effekten af testerne ikke separeres fra fejlkomponenten. Altså tillades alle tester ikke at score alle testdeltagerne. En envejs-model (ICC 1.1) er da betinget. Alle kilder til fejl er i denne model klumpet sammen i en fejlkomponent. Det eneste du mister ved denne model er evnen til at adskille tilfældig variation fra tester variation (McGraw et al., 1996; Weir, 2005). I et metodestudie hvor det er ønskeligt at belyse intra -og intertester reliabiliteten, fremstår det imidlertid nærliggende med et krydset design, med det formål at få information om effekten testerne har på reliabiliteten.

2. Kan systematiske fejl ignoreres eller ønskes det at inkludere absolute agreement? Ønskes der viden om i hvilken grad testervariation spiller ind på den totale variation inkorporeres absolute agreement. Dette er ofte ønskeligt tilknyttet et diagnostisk måleredskab, da enighed typisk er relevant (de Vet et al., 2011, pp. 104-105; McGraw et al., 1996).
3. Skal der generaliseres til alle tester i testerpopulationen (tester random factor) eller er testerne i studiet de eneste af interesse (tester fixed factor)? Er formålet at generalisere til populationen af testere anvendes model ICC1.1 eller ICC2.1. Her regnes testerne som en tilfældig faktor der repræsentativ for testerpopulationen. Hvis derimod generalisering kun omhandler testere inden for det enkelt studie anvendes ICC3.1, idet testerne her ikke anses som repræsentativ for hele testpopulationen. Flere studie omhandlende intra -og intertester reliabilitet anfører til tider, at reliabiliteten kan generaliseres udover det aktuelle studie. Dette forankret på to eller tre testere (Black et al., 2004; Skou & Aalkjaer, 2013; Sunding et al., 2016). Imidlertid synes det insufficient, at to tester skal kunne repræsentere variationen i en hel population af testere (McGraw et al., 1996; Van Genderen, De Bie, Helders, & Van Meeteren, 2003).
4. Skal reliabiliteten estimeres baseret på en enkelt måling eller gennemsnittet af flere målinger? Gennemsnit vil med baggrund i matematikken altid producere højere ICC-værdier. Viser reliabiliteten sig at være dårlig kan gennemsnittet af flere målinger være en af flere løsninger for at forbedre denne. Men i ud fra et klinisk perspektiv vil dette kræve mere tid og mere data at administrere. (McGraw et al., 1996).

Forvirring rundt de enkelte ICC modeller kan hurtig opstå. Anvendes Statistical Package for Social Science (SPSS) i dataanalysen bør man være opmærksom på at denne definitions-mæssigt læner sig op af McGraw and Wong (1996). Hertil er det netop essentielt, at forskellen mellem ICC2.1 og ICC3.1 kun er per definition. Begge modeller beskrives som følger inkluderende på absolute agreement:

$$ICC2(A, 1) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_o^2 + \sigma_{residual}^2}$$

$$ICC3(A, 1) = \frac{\sigma_s^2}{\sigma_s^2 + \theta_o^2 + \sigma_{residual}^2}$$

Her ses det at ICC3 (A,1) næsten er identisk til ICC2 (A,1) og begge inkorporerende absolute agreement. Den eneste forskel er θ_o^2 mod σ_o^2 , hvilket kun illustrere definitionsforskellen relateret til tolkning af de to modeller. Skal der generaliseres udover studiet eller ikke. Selve kalkulering af ICC formlerne er den samme og ICC-værdien vil ligeledes blive den samme. Dette illustreres ved udregningen, hvor der for begge anvendes samme ligning bestående af MS værdier høstet fra ANOVA:

$$ICC2(A, 1) = \frac{MS_R - MS_\epsilon}{MS_R + (k - 1)MS_\epsilon + \frac{k}{n}(MS_C - MS_\epsilon)}$$

$$ICC3(A, 1) = \frac{MS_R - MS_\epsilon}{MS_R + (k - 1)MS_\epsilon + \frac{k}{n}(MS_C - MS_\epsilon)}$$

Således tillades det at indlemme absolute agreement i ICC3, hvilket ikke er muligt i Shrout og Fleiss (1979) definition af ICC3. Dermed er der grundlæggende ikke ændret på beregninger fra Shrout og Fleiss (1979). Det essentielle her er tolkningen af ICC ‘en. McGraw og Wong (1996) har udtrykt det på følgende måde: *“the random-fixed effects distinction is in its effect on the interpretation, but not calculation, of an ICC”*. (McGraw et al., 1996, p. 37) Altså er formålet at generalisere til populationen af tester anvendes ICC2.1. Gælder reliabilitet kun for testerne i studiet anvendes ICC 3.1

Med udgangspunkt i ICC formlerne og hvorvidt de inkorporerer absolute agreement eller kun consistency er afgørende for kvantificering af de enkelte ICC formler. Hvis effekten af denne variationskomponent er begrænset vil ICC-værdierne fremstå tilnærmelsesvis ens. Altså vil forskellen mellem modeller som indregner absolute agreement, og de som kun ser på consistency være ens eller tæt på. Jo større effekt den systematisk variation har, jo mere forskel vil komme til udtryk i ICC-værdierne. Specielt mellem ICC1,1, ICC2 og 3 (A,1) på den ene side og ICC2 og 3 (C,1) på den anden side (Streiner et al., 2015, pp. 167-169; Weir, 2005).

Hvorvidt den forekommer en systematisk fejl mellem testerne kan udledes ved at sammenligne ICC2(C,1) med ICC2(A,1), ICC3(C,1) med ICC3(A,1) og ICC1.1 med ICC2 og 3(C.1). Imidlertid fås der kun en relativ værdi for fejlkomponenten og ikke den aktuelle værdi. Siden varianskomponenterne kommer fra en analysis of variance (ANOVA) kan den aktuelle værdi hentes herfra mean square (MS). I dette tilfælde tester mean square. Samtidig kan den tilhørende *F*-test vise om denne er signifikant. ICC kan nemlig godt indikere, at reliabiliteten er god, selvom der er en signifikant systematisk fejl. ICC er nemlig afhængig af bredden af målingerne i udvalget. Det vil sige, hvis variation mellem testpersonernes patellarsene tykkelse er stor, er det muligt at selv en betydelig målefejl ikke tilslører forskellen mellem testpersonerne og ICC-værdien bliver høj (de Vet et al., 2006). Således er en analyse som kun fokuserer på ICC og ikke varianskomponenterne eller målefejlen ikke fyldestgørende, idet man risikerer at konkludere at reliabiliteten er god, selvom den i virkeligheden er befængt med betydelig målefejl (Weir, 2005).

ICC er en reliabilitetskoefficient der som en dimensionsløs ratio udtrykker graden af variation forårsaget af forskel mellem testpersoner i forhold til den totale variation. Problemet med en dimensionsløs ratio er at det er vanskeligt at tolke dens betydning relateret til en enkelt måling. Således er det en relativ kvantificering af reliabilitet. Siden ICC kan inkludere absolute agreement, kan der stilles spørgsmål ved om parametre for målefejl er nødvendige. Det er jo muligt at uddrive værdier for målefejl ($\sigma_0^2 + \sigma_{\text{residual}}^2$) fra ICC formlerne. Selvom en ICC tager højde for absolut agreement, så har den ingen egentlig måleenhed og er derfor stadig et relativ udtryk for reliabiliteten. ICC foreviser herved ingen absolut indikation for hvor meget målefejl, der kan forventes i millimeter

eller kilo fra måling til måling. Videre hører det til sjældenheder at varianskomponenterne er præsenteret i et studie. Det sker, at det kun er ICC værdien og modelnummer som er frembragt. Information om formelen indbefatter absolut agreement eller kun consistency er ofte manglende også i studier omhandlende RTUS og PT (Sunding et al., 2016). Varsomhed rundt tolkningen af ICC-værdier alene er derfor vigtig og indikator for den absolutte reliabilitet bør fremvises. (de Vet et al., 2006; Weir, 2005).

En ICC på 1,0 indikerer perfekt reliabilitet, og spørgsmålet bliver derfor, hvor stor en afvigelse fra 1 kan accepteres. ICC graderes ofte i dårlig, moderat og god reliabilitet. hyppigt angives værdier over 0.75 som høj reliabilitet (Lindahl & Juhl, 2010, p. 195). Imidlertid er ICC-værdien afhængig af heterogeniteten i udvalget, og forskellige ICC-formler vil kunne resultere i forskellige ICC-værdier på lidt forskellig grundlag (Streiner et al., 2015, p. 183). Derfor bør graderingen af ICC tolkes udfra hvad målemetoden skal bruges til. Måling af livstruende sygdomme vil selvsagt kræve, at ICC værdien er højere end 0.75. Måske så høje som over 0.95. Således er det vanskelig at vurdere hvad, der kan accepteres som god reliabilitet tilknyttet målinger på symptomatiske og asymptomatiske sener. Accepteres en ICC på 0.75 accepteres det samtidig at 25% skyldes fejlvariation (Carter & Lubinsky, 2016, pp. 243-244).

Alligevel synes der at være et grundlag for at kræve højere ICC-værdier for en metode, der skal anvendes i klinikken på individniveau end for målinger på gruppeniveau. Typisk fordi målingen i klinikken lægger op til bestemmelser rundt den enkelt patient. Derfor ønsker klinikerne at have stor sikkerhed rundt sin beslutning, specielt hvis der måles på en livstruende sygdom. Modsat er det på gruppeniveau for eksempel i forskning. Her kan en mindre ICC-værdi accepteres. Videre vil det, i kliniske randomiserede studier forventes, at kilder til fejl fordeler sig ligeligt mellem interventionsgruppen og kontrolgruppen (Beyer et al., 2012, pp. 68-69; Carter & Lubinsky, 2016, pp. 243-244).

2.6.5. Absolut reliabilitet

Grundlæggende kan det siges at RTUS søger at diskriminere mellem patienter. Her synes den relativ reliabilitet primært at være indikeret. Dette eftersom variationen forårsaget af målefejl bør være begrænset i forhold til variationen mellem de personer målemetoden

skal skelne imellem. Imidlertid er ICC betydelig afhængig af variationsbredden i det udvalg den analyseres i og har en begrænset evne til at opdage systematiske fejl. SEM på den anden side er ikke afhængig af variation mellem testpersoner og den giver en absolut størrelsesorden for målefejl. Samtidig anvendes RTUS til evaluering i studier og dermed indenfor forskning (Drew et al., 2014). Her er det vigtigt med en kvantificering af målefejlen for at kunne bestemme om en målt ændring kan tillægges målefejl eller en virkelig ændring (Weir, 2005).

Den absolutte reliabilitet er således relateret til målefejl og kan kvantificeres på flere måder. Absolut reliabilitet angiver en dimensionsspecifik værdi for målefejlen. Det vil sige den angives i samme måleenhed som testen. *Standard error of measurement (SEM)* er et parameter for målefejl. SEM angives i samme måleenhed som målemetoden og er således en kvantificering af absolut reliabilitet. SEM fortæller noget om variation af gentagende målinger inden for et individ eksempelvis gentagende målinger med RTUS af en testpersons tykkelse af patellarsenen. Dermed er det variationen hos det enkelt individ (*within-subject variation*) som estimeres. Variation af den enkelte testdeltagers målinger kan enkelt beskrives med standard afvigelsen af målingerne (*SD*). Er SD estimeret med baggrund i variationen som udvises når en testdeltager udsættes for gentagende målinger, vil denne SD repræsentere SEM. SEM er således SD rundt en enkelt måling og hermed et mål for hvor langt fra hinanden gentagende målinger er, eller omvendt for hvor enige de er (de Vet et al., 2006; Hopkins, 2000).

SEM kan udregnes på flere måder. Igen kan SEM som ICC inkludere den systematisk fejl eller kun tilfældig fejl. Hvis ICC-værdien skal indgå i udregningen, er det essentielt at stadfæste, om den er baseret på consistency eller absolute agreement. Dette påvirker dermed størrelsen på SEM specielt hvis den systematiske variation er stor. SEM kan udledes fra den totale varianskomponent σ_{ϵ}^2 som indgår i ICC formelen og kan generelt beskrives således:

$$SEM = \sqrt{\sigma_{\epsilon}^2}$$

Fra tidligere står det klart at den σ_{ϵ}^2 inkluderet systematisk fejl og tilfældig fejl er lig $\sigma_{\epsilon}^2 = \sigma_o^2 + \sigma_{residual}^2$. Dermed kan SEM agreement beskrives som følger:

$$SEM_{agreement} = \sqrt{(\sigma_o^2 + \sigma_{residual}^2)}$$

Alle varianskomponenterne kan som nævnt uddrives fra MS i en ANOVA tabel.

På trods af at SEM er velegnet til at belyse den gennemsnitlige målevariation for af patellarsenens tykkelse, så genspejler den gennemsnitlige målevariation ikke nødvendigvis målevariation på alle scoreniveauer på skalaen af målinger. Målevariation kan være forskellige for de som har høje score og de som har lave. Er dette tilfældet er dataene heteroscedastiske. En forudsætning for at SEM er at dataene er homoscedastiske for at SEM afspejler den gennemsnitlige målevariation over hele skalaen af målinger (Atkinson & Nevill, 1998; Beyer et al., 2012, pp. 62-63).

SEM er egentlig SD om målefejl og når målefejl antages at være normalfordelte samt med viden om z-score er der 68% sandsynlighed for at en testpersons sande måling ligger indenfor en SEM eller x-antal millimeter, kilo el.lign. Videre vil der være cirka 95% sandsynlighed for at en testpersons score falder inden for 1,96 SEM (Carter & Lubinsky, 2016, pp. 237-239). Det vil sige hvis der foretages en enkelt måling og den overstiger en SEM er vi 68% sikker på at der er tale om en sand forskel. Foretages der en måling af patellarsenens tykkelse og på et senere tidspunkt yderlig en måling som er lavere er historie med konfidensinterval en lidt anden. Nu er der to målinger fra en person. Fortsat er det ønskelig at vide om ændringen er indenfor det man ville forvente skyldes målefejl. *Smallest detectable change (SDC)* udtrykker dette (Weir, 2005). SEM anvendes her til at bestemme SDC ud fra nedenstående formel:

$$SDC_{Individ} = SEM \times 1.96 \times \sqrt{2}$$

Der er altså to målinger fra samme person. Hver af disse målinger har en sand komponent og en fejl komponent. Derfor må der tages højde for komponenterne forbundet med begge målinger. Det gøres ved SD for forskellen mellem de to målinger som er $\sqrt{2} \times SEM$. De 1.96 er z-scoren for 95% konfidensinterval og resultatet bliver en SDC der med 95% sikkerhed angiver ændringen på senetykkelse er stor nok til ikke at kunne tilskrives målefejl. Et konfidensinterval på 95% synes at være ganske konservativ. Målevariation på individniveau har typisk en større variation end på

gruppeniveau og med et konfidensinterval på 95% er der fare for at kliniske relevante ændringer vurderes som målefejl og dermed risiko for at begå en type II fejl. Derfor hvis en mere liberal tilgang ønskes kan man med viden om z-score ændre denne så konfidensinterval eksempelvis bliver 80% (Streiner et al., 2015, pp. 181-183; Weir, 2005). Denne definition af SDC beregner målevariation på individniveau. Ved at dividere den individuelle målvariation med \sqrt{n} , hvor n er antal testpersoner, fås målevariationen på gruppe niveau (de Vet et al., 2011, pp. 244-245).

2.7. Opsummering

Grundlæggende er der to vigtige spørgsmål, der vedrører reliabilitet som et overordnet begreb. Hvor godt kan vi skelne testpersoner fra hinanden til trods for tilstedeværelsen af målefejl? Spørgsmålet vedrører hvorvidt måleinstrumentet i en given kontekst er reliabel. Hvor enige er testerne mellem gentagende målinger? Dette omhandler målefejl og fortæller noget om, hvor ens målingerne er ved gentagende målinger. Samlet udfylder de begrebet reliabilitet (de Vet et al., 2006; Mokkink et al., 2010b).

I ICC indgår variationen mellem testpersoner og derfor er ICC påvirket af heterogeniteten i udvalget. Således er det vigtigt, at være opmærksom på at en høj ICC-værdi kan skyldes en høj SD, og at dette kan skjule betydelige målefejl. Dette understreger vigtigheden i at kvantificere målefejlen sammen med den relative reliabilitet. Reliabilitet er ikke uafhængig af tid og sted, men derimod stærk påvirket af populationen og konteksten. Reliabilitet referer til målingerne hentet fra en given population, opnået på en bestemt måde med et måleinstrument (Streiner et al., 2015, pp. 159-164).

På hvilken måde relativ og absolut reliabilitet kan kvantificeres på, er blandet andet afhængig af datatype. ICC, SEM og SDC er parametrisk statistik og bygger på de antagelser, at datamaterialet er normalfordelt og homoscedastisk (Atkinson & Nevill, 1998).

2.8. Kilder til variation

I forbindelse med en reliabilitetsundersøgelse er det vigtigt at analysere kilder til fejl. Kilder som åbenbart kan føre til systematisk fejl bør søges kontrolleret. For RTUS kan kilder til fejl være manglende retningslinjer for hvor senen skal måles. Det kan ved skanning være forskellig placering af transduceren. Kilder til variation søges typisk kontrolleret ved at nedfælde en forsøgsprotokol. Denne baserer sig på en grad af standardisering. Graden af standardisering bestemmes ud fra, hvad målemetoden skal anvendes til, konteksten samt om det ønskes at generalisere til andre tester. Målemetoder tilknyttet alvorlig patologi vil ofte kræve at så mange kilder til variation som mulig elimineres. Dette ved en høj grad af standardisering. Samtidig kan et unøjagtigt måleinstrument kræve, at der udføres en høj grad af standardisering. For målemetoder som ønskes anvendt i klinisk praksis eller generelt fremstå overførbare til andre baseres typisk på en delvis standardiseret tilgang (Carter & Lubinsky, 2016, pp. 250-251).

3. Metode

I dette afsnit tydeliggøres og begrundes valget af forskningsdesign. Videre beskrives valg af materiale og der redegøres for etiske overvejelser. Endelig redegøres der for den metodisk fremgangsmåde for dataindsamling –og bearbejdning.

3.1. Design

Jævnfør problemstillingen søger studiet at undersøge intra -og intertester reliabiliteten. Undersøgelsen er således en beskrivende undersøgelse af intra -og intertester reliabilitet for målinger af den posterior-anterior tykkelse af patellarsenen på indsamlede sonografiske billeder fra en gruppe individer med PT (symptomatiske sener) og en gruppe friske individer (asymptomatiske sener) udført af tre fysioterapeuter. Studiet kan også kaldes et kvantitativt metodestudie, idet hensigten er at dokumentere intra -og intertester reliabiliteten (Carter & Lubinsky, 2016, pp. 249-253).

Selve designet er et tværsnitsdesign, da der indsamles information om intra -og intertester reliabiliteten fra en afgrænset gruppe af deltagere og testere på en standardiseret måde tilknyttet et specifikt tidspunkt (Carter & Lubinsky, 2016, p. 189). Studiet vil derfor blive refereret til som en tværsnitsundersøgelse af intra -og intertester reliabiliteten.

Undersøgelse må således inkorporere reliabilitetskomponenterne intra -og intertester reliabilitet. Disse kan belyses med udgangspunkt i definition af reliabilitetskomponenter af COSMIN:

The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. [...] by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater). (Mokkink et al., 2010b)

Studiedesignet må således begribe intertester reliabiliteten som hvorvidt testerne i studiet vil opnå samme måleresultat ved to eller flere målinger. Her repræsenterer forskellige fysioterapeuter, *different persons*, og sonografiske billederne, *same occasion* i studiedesignet. For intratester reliabilitet, går det således på om den enkelte fysioterapeut opnår samme måleresultat, hvis ultralydsbillederne måles ved to eller flere anledninger. Den enkelte fysioterapeut agerer her, *same persons* i studiedesignet, og målinger ved forskellige anledninger, *different occasions*. (McGraw et al., 1996; Mokkink et al., 2010b).

Studiedesignet har derfor til hensigt, at analysere kilder til variation ved gentagende målinger som kan influere på reliabiliteten. Dette specielt relateret til systematisk variation. Kilder til variation søges i en eller anden grad kontrolleret ved standardisering af proceduren. Graden af standardisering er relateret til studiets problemstilling. Den grad af standardisering som dikteres, er graden af kilder til variation, som ønskes kontrolleret i studiet. (Carter & Lubinsky, 2016, pp. 249-253; de Vet et al., 2006). Det er ønskeligt at dokumentere reliabiliteten ud fra en standardiseringsprocedure som er anvendelig i klinisk praksis. I tillæg til at kunne benyttes i en større reliabilitetsundersøgelse, der inkluderer flere tester. Delvis standardisering anvendes derfor. Ingen standardisering vil utvivlsomt medføre dårlig reliabilitet specielt grundet patellarsenens anatomiske struktur (U. Fredberg et al., 2008). En meget standardiseret procedure giver mening hvis det ønskes at etablere højest mulig reliabilitet (Carter & Lubinsky, 2016, p. 251).

3.2. Begrebsdefinitioner

Definitionerne og operationalisering af intra -og intertester reliabilitet beskrives i relation til studiedesignet.

3.2.1. Intratester reliabilitet:

Nominel: I hvilket omfang målinger fra individer, der ikke har ændret sig, er de samme ved gentagende målinger udført af samme testere ved forskellige anledninger (Mokkink et al., 2010b).

Operational: Hvorvidt to målinger udført af samme fysioterapeut af patellarsenens tykkelse på sonografiske billeder ved to forskellige anledninger giver de samme resultater.

3.2.2. Intertester reliabilitet

Nominel: I hvilket omfang målinger fra individer, der ikke har ændret sig, er de samme ved gentagende målinger udført af forskellige testere ved samme anledning (Mokkink et al., 2010b).

Operational: Hvorvidt målinger udført af tre forskellige fysioterapeuter af patellarsenens tykkelse på sonografiske billeder giver samme resultat.

3.3. Udvalg

Udvalget bør reflektere populationen af interesse, fordi reliabiliteten er betydelig afhængig af fordelingen af de karakteristika som måles (de Vet et al., 2006; Weir, 2005).

Studiets population var således individer med klinisk diagnosticeret PT (symptomatiske sener) fra 18 år og over i Bergen og friske individer (asymptomatiske sener) uden tidligere diagnosticeret PT eller anden lidelse relateret til knæerne over 18 år.

I dette studie bestod udvalget af to grupper. En gruppe individer med symptomatiske sener og en gruppe friske individer med asymptomatiske sener jævnfør studiets hensigt og problemstilling.

3.3.1. Rekruttering

Deltagerne med klinisk diagnosticeret PT og friske blev søgt rekrutteret via primærhelsetjenesten (fastleger/fysioterapeuter) og lokale idrætsklubber. Der var to årsager til dette. PT er tidligere vist at forekomme specielt hos idrætsudøvere men også ikke-idrætsudøvere (Albers et al., 2016; Ø. B. Lian et al., 2005). Samtidig var hensigten at rekruttere patellarsener og dermed senetykkelse fra et bredt spekter af populationen og ikke kun elite idrætsudøvere med tanke på, at reliabiliteten skulle gælde for hele spekteret. I tillæg til dette blev der for at sikre et tilstrækkeligt stort udvalg rekrutteret deltagere fra familie og venner samt deres bekendtskaber. Udvalgt blev rekrutteret som et såkaldt bekvemmelighedsudvalg (Carter & Lubinsky, 2016, pp. 98-99).

Idrætsudøvere blev defineret som aktiv deltagelse i en idræt. Ikke-idrætsudøver blev defineret som ikke-aktiv deltagelse i en idræt, Ikke-idrætsudøvere kunne udøve fysisk aktivitet på egen hånd.

Ud fra de Vet et al (2011) sine anbefalinger var det ønskelig at rekruttere nærmere 50 symptomatiske og 50 asymptomatiske sener med minimum to målinger/testere (de Vet et al., 2011, p. 126). De fleste tidligere studier omhandlende reliabilitet tilknyttet ultralydsmålinger målinger af patellarsenen har anvendt op til 35 sener (U. Fredberg et al., 2008; Sunding et al., 2016). Nærmere 50 symptomatiske og 50 asymptomatiske sener blev derfor vurderet tilstrækkelig til at opnå et fornuftigt sikkerhedsinterval tilknyttet reliabilitetskoefficienten (Norman & Streiner, 2014, pp. 62-63).

3.3.2. Beskrivelse af udvalget

Beskrivelsen af udvalget illustreres i tabel II udfra inklusion -og eksklusionskriterier.

Tabel II: Inklusions –og eksklusionskriterier for deltagelse i studiet.

Inklusion	Eksklusion
Fra 18 år og over.	Under 18 år.
Klinisk diagnosticeret PT. <ul style="list-style-type: none">- Beretning om aktivitetsteltet smerte ved patellarsenens udspring ved apex patella- Smerte ved palpation analog til det smertefulde område.	<ul style="list-style-type: none">- Kirurgiske indgreb i knæet.- Reumatologiske sygdomme.- Neurologiske sygdomme.- Injektioner < 6 måneder- Tidligere diagnose med PT
Frisk. <ul style="list-style-type: none">- Ingen smerter eller funktionstab tilknyttet knæet sidste 12 måneder	
	Bekræftelse af differentiell diagnose ved grundig klinisk undersøgelse af knæet

Der blev fastsat en aldersgrænse på 18 år. Det primære argument for dette, var højere krav til etisk godkendelse af projektet, da individer under 18 ikke karakteriseres som myndige.

Næste inklusionskriterie var klinisk diagnosticeret PT. PT som klinisk diagnose blev operationaliseret som følger:

- En beretning om nuværende aktivitetsrelaterede smerter.
- Deltageren skulle præcis kunne angive smertelokalisationen til patellarsenens proximale del ved apex patella.
- Palpationsømhed analog til det smertefulde område.

Klinisk diagnosticeret PT er anvendt som inklusionskriterie, da dette tidligere er anvendt som inklusionskriterie i flere studier omhandlende PT (J L Cook et al., 2001; McAuliffe et al., 2016; Sunding et al., 2016; Visnes et al., 2015). Samtidig betragtes det fortsat som standarden for diagnosticering af PT i klinisk praksis (Brukner & Khan,

2012, pp. 700-707; J L Cook & Purdam, 2009; Malliaras et al., 2015). PT samt palpationsømhed er også fundet hyppigst at forekomme ved patellarsenens proksimale del rundt apex patella både klinisk og paraklinisk (J. L. Cook et al., 2000; Visnes et al., 2015). Palpationsømhed relateret til den nævnte lokalisation er begrænset undersøgt for reliabilitet og validitet. Et studie har tidligere fundet god test-retest reliabilitet for en tester ($r = 0,82$) men begrænset diagnostisk værdi (J L Cook et al., 2001). Erfaringsmæssigt beretter patienter imidlertid typisk, at det er ondt når de trykker på den proksimale del af senen, samtidig som det kontinuerlig anvendes som et inklusionskriterie i studier om PT (Comin et al., 2013; Visnes et al., 2015). Derfor indgik palpationsømhed som et inklusionskriterie i dette studie.

Beretninger om neurologiske eller reumatiske sygdomme var eksklusionsgrundlag. Først og fremmest fordi disse sygdomme ikke ville være repræsentativ for population. Der ville sandsynligvis også opstå problemer relateret til bestemmelser rundt symptombilledet. Specielt reumatologiske sygdomme ville kunne medføre generelle inflammatoriske anormaliteter i knæet og dets struktur og herved vanskeliggøre det rent diagnostisk at differentiere. Det gjorde sig ligeledes gældende for kirurgi. Yderlig ville kirurgi for PT antageligvis kunne medføre betydelige seneanormaliteter som arvæv og derved vanskeligheder tilknyttet afgrænsning af senen på RTUS (Bianchi & Martinoli, 2007, pp. 679-683; Docking et al., 2015). Injektioner blev også anvendt som eksklusionskriterie da disse er vist at kunne ændre på senedimensionen og muligvis hæmme afgrænsning af senen (Kongsgaard et al., 2009). For at udelukke anden årsag til symptomerne som patellafemoral smertesyndrom og afgrænse udvalget til symptomatiske og asymptomatiske sener var differentialdiagnoser eller mistanke herom yderlig et eksklusionsgrundlag. Med disse begrundelser indgik ovenfor nævnte elementer i eksklusionskriterierne.

In -og eksklusionskriterierne blev effektueret ved en klinisk undersøgelse. Tabel III viser det grundlæggende indhold i den kliniske undersøgelse:

Tabel III: Den klinisk undersøgelse

Kliniske diagnosticeret PT	
Anamnese	Sygdom debut/varighed og årsag smertelokalisation aktivitetsrelaterede smerter funktions/aktivitetsnedsættelse Behandlingshistorik Anden sygdom
Kliniske undersøgelse	Funktionsundersøgelse (ROM, fremprovokation af smerten ved funktion) Palpation Specifikke test (differentialdiagnostiske)

3.3.3. Variabler

Demografiske variabler relateret til beskrivelse af udvalget var alder, køn, idrætsudøver og ikke-idrætsudøver.

Testerne, deltagerne/sener på sonografiske billeder og testproceduren kunne i dette studie anses som uafhængige variabler eller faktorer, der kunne influere på variation i tykkelsen af patellarsenen og derved reliabiliteten som den afhængige faktor. Altså reliabilitet som variation mellem deltagerne, systematisk variation og tilfældig variation (Weir, 2005).

3.4. Testere

I dette studie indgik der i alt tre fysioterapeuter med forskellig erfaring indenfor RTUS. I form af deres grunduddannelse som fysioterapeut havde alle testerne basal kundskab om anatomi og fysiologi. I tillæg havde alle tre fysioterapeuter som minimum basal oplæring i RTUS. Disse tre agerede testere i dette studie. Testerne blev rekrutteret som et bekvemmelighedsudvalg fra universitet i Bergen og kollegaer.

To af fysioterapeuterne (tester A og C) var novicer med basal oplæring i RTUS med ingen til begrænset empirisk erfaring med RTUS. Den tredje fysioterapeut (tester B) var

erfaren med en mere avanceret oplæring i RTUS og mange års erfaring i brugen af RTUS indenfor blandt andet seneforskning.

Selvom det var ønskelig med hensyn til den eksterne validitet, blev der ikke inkluderet flere testere i dette studie grundet tidsperspektivet i forhold til oplæring og måling (de Vet et al., 2011, pp. 124-126; Van Genderen et al., 2003).

3.4.1. Operationalisering af begreber og beskrivelse af testerne.

Novice: operationaliseres som nybegynder indenfor RTUS med basal oplæring i RTUS, ingen til begrænset erfaring med RTUS og ikke arbejdende regelmæssig med RTUS.

Ingen erfaring: Ingen tidligere erfaring med brug af RTUS udover det som er opnået i forbindelse med processen op til under studiet.

Begrænset er erfaring: Sporadisk brug af RTUS. Arbejder ikke regelmæssig med RTUS.

Basal oplæring: Kortvarig oplæring i de mest basale funktioner og principper for RTUS af patellarsenen og nogle timers praktisk øvelse. Oplæring kan være gennemført på universitet i Bergen i regi af dette projekt eller i forbindelse med undervisning eller anden forskningsarbejde.

Erfaren tester: ≥ 10 års erfaring med ultralydsskanning af det muskuloskeletale system samt regelmæssigt arbejde med RTUS i forskningsregi.

Avanceret oplæring: Gennemført flere kursus indenfor brugen af RTUS, samt oparbejdet solid kundskab i brugen af RTUS gennem mange års erfaring.

Fysioterapeuterne vil videre blive omtalt som tester A, B og C.

3.4.2. Beskrivelse af testerne i studiet.

Beskrivelse af de tre tester og specificering af deres baggrund er præsenteret i tabel IV.

Tabel IV: beskriver erfaringsniveauet til de tre testere

Testere	Erfaring.
Tester A Novice	Følgende karakterisere Tester A <ul style="list-style-type: none">- Mand- Fysioterapeut- 2 - 3 timers oplæring- 6 timers øvelse- Projektansvarlig
Tester B Erfaren	Følgende karakterisere Tester B <ul style="list-style-type: none">- Mand- Fysioterapeut/Ph.d.- ≥ 10 års erfaring med RTUS indenfor specielt tilknyttet forskning. <p>Naturlig indføring i projektet som Vejleder. Sammenlagt cirka 1-2 timers øvelse og indføring i måleproceduren.</p>
Tester C Novice	Følgende karakterisere Tester B <ul style="list-style-type: none">- Kvinne- Fysioterapeut- Undervisning i RTUS cirka 14 timer.- 1-2 times indføring i projektet, måleprocedure og øvelse.- Tidligere indføring i RTUS af skulder og hertil måling af sener i forbindelse med et studie.

Tester A var projektansvarlig for indsamling af ultralydsbilleder. Der blev i forbindelse med dette ikke foretaget måling af patellarsenerne, men erfaring og læring relateret til vurdering af patellarsenen på ultralydsbilleder samt skanning blev formentlig opnået. Læringen opnået ved indsamling af ultralydsbilleder var > 20 timer. Derfor blev Tester A og C vurderet som ligeværdige i udgangspunktet for måling af senerne, selvom Tester C havde 14 timers undervisning og erfaring med RTUS fra tidligere.

3.5. Etiske betragtninger

Dette reliabilitetsstudie fulgte de videnskabetiske krav opstillet af Helsinki deklARATIONEN vedrørende behandling af forsøgspersoners personlige oplysninger samt lov om deltagernes retsstilling (Association, 1964). Formålet med dette var at sikre de enkelte forsøgspersoners integritet og autonomi ved hjælp af:

- Informeret og skriftligt samtykke. Herunder frivillighed, anonymitet, fortrydelsesret.
- Sikre deltagerens anonymitet.
- Hensyntagen til mulige risici ved deltagelse i undersøgelsen vejet op over for undersøgelsens fordele.
- Vurdering af undersøgelsens relevans og videnskabelig metode.
- Publicering af undersøgelsesresultaterne (pligt til at publicerer resultater både positive og negative).

Der blev søgt godkendelse ved Regional Etisk Komite Vest og studiet blev vurderet til ikke at falde ind under helseforskningsloven og studiet kunne således udføres.

Reference nummer på godkendelse fra REK: 2016/703. Se Bilag 1 for REK godkendelse.

Der blev ikke vurderet at være nogen fysiske eller psykiske risici ved deltagelse i studiet, da RTUS ikke betragtes som en metode der forvolder forsøgspersonerne nogen form for skadelig virkning, men derimod anses som en smerte –og strålefri metode (Bianchi & Martinoli, 2007, pp. 11-18; Docking et al., 2015). Derimod er det vigtigt at sikre deltagerens anonymitet i studiedesignet og specielt ved eventuel publicering af resultatet. Alt data anonymiseret og det var ikke mulig at koble personidentificerbart data til dataene hverken direkte eller indirekte. Alle blev bedt om informeret samtykke før endelig deltagelse (se Bilag 2).

Forsøgspersonerne havde ikke nogen direkte fordel ved deltagelse i studiet relateret til symptomer, da der ikke var snak om et effektstudie. Deltagerne blev dog tilbudt råd og vejledning i forbindelse ved deltagelse tilknyttet PT. Studiet kan alligevel siges indirekte at være gavnlig for deltagerne, grundet anvendelse af RTUS som supplement i diagnosticeringen og vurderingen, men foreløbig med begrænset dokumentation for reliabiliteten (Kunnskapsdepartementet, 2006).

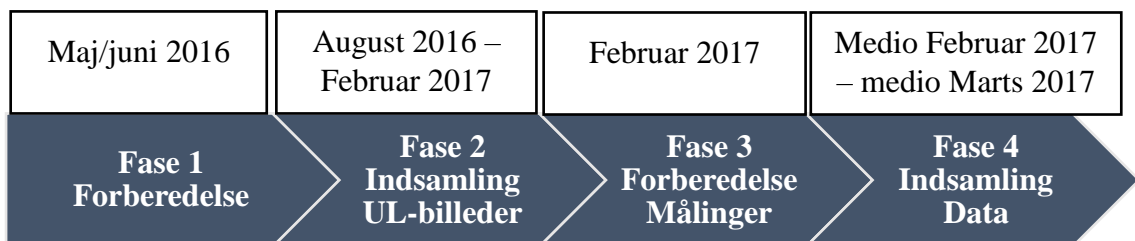
3.6. Dataindsamling

I dette afnit vil dataindsamling blive beskrevet. Hertil hvordan processen var opbygget, hvordan indsamlingen af sonografiske billeder blev udført, og hvordan selve målinger af patellarsenens tykkelse foregik.

Tidsperspektivet for forløbet skal ses som et tværsnit, idet de indsamlede data er tilknyttet et tidspunkt (Carter & Lubinsky, 2016, p. 189).

3.6.1. Dataindsamlingsfaser

Indsamlingen af data bestod af fire overordnede faser. Første fase var forberedelsesfasen efterfulgt af en rekrutterings -og indsamlingsfase, der igen blev efterfulgt af en ny forberedelsesfase, og til slut den endelige dataindsamlingsfase. Hele dette forløb startede i maj/juni 2016 med forberedelse og stræk sig til medio marts 2017. Figur 5. illustrere forløbet.



Figur 5. Viser en tidslinje for hele dataindsamlingsprocessen.

I hele dette forløb indgik testerne A, B og C i forskellig udtrækning. Tester A var projektansvalig og indgik i hele forløbet. Det var kun tester A, der indsamlede Sonografiske billeder. Selvom det var ønskelig at enten tester B eller C også skulle indsamle billeder fra samme individer, var dette desværre ikke muligt grundet tidsperspektivet. Tester B deltog kun delvist i fase 1 og fase 2 og fuldt ud i fase 3 og 4. Tester C blev hentet ind sent i forløbet, da det var ønskeligt at styrke den eksterne validitet. Tester C måtte derfor igennem en del af fase 3 før vedkommende kunne deltage i fase 4. Jævnfør afsnit 3.4. blev tester A og C karakteriseret som novicer og Tester B som specialist. Tester A stod for inklusionsundersøgelserne i forbindelse med indsamling af de sonografiske billeder. Tester A havde mere end seks års erfaring med klinisk arbejde specielt rettet mod muskuloskeletale lidelser og idrætsrelaterede skader.

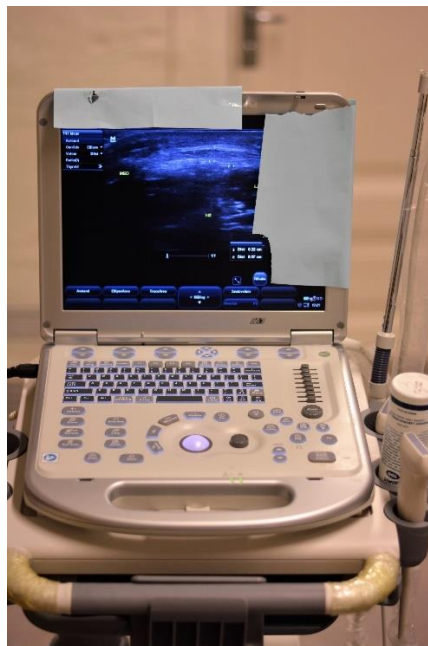
Første fase var forberedelsesfasen, hvor tester A og B først opnåede endelig enighed rundt forsøgsprotokollen relateret til selve sonografiske skanningen. Altså den tekniske

og praktisk fremgangsmåde for indsamling af sonografiske billeder. Dette foregik ved indsamling af ultralydsbilleder fra to individer, hvilket også tjente som en del af indføringen og den praktisk øvelse i RTUS for tester A. Der blev i denne proces også diskuteret hvorledes billeder skulle måles i fase 4, men ingen endelig bestemmelse blev nedfældet i denne fase. Videre gennemgik tester A praktisk øvelse tilknyttet skanning af yderlig et individ. Ingen af de individer som blev skannet i denne fase indgik i datamaterialet. Disse var kun et led i forberedelsen.

I anden fase startede rekrutteringen af deltagere. Der blev sendt ud skriftlig information til primærhelsetjenesten, idrætsklubber og mundtlig information til kollegaer, familie og venner. Individer som ønskede at deltage, tog selv kontakt til tester A (projekt ansvarlig) via e-mail eller telefon. Deltagerne afgav alle informeret samtykke til tester A ved fremmøde jævnfør afsnit 3.5. og bilag 2. Der blev kun registreret navn og telefonnummer i forbindelse med aftale af time til undersøgelse, og denne information blev slettet umiddelbart ved ankomst. Ved fremmøde blev der først udført en klinisk undersøgelse med henblik på inklusion og eksklusion. Resultatet blev nedfældet i et dataskema (se Bilag 3 for skemaet). Derefter blev der udført en skanning af begge knæ. Alle deltagerne blev skannet ved fremmøde uafhængigt om de opfyldte inklusions eller eksklusionskriterierne. Alle deltagerne blev anonymiseret ved kodenormering. Personidentificerbart data blev ikke indsamlet. Under skanningen blev der frosset og gemt billeder af patellarsenen i et longitudinal -og tværsnitsformat for begge knæerne. Der blev ikke udført nogen målinger i denne fase. Alle skanningerne blev udført på universitet i Bergen med undtagelse af 10 individer, som havde lyst til at deltage, men kun mulighed for at stille i forbindelse med deres idrætsaktivitet. Da dette var en mulighed for at rekruttere en del deltager, blev dette accepteret. Disse skanninger blev gjort over to seancer, hvor Tester B deltog i forbindelse med udlån af ultralydsapparat. Tester B udførte ingen skanninger og var blindet overfor anonymiseringen. Standardiseringsproceduren var akkurat den samme.

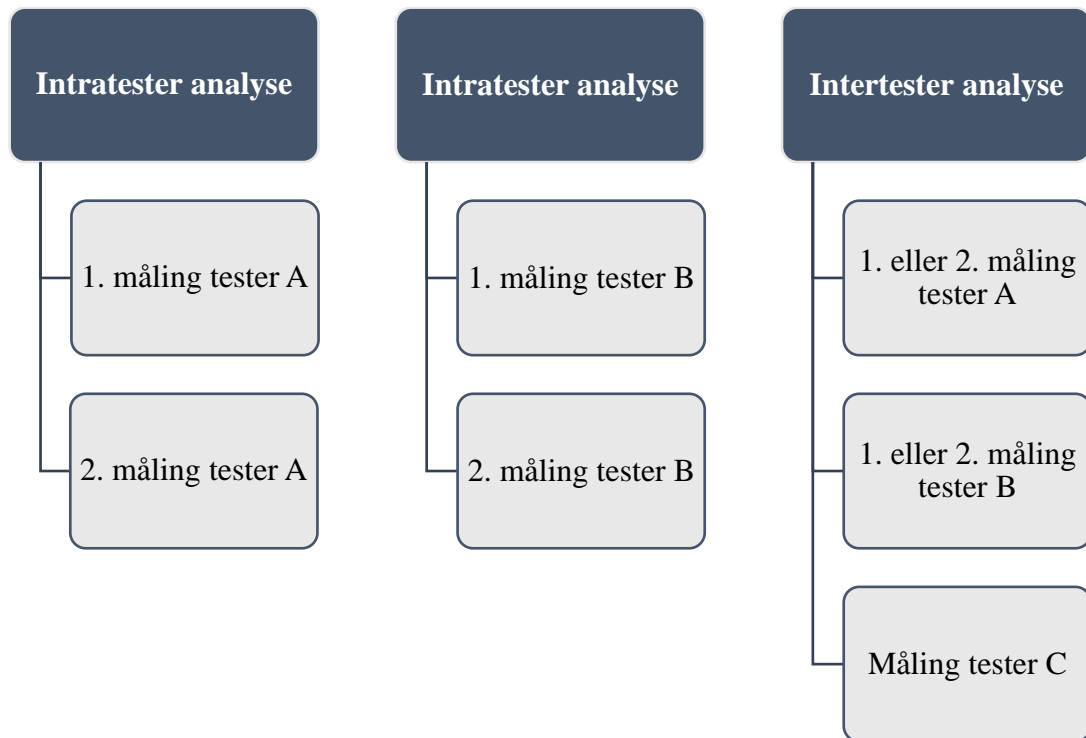
Efter endt rekruttering blev der i fase 3 opnået enighed mellem tester A og B vedrørende procedure for måling af billederne. Der blev efterfølgende udført prøvemålinger på ekskluderede sener og endelig enighed om den præcise fremgangsmetode blev nedfældet i forsøgsprotokollen af tester A.

Ved fase 4 blev målingerne udført jævnt før forsøgsprotokollen. Tester A og B udførte hver en måling af alle senerne på ultralydsapparatet uafhængigt af hinanden. Efter minimum fire ugers ophold foretog de uafhængigt af hinanden måling nummer to. Den ene af de to målinger fra Tester A og B kunne således indgå i intertester reliabilitetsanalysen og deres gentagende målinger *on different occasions* kunne agere i intratester reliabilitetsanalysen. Det blev vurderet at der med hensyn til den eksterne validitet burde indgå flere testere. Det lykkedes dog kun at rekruttere yderlig en tester. Tester C blev derfor indført i studiet, måleproceduren og gennemgik lidt øvelse i skanning for at skabe overblik tilknyttet vurdering af billederne. Yderlig udførte tester C et par prøve målinger på ekskluderede billeder. Dette med tanke på homogenitet mellem tester A og C. Tester C foretog kun en måling af alle senerne, og deltog derfor kun i intertester reliabilitetsanalysen. Alle målingerne blev udført på de gemte billeder på ultralydsapparatet. Tester B og C var blindet via anonymiseringen af billederne og kunne der udføre målingerne uden yderligere tiltag. Tester A havde stået for indsamlingen af billeder, og i realiteten var det muligt for tester A at sætte normeringsrækkefølgen op mod karakteristika på skanningsbillederne. Derfor blev tester A blindet ved hjælp af tester C som blindede normeringen af billederne. Se figur 6 for illustration af dette.



Figur 6. Information relateret til billedenummer og deltager ID blev blindet som vist i figuren. Således kunne tester A kun se senen.

Samtidig udførte tester C en tilfældig fordelingsprocedure for hvilke billeder, der skulle måles. Dette blev gjort for begge tester A's målingsseancer. Figur 7 illustrere hvordan testernes målingsseancer indgik i henholdsvis intra -og intertester reliabilitetsanalysen.



Figur 7. Fra venstre ses intratester analysen for tester A og i midten intratesteranalysen for tester B. Til højre ses intertester analysen. Figuren illustrerer hvordan de forskellige målingsseancerne indgik i analyserne.

Alle testerne blev således eksponeret for samme standardiseringsprocedure samt de samme skanningsbilleder, som alle var indhentet af tester A. Tester B og C indhentede ingen billeder.

3.6.2. Udstyr

Ultralydsapparatet som blev anvendt i studiet var '*Mindway M7 Portable Color Ultrasound Diagnostic System*' (Mindway, Kina) med en 11Mhz lineær transducer med indstilling i gråtone B-mode. Figur 8 viser ultralydsapparatet.



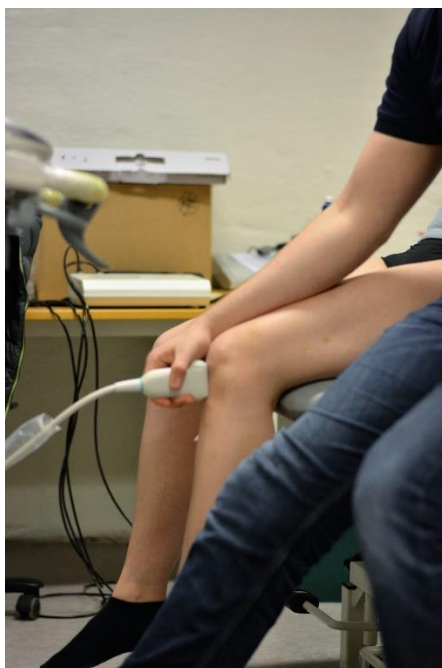
Figur 8. Mindway M7 blev anvendte til indsamling af sonografiske billede og målinger af patellarsenens tykkelse.

3.6.3. Procedure

Indsamlingsfaserne forstået som både indsamling af ultralydsbillederne og selve målingen af billederne var baseret på en forsøgsprotokol. Forsøgsprotokollen havde til hensigt at forankre sig i en fremgangsmåde som ville gøre sig mulig i klinisk praksis. Videre var hensigten at styrke den interne validitet. Forsøgsprotokollen var derfor baseret på en analyse af kilder til systematisk og tilfældig fejl. Kilder som blev vurderet til at kunne bidrage betydeligt til systematisk og/eller tilfældig fejl blev forsøgt kontrolleret ved en gradvis standardiseret tilgang (Carter & Lubinsky, 2016, p. 251).

3.6.4. Procedure for ultralydsskanning

For skanning af patellarsenen blev deltagerne placeret i siddende stilling på en behandlingsbænk med fuld støtte under lårene og underbenene hængende afslappet ned med cirka 90 graders fleksion i knæet. Både longitudinale billeder og tværsnitsbilleder var opnåelig i denne position. I figur 9 ses den standardiserede position af deltagerne for skanning.

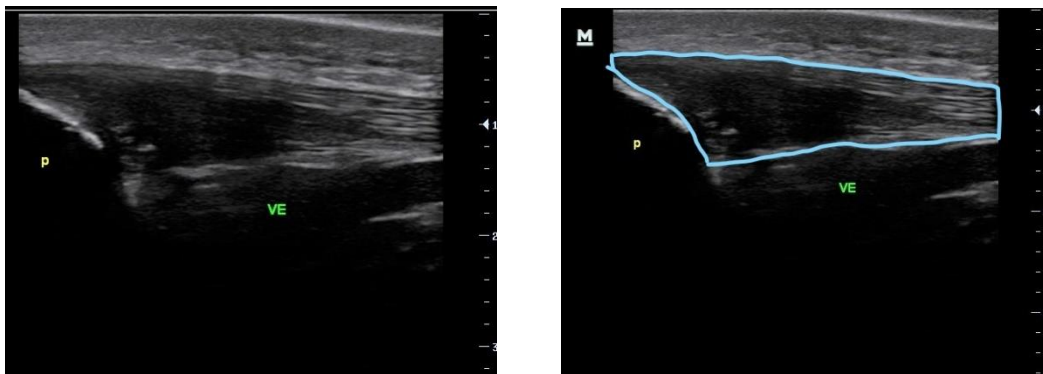


Figur 9. Skanningen blev foretaget med deltager siddende med knæet i cirka 90 grader.

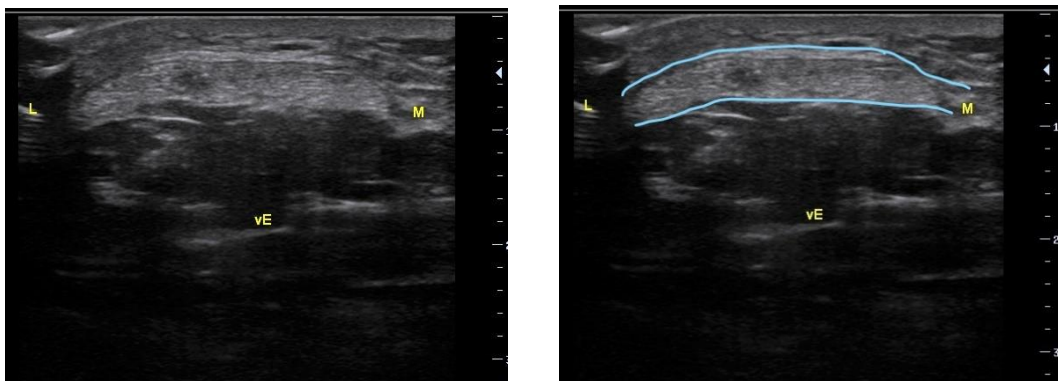
Efterfølgende blev der først foretaget en skanning af patellarsenen i en longitudinal retning. Transduceren blev som vist på figur 8 holdt longitudinal til patellarsenen. Der hvor senen fremstod tykkest og vurderet til at fremstå med en klar profund og superficiel seneafgrænsning, blev billedet frosset og gemt. Herefter blev der foretaget en transversal drejning af transduceren på samme sted for at sikre tilnærmelsesvis samme lokalisation. I både det longitudinale og transversale plan blev apex patella anvendt som et anatomisk holdepunkt. Hensigten var ligeledes at sikre samme anatomiske lokalisation. Proceduren videre for tværsnitsbilleder var ellers den samme, blot med transduceren transversal til patellarsenen. Alle skanningerne blev udført på begge knæ.

Som nævnt tidligere er klinisk diagnosticeret PT vist at være betydelig mere frekvent i den proksimale del af patellarsenen. Dette er fundet at stemme overens med de patologiske forandringer, der hyppigst er vist at forekomme ved patellarsenens udspring ved apex patella (Visnes et al., 2015). Derfor blev det på forhånd fastlagt i forsøgsprotokollen at den proksimale del af patellarsenen, skulle være den standardiserede lokalisation for undersøgelse og billede indsamling. Således var fokus på, at få et godt billede af den proksimale del af senen.

Stilling med 90 graders fleksion i knæet blev anvendt med tanke på at undgå anisotropy, hvilket lettere vil kunne forekomme ved knæet fuldt ekstenderet og patellarsenen afslappet på en bænk. Således blev transduceren holdt lodret til senen (U. Fredberg et al., 2008). Forskellige skanningspositioner er blevet anvendt i forskellige studier. Samtidig eksisterer der ingen universel konsensus angående en specifik standardiseret position for skanning. Position i dette studie er tidligere anvendt i flere studier på symptomatiske og asymptomatiske sener (Bjordal et al., 2003; U. Fredberg et al., 2008; Kongsgaard et al., 2009; Sunding et al., 2016). Figur 10 viser et longitudinal snit af patellarsenen og figur 11 et transversal snit også omtalt som tværsnitsbilleder.



Figur 10. Fra venstre ses et longitudinal snit af patellarsenen. P markerer patella og VE står for venstre knæ. Senen fremstår her fortykket og hypoekkoïd. På billedet til højre er senen afgrænset med et blå område.



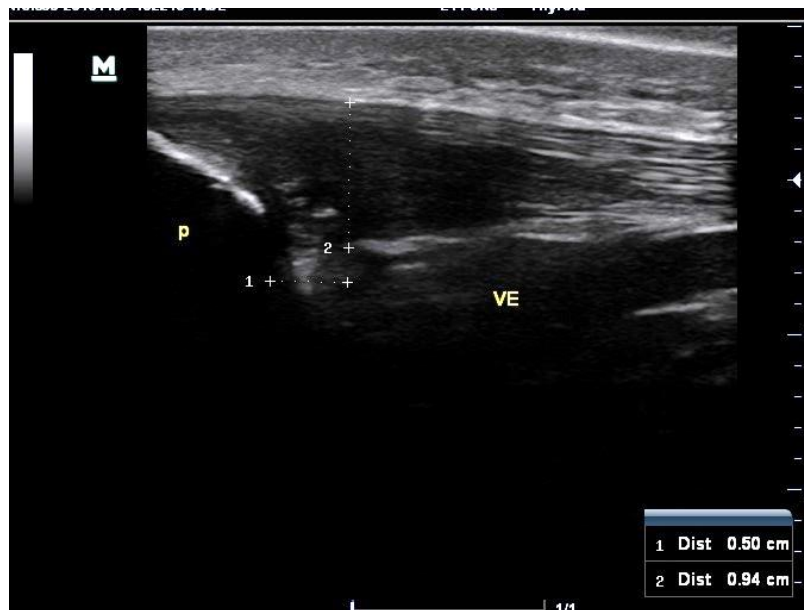
Figur 11. Begge billeder viser et tværsnitsbillede af en patellarsene. M står for medial og L for lateral. Senen fremstår uden seneændringer og hyperekoïd. Det blå område til højre illustrer seneafgrænsningen.

3.6.5. Procedure for måling af senerne.

Målingerne blev udført på billeder af patellarsenen i et longitudinal og et transversal snit på RTUS. Disse målinger blev kvantificeret i centimeter af ultralydsapparatets selvkalibrerende software og registreret i et excel-ark (se Bilag 4) for illustration af dataskemaet).

3.6.5.1. Longitudinale målinger.

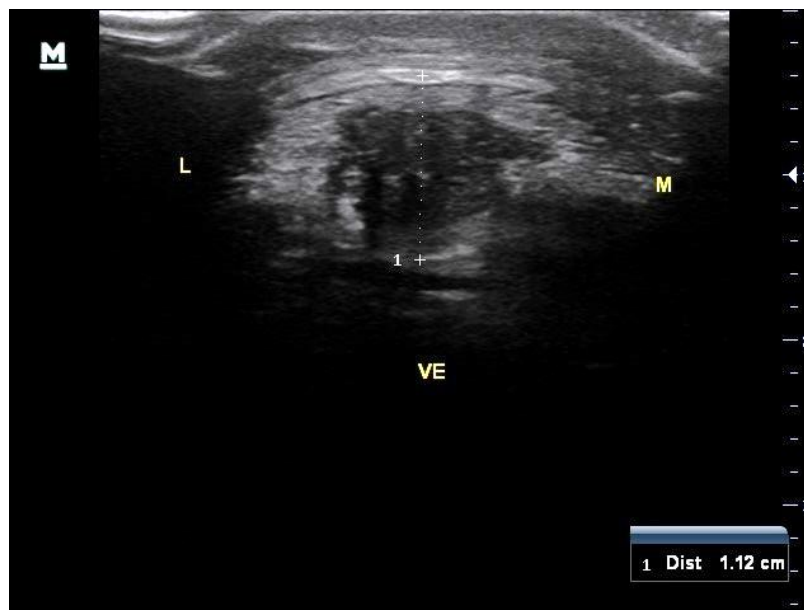
Den longitudinale tykkelse blev målt i posterior-anterior retning fra den profunde til den superficielle del af seneafgrænsning. Først blev apex patella lokaliseret og derefter den profunde afgrænsning (peritendineum) af patellarsenen. Udfra dette krydsningspunkt blev der sat en måling på 0,5 centimeter ret under den profunde afgrænsning af patellarsenen i en distal retning mod tibia. Herfra blev det første målepunkt sat i den ydre del af den profunde seneafgrænsning. I en posterior- anterior retning blev en ret linje ført op til den indre del af den superficielle seneafgrænsning. Figur 12 viser hvordan patellarsenen blev målt i det longitudinale snit.



Figur 12. Figuren illustrerer målingen på et longitudinal snit af en patellarsene med seneforandringer. På billedet ses et P som refererer til patella og VE for venstre knæ. 1. måling angiver de 0,5 cm ud fra apex patella ret under den profunde seneafgrænsning. 2. måling viser den posterior-anterior måling af senetykkelsen fra den ydre profunde seneafgrænsning til den indre superficielle seneafgrænsning.

3.6.5.2. Transversale målinger

Den transversale tykkelse blev ligeledes målt i en posterior-anterior retning mellem den profunde og superficielle seneafgrænsning. Målingen foregik over den midterste del af senen som den fremstod på billedet. Det første målepunkt blev afsat på den ydre kant af den profunde seneafgrænsning. Herfra blev en ret linje trukket anterior til midt i den superficielle seneafgrænsning for måling af senen posterior-anterior tykkelse. På figur 13 ses hvordan målingen blev udført.



Figur 13. Figuren viser et tværsnit af den proximale del af en patellarsene med seneforandringer. L og M indikerer laterale og mediale side. VE referer til venstre knæ. Måling ses som den stiplede linje fra den ydre profunde seneafgrænsning til midt i den superficielle seneafgrænsning. Som det ses, blev målingerne foretaget midt i senen.

3.7. Databearbejdning

De statistiske analyser blev udført ved hjælp af softwaren Statistical Package for Social Science (SPSS) version 24 og Excel 2016.

Demografiske data blev beskrevet som deskriptiv statistik. Den gennemsnitlige senetykkelse blev præsenteret for gruppen med symptomatisk og asymptomatiske sener. Dette som gennemsnittet af alle målinger udført af alle tre tester. Da dataene var på ratio niveau blev der udført en uparret t-test for at analysere, om der var en signifikant forskel i senetykkelse mellem de to grupper. Signifikansniveauet blev sat til 5% ($\alpha = 0.05$) (Beyer et al., 2012, pp. 51-53).

For belysning af reliabiliteten relateret til problemformulering blev intra -og intertester reliabiliteten kvantificeret i relativ og absolut reliabilitet.

3.7.1. Relativ reliabilitet

Siden datamaterialet præsenterede kontinuerligt data på ratioskala kvantificeret i millimeter blev ICC med 95% konfidensinterval anvendt til belysning af den relative reliabilitet. ICC beregnes på baggrund af en variansanalyse og genspejler graden af sammenhæng og enighed mellem to eller flere målinger (Beyer et al., 2012, p. 58).

ICC blev baseret på ICC3 (A.1) som i SPSS er analog til '*Two-Way-Mixed Absolute Agreement*'. Tretallet i denne formel implicerer at testerne ikke er repræsentativ for hele population af tester. Således forudsætter denne formel tolkningsmæssigt, at der ikke kan generaliseres til andre testere. Dette blev valgt, da der kun indgik 3 tester og de næppe kunne antages at repræsentere hele populationen af tester. Videre implicerede 'A' at systematisk variation mellem testerne indgik i beregningen af ratioforholdet mellem variationskilder. Dette blev valgt, da det også var ønskelig at belyse det reelle forhold af variation mellem testere/gentagende målinger. I tillæg blev ICC3 (C.1) valgt. Denne inkluderede ikke systematisk variation mellem testerne. Således kunne det udledes ved sammenligning med ICC3 (A.1), om systematisk fejl var nævneværdig tilstede. (McGraw et al., 1996; Weir, 2005). Formelen kan defineres som:

$$ICC3 (A.1) = \frac{\sigma_s^2}{\sigma_s^2 + \theta_o^2 + \sigma_{residual}^2}$$

$$ICC3 (C.1) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_{residual}^2}$$

Selve beregningen er baseret på høstede MS-værdier fra en tovejs ANOVA og kan kvantificeres udfra følgende formel:

$$ICC3 (A, 1) = \frac{MS_R - MS_\epsilon}{MS_R + (k - 1)MS_\epsilon + \frac{k}{n}(MS_c - MS_\epsilon)}$$

$$ICC3 (C.1) = \frac{MS_R - MS_E}{MS_R + (k - 1)MS_E}$$

F-testen fra ANOVA-tabel blev vurderet, med henblik på om en eventuel forskel mellem ICC3 (C.1) og ICC3 (A.1) indikerende systematisk variation mellem testerne var signifikant. Hvis denne var signifikant således under $p < 0.05$, indikerede dette en signifikant systematisk fejl (Weir, 2005).

Det blev i denne opgave valgt at gradere ICC. Dette kun som en metode til at karakterisere reliabilitetskoefficienten på, og ikke hvorvidt reliabiliteten var acceptabel i given kontekst (Lindahl & Juhl, 2010). ICC blev graderet på baggrund af følgende tre kategorier:

0.00-0.50 Ingen til lav reliabilitet

0.50-0.75 Moderat reliabilitet

> 0.75 God reliabilitet

I metodelitteraturen henvises der ofte til at måleinstrumenter som skal anvendes på individniveau bør have ICC-værdier >0.90 , og ICC-værdier for måleinstrumenter som skal bruges på gruppeniveau eller i forskning bør være >0.70 . Dette kan anvendes som et estimat for, hvad der kan accepteres tilknyttet RTUS som en specifik målemetode i perspektiv til diagnosticering og evaluering (Beyer et al., 2012, p. 58; Streiner et al., 2015, pp. 183-184).

3.7.2. Absolut reliabilitet

Da ICC har en begrænset effekt til at informere om størrelsen af målefejlen og i tillæg er afhængig af fordelingsbredden, blev absolut reliabilitet beregnet. Målefejlen blev videre kvantificeres målefejlen typisk i dimensionsspecifikke værdier og i dette tilfælde millimeter. Således kvantificeres størrelsen på målefejlen, hvilket er relevant i et evalueringsperspektiv (de Vet et al., 2006).

I dette studie blev SEM og SDC udregnet. SEM fortæller at et individs sande score med 68% sikkerhed befinder sig indenfor 1 SEM. SDC fortæller at et individs sande score

ved gentagende målinger med 95% sikkerhed befinder sig indenfor 1 SDC. Herved opnås der information om målefejlen for en måling og for gentagende målinger. SEM blev udledt jævnfør de Vet et al (2011) fra en varianskomponent analyse i SPSS som bygger på en ANOVA (de Vet et al., 2006; de Vet et al., 2011, p. 106). I denne model blev variationen analyseret og kvantificeret i varianskomponenter, som ses indgå i formelen under. Således blev formelen til udregning af SEM udledt.

$$SEM_{agreement} = \sqrt{(\sigma_o^2 + \sigma_{residual}^2)}$$

Denne formel indeholder systematisk variation mellem testere og beskriver den gennemsnitlige målevariation. SDC blev udledt fra SEM værdien på følgende måde:

$$SDC_{Individ} = SEM \times 1.96 \times \sqrt{2}$$

Som det ses af formelen udregner denne målevariationen på individniveau (de Vet et al., 2006).

3.7.3. Normalfordeling

De ovenfor nævnte statistiske parametre forudsætter at dataene er normalfordelte og homoscedastiske. ICC som bygger på en ANOVA er mere robust overfor skævt fordelte data og heterogen variation mellem dataene. Samtidig bygger ANOVA på forskelle mellem middelværdier. Relateres dette til Central Limit Theorem vil en fordeling typisk uafhængig af den oprindelige fordeling forekomme normalfordelt (Norman & Streiner, 2014, p. 88).

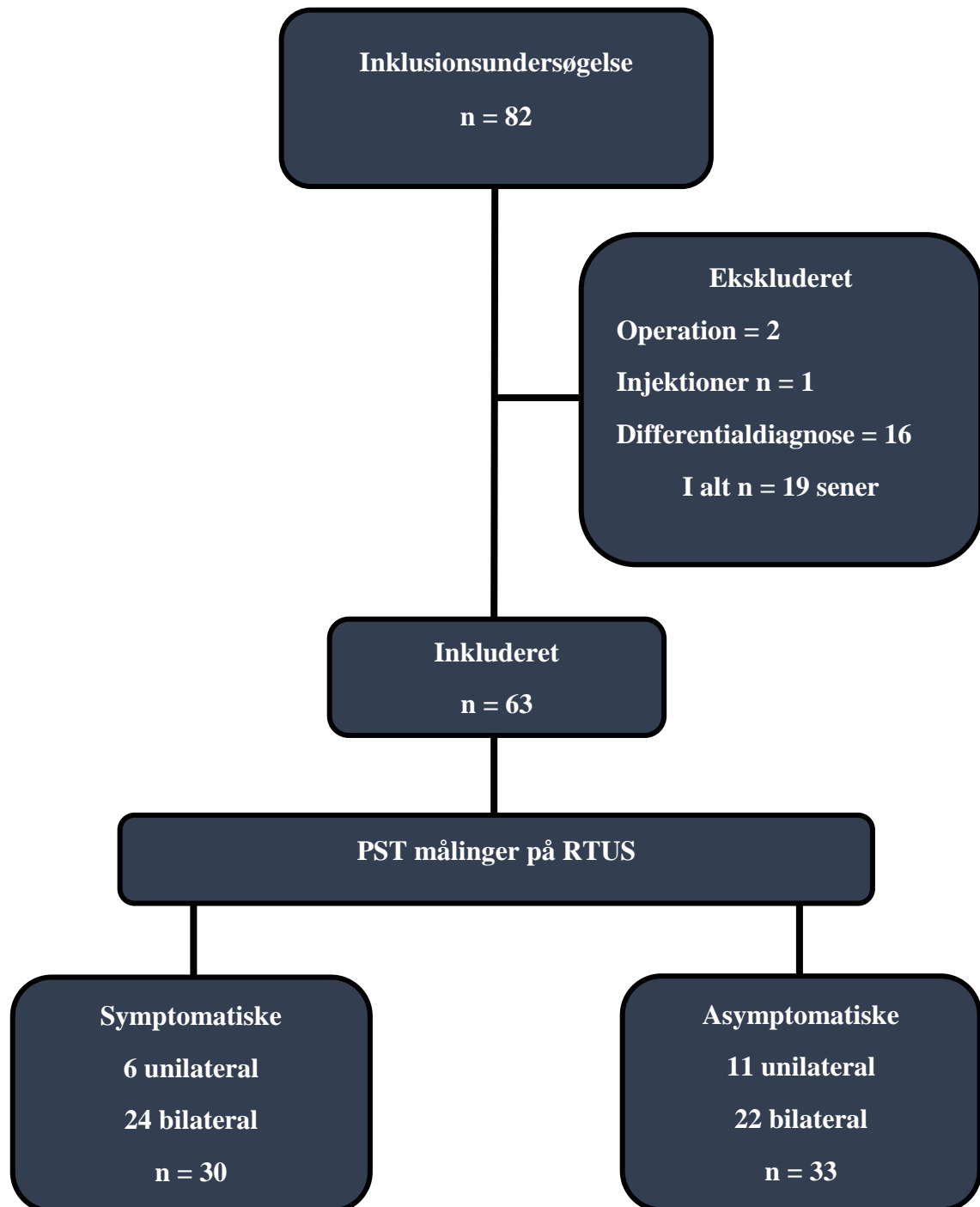
SEM derimod vil nødvendigvis ikke genspejle målevariation så godt, hvis målevariationen eksempelvis stiger i takt med scoren (Beyer et al., 2012, pp. 62-63).

Datamaterialet for de enkelte målingsseancer tilknyttet gruppen med symptomatiske sener og gruppen med asymptomatiske sener blev undersøgt for normalfordeling. Alt datamaterialet for måling af patellarsene tykkelsen på sonografiske billeder blev grafisk testet for normalfordeling med histogrammer og normal probability plots (*Normal Q-Q plot*). I tillæg blev scatterplots kontureret for datamaterialet som skulle anvendes i

dataanalysen for en visuel vurdering af variationsforholdet mellem variableerne (tilstedeværelsen af heteroscedastisitet) (Norman & Streiner, 2014, pp. 25-27).

4. Resultat

Der deltog i alt 41 deltager i indeværende undersøgelse. Samlet indgik der 82 patellarsener i inklusionsundersøgelsen. Af disse blev 19 sener ekskluderet grundet tidligere operation, injektioner i senen eller grundet differentialdiagnose. I alt blev 63 patellarsener inkluderet i studiet, hvoraf 33 var asymptomatiske sener og 30 var symptomatiske sener. Processen er illustreret ved et flowdiagram i Figur 14.



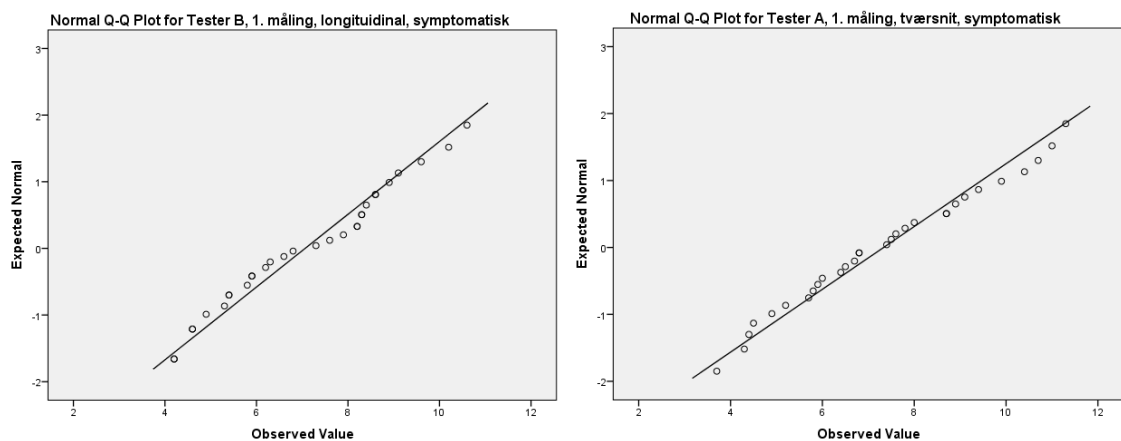
Figur 14. Flowdiagram af reliabilitetsundersøgelsen.

4.1 Normalfordeling og variations homogenitet

Datamaterialet indsamlet fra symptomatiske sener fremstod acceptabelt normalfordelt på histogrammer (se Bilag 5). Dog viste dataene fra første måling udført af tester B på tværsnitsbilleder af symptomatiske sener en mindre højre skæv fordeling.

Materialet fra asymptomatiske sener viste gennemgående en tendens til højere skævhed for tester A og B, hvorimod datamaterialet for tester C fremstod normalfordelt. Højre skævheden og kurtosis var for alt datamaterialet mindre end 2.0 (*Skewness og kurtosis* < 2.0). Dette var ikke nok til at forkaste hypotesen om normalfordelte data.

Samtlige normal Q-Q plots viste en tilnærmelsesvis acceptabel lineær sammenhæng mellem de observerede data og den estimerede normalfordeling. I Figur 15. ses to eksempler på fordelingen af data på normal Q-Q plots. Disse afspejlede den generelle tendens.



Figur 15. Betragtes graferne afviger enkelt målinger af patellarsenens tykkelse fra middelværdien. Dette er imidlertid ikke nok til at forkaste antagelse om normal fordelte data.

Forholdet mellem to målinger blev visuelt fremstillet i scatterplots. Disse viste gennemgående en visuel fornuftig homogen variation mellem målingerne (Se Bilag 6).

Da dataene var på rationiveau, variationen mellem score vurderet tilnærmelsesvis homogene og datamaterielt blev antaget at tendere mod en normalfordeling, kunne datamaterialet behandles med parametrisk statistik.

I intertester analysen blev den 2. målesæt for tester A og B anvendt. Denne forekom mere normalfordelt for tester B. For den resterende intertester analyse blev 1. målesæt fra tester A og B anvendt.

4.2. Deskriptiv statistik og forskel mellem grupperne.

Tester er beskrevet i afsnit 3.4.

4.2.1. Demografiske variabler

Grupperne var homogene med hensyn til alder, som det ses i Tabel V. Frekvensen af mænd var højere i den symptomatiske gruppe og frekvensen af kvinder var højere i den asymptomatiske gruppe. I Tabel VI ses det at forekomsten af idrætsudøverere var højere i den symptomatiske gruppe (90%) i forhold til den asymptomatiske gruppe (58%).

Tabel V. Beskriver aldersfordelingen for individer i henholdsvis den symptomatiske og den asymptomatiske gruppe

Alder	Gns.	SD	Min.	Maks.
Symptomatiske	26.67	11.62	18	56
Asymptomatiske	28.55	10.29	18	60

Gns = gennemsnit, SD = Standard deviation, Min-Max = minimum og maksimum alder for deltagerne.

Tabel VI. Frekvensen af mænd og kvinder samt idrætsudøver kontra ikke idrætsudøver i afrundede tal.

Variabler	Symptomatisk (n = 30)		Asymptomatisk (n = 33)	
	Antal	Procent %	Antal	Procent %
Mand/kvinde	20/10	67/33	18/15	55/46
Idræt/ikke idræt	27/3	90/10	18/18	58/42

Antal = frekvens antal, Procent % = frekvens %

4.2.2. Senetykkelsen

Den gennemsnitlige posterior-anterior tykkelse af patellarsenen for symptomatiske sener var både i det longitudinale og transversale snit signifikant tykkere end asymptomatiske sener ($p = 0.001$). Forskellen mellem symptomatiske og asymptomatiske sener er præsenteret i Tabel VII.

Tabel VII. Her ses den gennemsnitlige forskel for patellarsenens tykkelse mellem grupperne. Gennemsnittet blev udregnet baseret på gennemsnittet alle udførte målinger af alle testerne. Resultatet er baseret på en uparret t-test.

Variabler	Symptomatisk		Asymptomatisk		Gns. forskel	p-værdi
	Gns.	SD	Gns.	SD		
Patellarsenen, longitudinal	7.16	1.68	4.28	0.71	2,89	< 0.001
Patellarsenen, transversal	6.99	2.07	4.25	0.95	2,74	< 0.001

4.3. Intratester reliabilitet

I Tabel VIII og IX er resultatet af den relative og absolute reliabilitet præsenteret sammen med middelværdier og standard deviationen (SD) for patellarsenens tykkelse vurderet af testerne på symptomatiske og asymptomatiske sener.

Tabel VIII. Intratester reliabiliteten for målinger på symptomatiske sener.
Angiver afrundede værdier for relativ og absolut reliabilitet

Longitudinal	Gns.	SD	ICC3(A.1)	95%CI	ICC3(C.1)	95%CI	SEM(mm)	SDC(mm)
Tester A 1. måling	7,21	1,65	0,96	0,92-0,98	0,96	0,92-0,98	0,33	0,91
Tester A 2. måling	7,26	1,72						
Tester B 1. måling	7,06	1,83	0,95	0,89-0,98	0,95	0,90-0,98	0,39	1,08
Tester B 2. måling	7,15	1,7						
Tværsnit								
Tester A 1. måling	7,33	2,13	0,89	0,78-0,85	0,89	0,78-0,95	0,72	2,00
Tester A 2. måling	7,15	2,19						
Tester B 1. måling	6,49	2,25	0,91*	0,82-0,96	0,92*	0,83-0,96	0,67	1,86
Tester B 2. måling	6,78	2,27						

Gns. = gennemsnittet i millimeter, SD = standard deviationen, ICC = intra class correlation coefficient, 95%CI = 95% konfidensinterval, SEM = standard error of measurement i millimeter, SDC = smallest detectable change i millimeter
*Ikke signifikant systematisk shift (F -test, $p = 0.096$)

Tabel IX. Intratester reliabiliteten for målinger på asymptomatiske sener.
Angiver afrundede værdier for relativ og absolut reliabilitet

Longitudinal	Gns.	SD	ICC3(A.1)	95%CI	ICC3(C.1)	95%CI	SEM(mm)	SDC(mm)
Tester A 1. måling	4,22	0,64	0,88	0,77-0,94	0,88	0,77-0,94	0,24	0,67
Tester A 2. måling	4,31	0,73						
Tester B 1. måling	4,17	0,69	0,93	0,87-0,97	0,93	0,87-0,97	0,19	0,53
Tester B 2. måling	4,24	0,77						
Tværsnit								
Tester A 1. måling	4,35	0,89	0,87	0,76-0,93	0,87	0,76-0,93	0,33	0,91
Tester A 2. måling	4,27	0,92						
Tester B 1. måling	4,05	1,05	0,96	0,92-0,98	0,96	0,92-0,98	0,20	0,55
Tester B 2. måling	4,12	1,02						

Gns. = gennemsnittet i millimeter, SD = standard deviationen, ICC = intra class correlation coefficient, 95%CI = 95% konfidensinterval, SEM = standard error of measurement i millimeter, SDC = smallest detectable change i millimeter

4.3.1. Relativ reliabilitet

Den relative reliabilitet var gennemgående høj til meget høj for både tester A og B ved målinger af symptomatiske og asymptomatiske sener på longitudinale og transversale billeder. Hertil ses det i Tabel VIII at den relative reliabilitet for symptomatiske sener varierede fra ICC3(A.1) 0.89 - 0.96 og i Tabel IX for asymptomatiske sener ICC3(A.1) 0.87-0.96. Den relative reliabilitet var for begge testere højere for målinger af symptomatiske sener på longitudinale billeder end for asymptomatiske sener. Dertil var den gennemsnitlige spredning (SD) af målinger størst for symptomatiske sener i både det longitudinale -og transversale plan jævnfør Tabel VIII og IX.

I Tabel IX ses det, at Tester A udviste en større variation i ICC-værdier end tester B tilknyttet målinger på asymptomatiske sener. Den relative reliabilitet var i tillæg generelt højere for tester B end for tester A på tværsnitsbilleder både tilknyttet symptomatiske og asymptomatiske sener. For målinger på tværsnitsbilleder udført af tester A udgjorde den sande variationen mellem scorerne af patellarsenens tykkelse på symptomatiske sener 89% og 87% for asymptomatiske. Den sande variation mellem deltagerne udgjorde en større grad af den totale variation hos Tester B med henholdsvis 91% og 96% hos symptomatiske og asymptomatiske sener. Dette illustreres i Tabel VIII og IX.

Den relative reliabilitet for tester A var kontinuerlig høj til meget høj for målinger på longitudinale billeder kontra moderat på tværnsnitsbilleder. For tester B var den relative reliabilitet gennemgående meget høj varierende mellem ICC3(A.1) 0,91-0,96.

Den RR tilknyttet målinger på tværnsnitsbilleder fra symptomatiske sener for tester B viste tilstedeværelsen af en systematisk fejl ved forskellen mellem ICC-absolute agreement og consistency (0.92-0.91). Forskellen var imidlertid ikke signifikant (*F-test*, $p = 0.096$)

4.3.2. Absolut reliabilitet

Målefejlen kvantificeret i SEM for begge testere var mindre ved målinger af asymptomatiske sener. SEM varierede tilknyttet symptomatiske sener mellem 0.33mm – 0.72mm og tilknyttet asymptomatiske sener mellem 0.19mm – 0.33mm. Den laveste målefejl var for tester B relateret til målinger af asymptomatiske sener i et longitudinal snit (SEM 0.19mm). Altså ville et individs sande score med 68% sandsynlighed ligge mellem individets score ± 0.19 mm. Den største målefejl var tilknyttet tester A ved målinger af symptomatiske sener på tværnsnitsbilleder (SEM 0.72mm). Her ville et individs sande score med 68% sandsynlighed ligge mellem individets score ± 0.72 mm.

SDC var lavere for begge tester tilknyttet målinger af patellarsenens tykkelse på asymptomatiske sener. For symptomatiske sener varierer SDC mellem 0.91mm – 2.00mm og for asymptomatiske sener 0.53mm – 0.91mm for begge testere.

På det laveste angav SDC at en gentagen måling med 95% sandsynlighed skulle være større end ± 0.53 mm, for at den ikke kunne tilskrives målevariationen tilknyttet asymptomatiske sener i et longitudinal snit vurderet af tester B. Den højeste SDC tilknyttet Tester B var 1.86mm på tværnsnitsbilleder af symptomatiske sener.

For tester A var den laveste SDC 0.67mm relateret til asymptomatiske sener på longitudinale billeder. På det højeste skulle en gentagen måling med 95% sandsynlighed være større end ± 2.00 mm, for at den ikke kunne tilskrives målevariationen med RTUS tilknyttet symptomatiske sener på et tværnsnitsbillede vurderet af tester A.

4.4. Intertester reliabiliteten

Resultatet af intertester reliabilitetsanalysen er præsenteret i Tabel X og XI. Relativ og absolut reliabilitet er fremstillet sammen med middelværdier og SD for patellarsenens tykkelse vurderet af testerne på symptomatiske og asymptomatiske sener.

Tabel X. Intertester reliabiliteten for målinger udført på symptomatiske sener. Angiver afrundede værdier for relativ og absolut reliabilitet

Longitudinal	Gns.	SD	ICC3(A.1)	95%CI	ICC3(C.1)	95%CI	SEM(mm)	SDC(mm)
Tester A 1. måling	7,21	1,65						
Tester B 1. måling	7,06	1,83	0,95	0,92-0,98	0,96	0,92-0,98	0,37	1,03
Tester C	7,14	1,68						
Tværsnit								
Tester A 2. måling	7,21	2,19						
Tester B 2. måling	6,78	2,27	0,87	0,77-0,93	0,87	0,78-0,93	0,80	2,22
Tester C	7,14	1,68						

Gns. = gennemsnittet i millimeter, SD = standard deviationen, 95%CI = 95% konfidensinterval, ICC = intra class correlation coefficient, SEM = standard error of measurement i millimeter, SDC = smallest detectable change i millimeter

Tabel XI. Intertester reliabiliteten for målinger udført på asymptomatiske sener. Angiver afrundede værdier for relativ og absolut reliabilitet.

Longitudinal	Gns.	SD	ICC3(A.1)	95%CI	ICC3(C.1)	95%CI	SEM(mm)	SDC(mm)
Tester A 1. måling	4,22	0,64						
Tester B 1. måling	4,17	0,69	0,87	0,79-0,93	0,87	0,79-0,93	0,25	0,69
Tester C	4,24	0,77						
Tværsnit								
Tester A 1. måling	4,35	0,89						
Tester B 1. måling	4,05	1,05	0,84*	0,73-0,92	0,86*	0,77-0,94	0,39	1,08
Tester C	4,24	0,77						

Gns. = gennemsnittet i millimeter, SD = standard deviationen, 95%CI = 95% konfidensinterval, ICC = intra class correlation coefficient, SEM = standard error of measurement i millimeter, SDC = smallest detectable change i millimeter.

*Signifikant systematisk shift (*F-test*, $p = 0.003$). Tester B score systematisk lavere end tester A og C.

4.4.1. Intertester reliabiliteten

SD var størst for målinger udført på symptomatiske sener. Som det ses i tabel X og XI var den relative reliabilitet var høj til meget høj for begge tester ved vurdering af symptomatiske sener med ICC3(A.1) mellem 0.96-0.87. Ved målinger af asymptomatiske sener for begge testere er den relative reliabilitet høj med ICC3(A.1) mellem 0.84-0.87. Tabel X og XI viste videre, at den relative reliabilitet for målinger på

longitudinale billeder gennemgående var højere eller lig med målinger af patellarsenens tykkelse på tværsnitsbilleder (ICC3(A.1) $0.95-0.87 \geq 0.87-0.84$).

Relativ reliabilitet var højest med en ICC3(A.1) på 0.96 ved målinger på longitudinale billeder og laveste tilknyttet vurderingen af asymptomatiske på tværsnitsbilleder med ICC3(A.1) på 0.84. Videre var intertester reliabiliteten for denne påvirket af en signifikant ($p = 0.003$) systematisk fejl mellem testerne, hvilket ses i tabel XI. Som det ses udfra middelværdien, scorede tester B systematisk lavere end Tester A og C. Forholdet mellem målingerne kan ses på scatterplots i Bilag 6.

4.4.2. Absolut reliabilitet.

Målefejlen er kvantificeret som SEM og SDC i tabel X og XI er større vedrørende målinger af symptomatiske sener kontra asymptomatiske sener.

SEM var mindst på longitudinale billeder både for symptomatiske og asymptomatiske sener med en SEM på henholdsvis 0.37mm - 0.25mm. Tilknyttet vurderinger af tværsnitsbilleder var SEM størst for symptomatiske sener (SEM 0.80mm).

SDC var størst ved vurdering af patellarsenens tykkelse for symptomatiske sener på tværsnitsbilleder. Her skulle en gentagen måling være større end ± 2.22 mm, for at den med 95% sandsynlighed kunne tilskrives en reel forandring af patellarsenens tykkelse. For asymptomatiske sener på longitudinal billeder kunne en gentagen måling være cirka 3 gange mindre (± 0.69 mm), for at den med 95% sandsynlighed ikke kunne tilskrives målefejl.

5. Diskussion

RTUS synes med teknologiske fremskridt at være under udvikling som en tilgængelig målemetode for fysioterapeuter i primærhelsetjenesten. Specielt i et diagnostik øjemed tilknyttet muskel-sene-apparatet. Målinger af senetykkelsen med RTUS kan indikere hævelse i senen og sammen med strukturelle forandringer og hyperæmi bidrage med værdifuld indsigt i de intratendinøse forhold ofte tilknyttet PT (Abate et al., 2009; Docking et al., 2015).

I tillæg anvendes RTUS kontinuerlig i studier vedrørende PT. Dette omhandler både RTUS som målemetode tilknyttet diagnostik og klassificering, men også evaluering. Der refereres ofte til intratendinøse forandringer såsom strukturelle forandringer, vaskularitet og senetykkelse. Alligevel er reliabiliteten for disse parametre sjældent rapporteret (Black et al., 2004; McAuliffe et al., 2016; Sunding et al., 2016).

Formålet med dette studie var at beskrive intra -og intertester reliabiliteten for målinger af patellarsenens tykkelse med RTUS som målemetode på symptomatiske og asymptomatiske sener.

Resultatet af dette studie viste, at intra -og intertester reliabiliteten tilknyttet indeværende testere var meget god og bedre for posterior-anterior målinger af patellarsenens tykkelse longitudinal end transversal på symptomatiske og asymptomatiske sener. Dog var det mulig for en erfaren tester at opnå meget høj intratester reliabilitet på tværsnitsbilleder af asymptomatiske sener. Generelt var intra -og intertester reliabiliteten imidlertid lavere for målinger tilknyttet tværsnitsbilleder og befængt med større målefejl. Intertester reliabiliteten tilhørende tværsnitsbilleder af asymptomatiske sener var tilmed influeret af en signifikant systematisk målefejl.

5.1. Resultat diskussion

5.1.1. Forskel i Senetykkelse

Resultaterne fra dette studie viste at den gennemsnitlige posterior- anterior tykkelse på sonografiske billeder var signifikant tykkere for symptomatiske patellarsener sammenholdt med asymptomatiske sener. Dette var gældende for målinger både i det longitudinale og transversale snit. I det longitudinale snit var symptomatiske sener i gennemsnit 2.89(1.55) mm tykkere end asymptomatiske ($p < 0.001$). På tværsnitsbilleder var symptomatiske sener i gennemsnit 2.74(2.05) mm tykkere end asymptomatiske sener ($p < 0.001$). Imidlertid blev der ikke samlet ind demografiske variabler for vægt og højde, hvilket kan være en mulig bias med tanke på senetykkelsen. I tillæg var grupperne ikke helt homogene. Gruppen med symptomatiske sener havde en højere andel mænd (68%) og idrætsudøvere (90%) end gruppen med asymptomatiske sener (55% mænd/idrætsudøvere 55%). Selvom dette muligvis kunne tænkes at reducere senetykkelsen i den asymptomatiske gruppe har dette ingen effekt på reliabilitetsudregningerne. Alligevel var der samsvar med den gennemsnitlige

signifikante forskel mellem grupperne i denne undersøgelse og med fund fra tidligere studier med mere homogene grupper. Disse studier fandt ligeledes en signifikant gennemsnitlig forskel mellem tykkelsen på symptomatiske og asymptomatiske patellarsener (O. Lian et al., 1996; Sunding et al., 2016).

5.1.2. Intratester reliabilitet

Intratester reliabiliteten var gennemgående god for begge testere på målinger i et longitudinal snit. Relativ reliabilitet kvantificeret i ICC var meget høj for tester A og B ved målinger af symptomatiske sener i et longitudinal snit (ICC 0.96 – 0.95). For målinger af asymptomatiske sener var den relative reliabilitet mindre, men fortsat høj (ICC 0.88 for tester A til 0.93 for tester B). Målefejlen kvantificeret i SEM og SDC var mindre ved målinger udført på asymptomatiske sener i et longitudinal snit for begge testere. Generelt var intratester reliabiliteten bedre for målinger af senerne på longitudinale billeder. Dog med en undtagelse tilknyttet tester B ved tværsnitsmålinger af asymptomatiske sener (ICC 0.96, SEM 0.22mm). Specielt var målefejlen større tilknyttet tværsnitsmålinger for begge testere. SEM og SDC var cirka dobbelt så stor ved målinger af symptomatiske sener på tværsnitsbilleder kontra longitudinale. Den relative reliabilitet ved tværsnitsmålinger var høj for tester A og meget høj for tester B.

Begge testere havde lavere ICC-værdierne ved målinger på asymptomatiske sener modsat symptomatiske sener, hvilket umiddelbart stod i kontrast, til det som var forestillingen. Grundet intratendinøse forandringer tilknyttet symptomatiske sener var tanken, at den profunde del af patellarsenen kunne være vanskeligere at afgrænse fra omkringliggende væv og herved medføre mere målevariation (Bianchi & Martinoli, 2007, pp. 662-664; Docking et al., 2015; U. Fredberg et al., 2008). Imidlertid var den absolute reliabilitet kvantificeret i SEM og SDC højere for symptomatiske sener kontra asymptomatiske sener. Således var enigheden mellem gentagende målinger bedre på asymptomatiske sener, men evnen til at skelne mellem deltagerne mindre. Forklaringen på dette kan muligvis tilskrives spredningen af data (se Tabel VIII og IX). Fordelingsbredden for patellarsenens tykkelse var over dobbelt så stor hos gruppen med symptomatiske sener. Perspektiveres dette til beskrivelsen af ICC ville størrelsen af variation mellem deltagerne påvirke koefficienten, siden denne både indgik i tælleren og nævneren. Dermed påvirkede målefejlen trods større hos gruppen med symptomatiske sener ikke i samme grad reliabilitetskoefficientens evne til at

diskriminere mellem deltagerne. Dette netop fordi variationen mellem deltagerne var stor. ICC bør derfor tolkes ud fra fordelingsbredden mellem deltagerne eller mere specifikt hertil mellem senetykkelsen (de Vet et al., 2006; Weir, 2005). Således var den relative reliabilitet lidt højere tilknyttet symptomatiske sener, men også befængt med en større målefejl indenfor gentagende målinger.

På tværsnitsbilleder var den relative reliabilitet lavere og målefejlen større kontra longitudinale billeder. Blandt andet var SEM og SDC tilknyttet målinger af symptomatiske sener transversal mere end dobbelt så store kontra longitudinal for tester A. En mulig årsag til dette kunne muligvis tilskrives en tilsløret profund afgrænsning af patellarsenen på tværsnitsbilleder grundet seneformændring. Alle testerne nedfældede kommentar om tilslørede profunde seneafgrænsninger. Dette især vedrørende tværsnitsbilleder. I visse tilfælde var den profunde del af senen og den infrapatellar fedtpude næsten isoekkoid, hvilket gjorde det specielt vanskeligt at fæste et nedre målepunkt. Foreliggende blev også kommenteret tilknyttet asymptomatiske tværsnitsbilleder. Erfaring med RTUS spillede antageligvis også en rolle, idet tester B opnåede bedre intratester reliabilitet tilknyttet tværsnitsbilleder og især for asymptomatiske sener (Tester A: ICC 0.87, SDC 0.91, tester B: ICC 0.96, SDC 0.55). Tester B kunne muligvis bedre diskriminere mellem vævsstrukturene ved den profunde del af senen, jævnfør RTUS som tillagt operatør afhængighed (Docking et al., 2015; U. Fredberg et al., 2008).

5.1.3. Intertester reliabilitet

Intertester reliabilitet var bedre for målinger af symptomatiske og asymptomatiske sener i et longitudinal snit kontra et transversal snit. Den relative reliabilitet var meget høj for symptomatiske sener og høj for asymptomatiske sener på longitudinale billeder. Målefejlen var betydeligt større tilknyttet tværsnitsbilleder specielt for symptomatiske sener (symptomatiske SDC 2.22mm, asymptomatiske SDC 1.08mm). Som nævnt ovenfor under intratester reliabiliteten var det ligeledes muligt, at en isoekkoid tilsløring kunne have medført større fejlvariation mellem testerne. Det er muligt, at en erfaren operatør tilsvarende tester B muligvis kunne have opnået kvalitetsmæssige bedre sonografiske billeder (Bianchi & Martinoli, 2007). Dermed måske en tydeligere afgrænsning af vævsstrukturene ved den profunde del af senen. En vurdering af dette

kunne være opnået ved at tester B i tillæg til tester A havde indsamlet billeder for reliabilitetsanalysen.

Den bedre intertester reliabilitet ved longitudinale billeder kunne skyldes muligheden for ossøse holdepunkter til afgivelsen af et konstant målepunkt. På longitudinale billeder blev der målt 0.5 cm i distal retning fra apex patella. For tværsnitsbilleder var det ikke mulig at anvende ossøse holdepunkter. Målepunktet var derfor baseret på, hvor midten af senen visuelt fremkom samt den nedre profunde afgrænsning. Dette har formentlig krævet en større grad af subjektiv vurdering fra testerne. Herved muligvis bevirket større variationen mellem testerne og deltagerne. SD var således også større tilknyttet patellarsenens tykkelse for alle testerne ved målinger på tværsnitsbilleder. At målinger på tværsnitsbilleder kunne medføre en større variation mellem testerne, kunne muligvis være en forklaring på den systematisk fejl mellem testerne tilknyttet tværsnitsbilleder af asymptomatiske sener. Tester B score systematisk lavere end Tester A og C. Hertil var ICC-consistency (0.86, 95%CI 0.77-0.94) større end ICC-absolute agreement (0.84, 95%CI 0.73-0.92). Og da absolute agreement indbefattede systematisk variation relateret til testerne, var der en systematisk forskel mellem testerne. Videre viste F-testen fra ANOVA-analysen en signifikant systematisk målefejl mellem testerne ($p = 0.003$). Det kan betyde, at måleproceduren for måling af tværsnitsbilleder ikke var standardiseret tilstrækkeligt, at tester A ikke havde nok erfaring til at indsamle gode nok billeder af tværsnittet eller, at patellarsenen i et transversal snit er strukturelt vanskeligere at måle blandt andet grundet få anatomiske landemærker. Fredberg et al (2008) fandt ligeledes, at målinger på tværsnitsbilleder var vanskeligere at reproducere selv for en tester specialiseret i RTUS. De konkluderede med at målinger af patellarsenens tykkelse burde foregå med senen i et longitudinal snit, hvor afstanden fra patella til målepunktet kunne anføres (U. Fredberg et al., 2008). Dette er ikke mulig på tværsnitsbilleder.

5.1.4. Diskussion af eksisterende litteratur

Få studier har så langt undersøgt intra -og intertester reliabiliteten tilknyttet målinger patellarsenens tykkelse på symptomatiske og asymptomatiske sener. Sunding et al (2016) undersøgte tilsvarende dette studie intra -og intertester reliabiliteten for målinger af patellarsenens tykkelse med RTUS på longitudinale billeder. Analog til dette studie blev målingerne udført på indsamlede sonografiske billeder. I alt indbefattede studiet

fire testere og 26 patellarsener. Testerne i studiet havde forskellig grad af erfaring med RTUS samt klinisk erfaring. De fandt i samsvar med denne undersøgelse en meget god intra -og intertester reliabilitet uafhængigt af erfaringsniveauet for målinger af PST i et longitudinal snit. Den relative reliabilitet var ved tilsvarende indeværende studie generelt meget høj (ICC 0.99). Målefejlen var meget lav og blev kvantificeret i CV og varierede mellem CV 1.6% - 2.7%.

En af styrkerne ved studiet af Sunding et al (2016) var kvantificering af målefejl og relativ reliabilitet samt en velbeskrevet fremgangsmåde for måling. Samtidig undersøgte de intertester reliabiliteten for to testere tilknyttet hele proceduren inkluderet skanning og fandt en samsvarende intertester reliabilitet. Erfaringsniveauet og baggrunden for testerne var ligeledes godt beskrevet. Alligevel var der flere svagheder ved studiet. Der indgik fire testere i studiet, men intertester reliabiliteten blev ikke testet for alle fire tester sammenlagt men i par af to. Havde de testet intertester reliabiliteten sammenlagt ville resultatet have været mere robust. Dette fordi det ville have inkluderet flere målinger og testere. Samtidig havde testerne forskellig professions baggrund, med forskellige specialer, de var derfor næppe repræsentative for populationen af testere. Selve proceduren for hvorledes målinger af tykkelse blev udført var ikke udførligt beskrevet, hvilket kan gøre det besværligt at overføre metoden til andre studier. Endelig indgik symptomatiske og asymptomatiske individer i en samlet gruppe for analyse. Der var således kun 13 symptomatiske individer inkluderet. Dette kunne medføre en større SD. Yderligere ville en eventuelt større målefejl tilknyttet asymptomatiske sener formentlig være blevet udjævnet af en mindre målefejl tilknyttet asymptomatiske sener. Således heteroscedastiske data. Det kan derfor diskuteres i hvilken grad reliabiliteten præsenterer målinger tilknyttet PT (Sunding et al., 2016).

I et ældre studie af Black et al (2004) søgte de at belyse intertester reliabiliteten for to sammenlignbare radiologer med speciale i RTUS og konkluderede med at intertester reliabiliteten var høj. Studiet omhandlede hele skanningsproceduren og ikke kun måling på billeder. Otte ud af tolv individer havde PT i studiet. De fandt i modsætning til dette studie bedre intertester reliabilitet ved måling af den P-A tykkelsen på tværsnitsbilleder end ved longitudinale målinger. Dårligere reliabilitet på målinger af patellarsenen i et longitudinal snit kunne muligvis skyldes manglende brug af standardiserede ossøse

holdepunkter. Dette kunne tænkes skabende mere fejlvariation tilknyttet målingerne. Specielt med tanke på patellarsenens kegleformede figur. Et andet aspekt kunne være at tester var specialiserede radiologer og derved opnåede bedre reliabilitet på tværsnitbilleder. Aktuelt kunne dette tilknyttet deres erfaring medføre en tydeligere afgræsning af senens profunde del og således klarere målepunkter (Black et al., 2004).

Studiet til Black et al (2004) er imidlertid ikke direkte komparabel med indeværende undersøgelse. Først og fremmest indbefatter studiet hele skanningsprocedure kontra kun sonografiske billeder i dette studie. Reliabiliteten er således eksponeret for variation tilknyttet positionering af deltager samt skanningen udført af testerne, hvilket ikke er tilfældet i indeværende studie. Samtidig har Black et al anvendt korrelationskoefficienten Pearsons r som udtryk for den relative reliabilitet modsat ICC i denne undersøgelse. Dette er en svaghed idet Pearsons r udelukkende bringer information om den lineære sammenhæng mellem to variabler og ikke enigheden imellem. Således vil systematisk fejl tilknyttet testerne ikke udtrykkes i Pearsons r modsat ICC (Streiner et al., 2015, pp. 172-173). I studiet til Black et al fremstod målefejlen umiddelbart stor på longitudinale billeder med en øvre LOA på 3.6 mm samtidig med at den gennemsnitlige forskel mellem testerne næsten var signifikant ($p = 0.07$). Videre var reliabiliteten udført på en gruppe bestående af både symptomatiske og asymptomatisk, hvilket muligvis kunne medføre en større fordelingsbredde. Det var således muligt at den høje korrelationskoefficient tilknyttet longitudinale billeder var påvirket af fordelingsbredden og det begrænsede antal af deltager (Black et al., 2004).

Tilknyttet friske sener ($n = 35$) fandt Fredberg et al (2008) en acceptabel intratester reliabilitet for en tester. De fandt i association til denne undersøgelse ligeledes en lavere intertester variation for longitudinale skanninger end for transversale skanninger. I tillæg anvendte de samme ossøse holdepunkter for målinger i det longitudinale snit som denne undersøgelse. Selvom studiet af Fredberg et al indbefattede reliabilitet af hele skanningsprocessen var den sammenlignbar med indeværende undersøgelse i det tester A indsamlede alle de sonografiske billeder og efterfølgende indgik i intra -og intertester analysen. Imidlertid var det kun den absolute reliabilitet, der blev belyst ved CV studiet af Fredberg et al. Reliabilitet kunne således kun tilknyttes målemetoden i et evaluerende perspektiv og ikke diagnostisk, idet den absolute reliabilitet som beskrevet af de Vet et

al (2006) ikke informere om metodens evne til at diskriminere mellem deltagerne (de Vet et al., 2006; U. Fredberg et al., 2008).

Skou & Aalkjaer (2013) undersøgte ligeledes intra -og intertester reliabiliteten tilknyttet longitudinale måling af PST dog kun på asymptomatiske sener. De fandt i modsætning til denne undersøgelse en lav til moderat relativ reliabilitet både for intra -og intertester reliabilitet. Målefejlen var ligeledes større i dette studie både for intra -og intertester reliabiliteten (95% LOA \pm 0.9mm). Dette kunne muligvis tillægges at reliabiliteten også inkluderede variation tilknyttet skanning og ikke kun variation relateret til selve det at sætte målinger på sonografiske billeder. Skanning blev udført af to fysioterapeuter specialiseret i brugen af RTUS. Dog viste studiet god reliabilitet ved gennemsnittet af to målinger. En styrke ved dette studie var at måleproceduren samt standardiseringsgraden var udførlig beskrevet og tilsvarende standardisering i indeværende undersøgelse. Reliabiliteten i studiet var imidlertid kun baseret på 18 friske sener og to testere og generaliserbarheden af studier syntes dermed begrænset. Flere deltager eller testere/gentagende målinger kunne muligvis have udjævnet målefejlen. (Skou & Aalkjaer, 2013).

5.1.5. Opsummering

Gennemgående var den relative reliabilitet som et parameter for intra -og intertester reliabilitet i indeværende undersøgelse god og i samsvar med lignende studier (Black et al., 2004; Sunding et al., 2016). I dette studie var der i samsvar med en tidligere undersøgelse en indikation for, at brugen af ossøse holdepunkter på longitudinale billeder var en medvirkende årsag til en mindre intra -og intertester variabilitet (U. Fredberg et al., 2008). Videre var patellarsenens tykkelse for symptomatiske sener signifikant tykkere end asymptomatiske sener i samsvar med tidligere studier (O. Lian et al., 1996; Sunding et al., 2016).

Indeværende studie viste ligeledes, at det var mulig for to fysioterapeuter med basal oplæring i RTUS at opnå god reliabilitet. Generelt var der ikke ingen forskel på reliabilitet mellem en erfaren tester og en novice på målinger af longitudinale billeder. Målinger på tværsnitsbilleder indikerede imidlertid, at intratester reliabiliteten var lidt bedre for en erfaren tester.

5.2. Valg af statistiske parametre

Reliabilitet består som nævnt i afsnit 2.5. af reliabilitet og målefejl som kan operationaliseres i relativ og absolut reliabilitet, der igen kan kvantificeres i forskellige statistiske parametre. Det som studeres er variation mellem målinger. Nærmere betegnet hvordan influere variation på evnen til opnå samme resultat ved gentagende eller flere målinger (Hopkins, 2000; Mokkink et al., 2010a, 2010b). De Vet et al (2006) fremhævede vigtigheden af både at benytte parametre for reliabilitet og målefejl. Egenskaberne er relateret til hinanden, men fortæller forskellige ting (de Vet et al., 2006). Alligevel blev det valgt i flere af de omtalte studier i denne opgave, kun at belyse et af de måletekniske egenskaber og således kun et aspekt af reliabilitet (Bjordal et al., 2003; U. Fredberg et al., 2008).

5.2.1. Relativ reliabilitet

ICC3 (A.1) blev anvendt i dette studie. Dette implicerede at absolute agreement var indbefattet. Forstået som systematisk forskel mellem testerne. Årsagen til dette var, at patellarsenens tykkelse som beskrevet med RTUS kan indgå i et diagnostisk regi, og derfor var det ønskelig, at opnå viden om hvorvidt testerne var enige om graden af senetykkelse. Information om rangeringen af senetykkelsen alene som ved consistency repræsenterer ikke den reelle værdi, da den systematisk fejl er undladt. Alligevel indgik ICC3 (C.1) som et statistisk parameter. Da denne ikke indeholdte systematisk fejl kunne det ved sammenligning af disse to, udledes hvorvidt der var en systematisk fejl, tilstede (de Vet et al., 2006). I dette studie var den systematisk fejl generelt begrænset. Dette kom til udtryk ved ingen eller en minimal forskel mellem ICC-værdierne. Dog med undtagelse af intertester reliabiliteten tilknyttet tværsnitsbilleder af asymptomatiske individer i Tabel XI. Her var det synlig at ICC3 (C.1) havde en højere ICC-værdi og et mindre sikkerhedsinterval end ICC3 (A.1). Siden systematisk variation mellem testerne ikke var med i ICC3 (C.1) blev denne højere. Jo større forskel des større er variationskomponenten for systematisk fejl tilknyttet testerne. Ved inklusion af både consistency og absolute agreement opstod der en hvis fornemmelse af størrelsen på den systematisk fejl. SEM og SDC bidrog ikke med information om dette da disse informerede om målefejlen som helhed. Inddragelse af consistency og absolute agreement fremstod således som en fordel i dette studiedesign, idet det vil være muligt

at analysere mulige kilder til den systematiske fejl og herefter søge at kontrollere disse. Dette lader sig ofte ikke gøre for tilfældig fejl (de Vet et al., 2006; Weir, 2005).

5.2.2. Absolut reliabilitet

Absolut reliabilitet blev kvantificeret som SEM og SDC i denne undersøgelse. RTUS blev i flere studier anvendt som et evaluerende måleinstrument og i den forbindelse er parametre for den absolutte reliabilitet helt essentielt. Information om hvor stor en ændring der må til for at den ikke kan tilskrives målefejl er af stor betydning i et evalueringsperspektiv. Alligevel er information som målefejlen frembringer kun anvendelig i det perspektiv at der foreligger en klar forståelse af hvor stor en forskel som kan betragtes vigtig. Der foreligger således ikke nogen gradering af hvornår målefejlen er høj eller lav, dette er tilknyttet den aktuelle genstand som skal måles (de Vet et al., 2006).

SEM kan som ICC indlemme systematisk fejl eller kun tilfældig fejl. I dette studie blev SEM baseret på absolute agreement og udregnet ud fra variationskomponenterne. Siden SEM kan være baseret på consistency og absolute agreement, er det essentielt, at det er beskrevet, hvordan den blev udregnet. SEM kun baseret på consistency vil være mindre, idet den ikke repræsenterer målefejl tilknyttet systematisk fejl mellem testere. Både SEM og SDC blev præsenteret, hvilket blev anset som en fordel. Dette fordi SEM angav hvor meget målevariation som kan forventes ved en måling, og SDC hvor meget målevariation der kan forventes ved gentagende målinger. Imidlertid er SDC baseret på et 95% sikkerhedsinterval, modsat SEM som er baseret på et 68% sikkerhedsinterval. Således kan SDC ikke i samme grad som SEM opfange mindre relevante kliniske forandringer. Alligevel kan sikkerhedsintervallet tilknyttet SDC ændres ved at indsætte en anden z-score i udregning som beskrevet i afsnit 2.6.4.

5.3. Metode diskussion

5.3.1. Udvalget

Antallet af deltagere vurderes som en styrke ved dette studie. Selvom målet i forkant af rekrutteringen var at inkludere nærmere 50 symptomatiske og 50 asymptomatiske sener, anses udvalget på 30 symptomatiske og 33 asymptomatiske sener som en styrke ved studiet sammenholdt med tidligere studier (Black et al., 2004; Sunding et al., 2016). Et udvalg på nærmere 50 sener blev valgt grundet de Vet et al (2011) sine anbefalinger om

minimum 50 deltager og to gentagende målinger/testere (de Vet et al., 2011, p. 126) Hertil har Streiner et al (2015) beskrevet at over 50 deltager i mange tilfælde vil være en statistisk overdrivelse (Streiner et al., 2015, p. 192).

Sunding et al (2016) opnåede med færre sener (n = 6, 13PT/13Friske) høje ICC- værdier med smalle sikkerhedsintervaller (Sunding et al., 2016). Sikkerhedsintervallet vil variere med størrelse af udvalget og SD. I indeværende studie var størrelse på sikkerhedsintervallet relativt fornuftigt. Det vil sige der var ikke et voldsomt stort spring i sikkerhedsintervallerne, hvor det med 95% sikkerhed kunne forventes, at den sande ICC lå (95% CI 0.73-0.92 til 0.92-0.98). Således fremstod udvalget stort nok og effektiv nok til at takle spredningen i dataene også selvom SD var cirka dobbelt så stor for symptomatiske sener. Dette var en styrke af studiets interne validitet (Norman & Streiner, 2014, pp. 62-63). Samtidig er udvalget af en størrelse så enkelte ekstreme værdier i mindre grad påvirker reliabilitetskoefficienten. Et stort udvalg producerer flere værdier og normalt en mere jævn spredning samt højere reliabilitetskoefficienter (Carter & Lubinsky, 2016, p. 252).

Udvalget kan påvirke generaliserbarheden til testpopulationen og dermed studiets eksterne validitet. Resultatet af dette studie kan derfor kun generaliseres til den type af individer, der indgår i udvalget som studeres (de Vet et al., 2006; Streiner et al., 2015, p. 164). Således individer med PT og friske individer fra 18 år og over. PT er en klinisk diagnose og derfor er populationen af individer over 18 år kendetegnet ud fra de kliniske kardinalsymptomer for PT. Inklusions og eksklusions kriterierne i denne undersøgelse bestod primært af kliniske kriterier høstet fra en grundig kliniske undersøgelse. Der findes så langt ingen gylden standardtest for diagnosticering af PT. Den kliniske definition af PT som beskrevet i afsnit 3.3.2. er derimod veldefineret i litteraturen og kontinuerligt anvendt som inklusionskriterier i forskning omhandlende PT samt i lignende reliabilitetsstudier (Comin et al., 2013; Sunding et al., 2016; Visnes et al., 2015; Warden et al., 2007).

Det vurderes derfor som en styrke for studiet eksterne validitet at der blev udført en grundig klinisk undersøgelse med hensyn til identificering af kliniske kardinalsymptomer for PT, men også med hensyn til differentialdiagnoser. Specielt diagnoser og tilstande som kunne medføre anterior knæsmerter. Som det fremgik, af

figur 14 i afsnit 4. blev flere netop ekskluderet grundet differentialdiagnoser. En styrke ved inklusionsprocessen var at tester A som udførte inklusionsundersøgelse havde flere års klinisk erfaring indenfor udredning og diagnosticering af overbelastningsrelaterede muskuloskeletale lidelser. Alligevel var det en mulig bias at tester A udførte hele inklusionsundersøgelsen inkluderet skanningen. Således var tester A potentielt ikke blindet, og kunne vurderingsmæssigt reelt være blevet påvirket af eventuelle fund på RTUS. Dette blev søgt kontrolleret. Den kliniske undersøgelse blev udført først og resultatet nedfældet i et dataskema. Derefter blev skanning med RTUS eksekveret. Alligevel ville det optimale have været at en udførte den kliniske undersøgelse og en anden skanningen adskilt og uafhængig fra hinanden. Dette var ikke mulig grundet ressourcer og tidsperspektivet.

At deltagerne blev rekrutteret som et bekvemmelighedsudvalg er mindre robust. Dette fordi risikoen for at deltagerne reelt repræsenterer populationen er mindre modsat et tilfældig udvalg af populationen (Carter & Lubinsky, 2016). I Tabel VI ses det at gruppen med PT bestod af flere idrætsudøvere end gruppen med friske individer, hvilket kan svække den interne og eksterne validitet (Carter & Lubinsky, 2016, pp. 92-103). Imidlertid var de andre variabler tilnærmelsesvis ligeligt fordelt. Samtidig fremstod den gennemsnitlige senetykkelse for begge grupper heller ikke påvirket af dette, i det disse var sammenlignelig med gennemsnitlige senetykkelser fra andre studier med tilsvarende udvalg (Kongsgaard et al., 2009; Skou & Aalkjaer, 2013; Sunding et al., 2016). Videre blev alle deltagerne inkluderet fra flere forskellige idrætsdiscipliner/klubber, primærhelsetjenesten og bekendte. Derfor blev både deltager med PT og friske i indeværende studie grundlæggende vurderet som et repræsentativt udvalg af populationen.

5.3.2. Testerne

Testerne i studiet kunne siges alle at tilhøre populationen fysioterapeuter med basal kundskab inden for anatomi og fysiologi samt basal oplæring i RTUS som minimum. Imidlertid var erfarings- og kompetenceniveauet tilknyttet RTUS ikke ens for testerne. Tester B havde lang erfaring og dermed mere kundskab indenfor RTUS. Således ville testernes forforståelse tilknyttet analytiske og vurderingsmæssige egenskaber af RTUS-billederne formentlig være forskellige og det kunne tænkes at tester B grundet erfaringen ville opnå bedre reliabilitet. Dette kunne være tilfældet vedrørende intratester

reliabiliteten tilknyttet tværsnitsbilleder for tester B specielt for asymptomatiske sener. Her opnåede tester B højere reliabilitet. Imidlertid blev alle testerne eksponeret for samme kvalitetsmæssige billeder, hvilket formentlig udjævnede erfaringsniveauet nået. Foreliggende kunne være en forklaring på at resultaterne for reliabiliteten overordnet ikke var nævneværdig forskellig mellem testerne med undtagelse af ovenfornævnte eksempel. Det er forenelig med fund i et andet studie som også anvendte tester med forskellig erfaringsniveau ved måling af patellarsenens tykkelse med RTUS (Sunding et al., 2016).

I dette studie blev der kun inkluderet to tester i intratester analysen og tre i intertester analysen. Dette svækker studiets eksterne validitet. Selvom flere studier har generaliseret resultater baseret på få tester syntes generaliserbarheden begrænset, hvilket netop præciseres af Van Genderen et al (2003). De argumenterede netop for at testerne ofte negligeres og reliabiliteten hyppigt generaliseres med baggrund i få inkluderet testere eller gentagende målinger (Van Genderen et al., 2003). Der kan derfor diskuteres hvorfor to tester skulle kunne repræsentere populationen af testere, når to deltager ikke kan. Alligevel er praksis i litteraturen at der tilknyttet intra-og intertester reliabilitet for målinger af patellarsenens tykkelse med RTUS anvendes mellem en og fire tester (Bjordal et al., 2003; Black et al., 2004; Mc Auliffe et al., 2016; Sunding et al., 2016). Dette er sammenlignbar med indeværende studie, hvilket sammen med de nævnte studier samlet kan bidrage til en at styrke den samlede eksterne validitet.

5.3.3. Indsamling af skanningsbilleder

Grundet praktiske forhold og begrænsede ressourcer blev indsamlingen af skanningsbilleder kun udført af en tester. Tester A med basal oplæring i RTUS udførte skanningerne. Et af kritikpunkterne til denne undersøgelse er netop, at der kun blev indsamlet materiale af en tester. Det medførte at reliabilitetsanalysen var baseret på kvaliteten af de sonografiske billeder som tester A producerede samt erfaringen tilknyttet tester A. Det var oplevelsen under skanningsprocessen at den profunde seneafgrænsning var svære at fremhæve i god kvalitet i et transversal snit kontra longitudinal. Dette kunne være en medvirkende årsag til at målefejlen stort set var større på målinger af tværsnitsbilleder i dette studie. Fredberg et al (2008) fandt lignede resultat og argumenterede ligeledes for at skanning transversal var mere operatørafhængig specielt tilknyttet anvendelsen af transduceren. Dette fordi

transduceren longitudinal og transversal kunne tiltes medial og lateral, men i det transversale plan kunne transduceren endvidere tiltes proksimal og distal, hvilket øgte risikoen for artefakter (U. Fredberg et al., 2008). Det var også oplevelse at det var lettere at holde transduceren i en 90 graders vinkel til senen longitudinal idet hele transduceren havde kontakt med hudoverfladen. Derimod havde hele transduceren ikke kontakt til hudoverfladen grundet knæets konvekse overfalde i det transversale plan. Det er muligt at dette var en indirekte medvirkende faktor til den signifikante systematiske fejl forbundet med intertester reliabiliteten for tværsnitsbilleder af asymptomatiske sener.

Havde tester B i tillæg til tester A indsamlet billeder fra de samme deltager ville dette have styrket studiets interne validitet. Dette ville have informeret om hvorvidt fejlvariationen var afhængig af billedkvaliteten associeret til tester A.

Til gengæld var det en styrke af designet at tester A gennemførte flere pilotskanninger af bekendte i forkant med henblik på optimalisering af sine færdigheder. Samtidig havde tester A formentlig en stigende læringskurve under indsamlingsfasen. Samlet set blev der i indsamlingsfasen indsamlet rundt 328 longitudinale og transversale billeder. Det gjorde i tillæg designet mere robust, at der blev anvendt en standardiseret positionering af deltager målrettet mod at undgå artefakter. Denne er tidligere anvendt forbundet med lav intratester variabilitet ved målinger asymptomatiske sener (Bjordal et al., 2003).

10 Individuer blev skannet i forbindelse med deres idrætsaktivitet. Det havde dog ingen effekt på reliabilitetsundersøgelse, idet skanningerne blev udført af samme tester følgende akkurat samme protokol og metode.

5.3.4. Målinger af patellarsenen

Intra -og intertester reliabiliteten i dette studie var baseret på variationen mellem gentagende målinger på samme billeder og således ikke processen med gentagende skanninger og målinger af deltagerne. Dette medførte at variation tilknyttet testerne under udførelse af skanning var kontrolleret. Således var variation vedrørende transducer placering samt angående nøjagtigheden af hvor patellarsenen blev afbilledet kontrolleret. Dette var ikke nødvendigvis en svaghed idet dokumentation af

reliabiliteten medførte en større afspejling af reliabiliteten for selve det at sætte målinger på patellarsenen variation hertil. Altså variationen tilknyttet måleproceduren for måling af senens posterior-anterior diameter.

En af de absolutte styrker ved dette studie var standardiseringsprotokollen tilknyttet målinger af patellarsenen. Specielt henvisende til måleproceduren af senen i det longitudinale snit. Selvom den profunde seneafgrænsning også kunne forekomme tilsløret her, medførte målepræferencen tilknyttet apex patella som et ossøs holdepunkt muligvis et mere konstant målepunkt. Dette kunne have været en medvirkende årsag til at målefejlen gennemgående var mindre ved målinger af patellarsenens tykkelse på longitudinale billeder. Samme måleproceduren blev benyttet af Fredberg et al (2008) og viste i samsvar med dette studie mindre målefejl tilknyttet longitudinale billeder. Samtidig var den delvise standardisering tilhørende indeværende måleprocedure af en grad som kan implementeres i klinikken og anden forskning (Carter & Lubinsky, 2016, pp. 250-251).

I forbindelse med målingerne af sener blev tester A blindet og en tilfældig fordelingsprocedure af rækkefølgen, som billederne blev målt i forekom. Således kunne tester A ikke sammenholde rækkefølgen på skanningerne med karakteristika fra billederne visualiseret under indsamlingen af billeder. I tillæg gik der næsten seks måneder fra første skanning til første målingsseance, hvilket også minimerede bias tilknyttet erindring. Tester B havde ingen relation til indsamlingen af billeder og blev derfor ikke blindet. Derfor blev der indført et intratester interval på minimum fire uger. Således blev bias tilknyttet erindring af billeder for tester B ved retest søgt kontrolleret. Generelt var forsøgsprotokollen tilhørende målinger af patellarsenens tykkelse en klar styrkelse af designets interne troværdighed.

5.3.5 Opsummering

Grundlæggende var intra -og intertester reliabiliteten forankret i et robust og troværdigt design bestående af et anseligt udvalg af deltagere. Hele forsøgsprotokollen var baseret på en analyse af mulige kilder til variation. Fremtrædende kilder til variation blev søgt kontrolleret ved en delvis standardisering af hele forsøgsprotokollen baseret på elementer fra tidligere forskning (Bjordal et al., 2003; U. Fredberg et al., 2008). Herunder anvendelse af ossøse holdepunkter ved scoring af sonografiske billeder.

Tester B og C var uafhængig indsamlingen af billeder og tester A blev blindet og udsat for en tilfældig fordelingsprocedure. Endelig blev alle deltagerne krydset med alle testerne. Dette tillod adskillelse af den systematiske kilde til variation tilknyttet testerne og således information om i hvilken grad målefejlen var påvirket af systematisk eller tilfældig fejl.

Det ville styrke studiets eksterne validitet, hvis flere testere havde været inkluderet i studiet. Desuden ville indsamling af sonografiske billeder udført af to testere antageligvis have skabt mere information om billede kvalitets indflydelse på variationen. Dette tilknyttet to uafhængige og adskilte undersøgere under inklusionsundersøgelserne vil have styrket studiet interne validitet. Tester A var projektansvarlig og stod samtidig for indsamling af billeder, udførte målinger og dataanalysen. Tester A var således påvirket af den forforståelse, blev opnået igennem studiet frem til dataanalysen. Selvom de indsamlede data ikke blev analyseret, før alt datamateriale var færdig indsamlet, ville det have styrket den eksterne validitet, hvis en uafhængig person havde udført enten skanning eller dataanalysen.

5.4. Kliniske implikationer

For fysioterapeuter i daglig klinisk praksis kan RTUS vise sig som et supplement i diagnosticering og klassificering. Der findes ingen gylden standardtest for diagnosticering af PT og den kliniske undersøgelse består af relativ for diagnostiske test rettet mod PT (Docking et al., 2015). Selvom om det kliniske billede ofte fremstår tydelig for PT, kan der opstå tvivlstilfælde rundt årsagen til de anterior knæsmærter. Her kan RTUS være et element i udredning til at skille symptomatiske sener fra asymptomatiske sener og således individer med PT fra friske individer. RTUS er vist at have en sensitiv evne til at identificere individer med klinisk diagnosticeret PT. Og en god evne til at identificere asymptomatiske sener (Warden et al., 2007). Hvis RTUS i tillæg etablerer sig som et reliabel måleinstrument tilknyttet målinger af symptomatiske og asymptomatiske sener, kan RTUS indgå som et diagnostisk og evaluerende element i håndtering af PT. For praktiserende fysioterapeuter og forskere medfører dette muligheden for at udføre præcise målinger af patellarsenen.

Parameteret relativ reliabilitet informerer om hvor gode målingerne er til at skelne patienterne fra hinanden. Således relevant i et diagnostisk perspektiv (de Vet et al.,

2006). Metodelitteraturen henviser ofte til et stringent grundlag for reliabiliteten på individniveau og et mindre stringent grundlag på gruppeniveau. Typisk ICC-værdier >0.90 for målemetoder som skal anvendes på individniveau og >0.70 på gruppeniveau (Beyer et al., 2012, p. 58; Streiner et al., 2015, pp. 183-184). Potentielt ud fra indeværende undersøgelse kan målinger af patellarsenens tykkelse på symptomatiske og asymptomatiske sener i et longitudinal snit med en ICC på op til $0.96(95\%CI\ 0.92-0.98)$ implicere, at ultralydsmålingerne er tilstrækkelig reliabel til brug på individ niveau i klinikken. Dette både aktuelt hvis målingerne udføres af samme tester eller forskellige testere. Ligeledes gør dette sig gældende for målinger udført af samme erfarne tester på tværnsnitsbilleder af asymptomatiske sener. Generelt tilknyttet målinger på tværnsnitsbilleder af symptomatiske og asymptomatiske sener implicerer dette, at disse er tilstrækkelig reliabel til at anvendes på gruppeniveau typisk tilknyttet forskning.

Selvom målinger fremstår som effektive til at skelne patienterne fra hinanden til trods for tilstedeværelsen af målefejl fortæller ikke nået om måleinstrumentets evaluerende evner. Her spiller målefejlen ind som informerer noget om hvor meget målefejl, der kan forventes (de Vet et al., 2006).

Kongsgaard et al (2013) fandt en signifikant reduktion i den posterior-anterior tykkelse af patellarsenen longitudinal på RTUS efter en 12 ugers intervention med tung langsom styrketræning. Efter en 12 ugers intervention var den gennemsnitlige reduktion af senetykkelsen 1.2mm . Betragtes dette i perspektiv af målefejlen frembragt i dette studie tilknyttet målinger af symptomatiske sener longitudinal, implicerer dette, at forskellen repræsenterer en reel ændring. I indeværende studie var SDC for flere testere 1.03mm tilknyttet symptomatiske sener i et longitudinal snit på individniveau. Beregnes SDC på gruppeniveau ($1.03\text{mm}/\sqrt{n = 30}$) bliver SDC 0.19mm . Dette implicerer at den gennemsnitlige reduktion af senetykkelse på 1.2mm med 95% sikkerhed både på individ -og gruppeniveau ikke kan tilskrives målefejl. Udføres de gentagende målinger af samme tester bliver SDC mindre ($SDC_{individ} 0.91\text{mm}$ og $SDC_{gruppe} 0.17\text{mm}$). Foreliggende er i tillæg baseret på, at Kongsgaard et al (2013) anvendte samme måleprocedure som indeværende studie for måling af senetykkelsen i et longitudinal snit (Kongsgaard et al., 2009).

Imidlertid kan RTUS ikke erstatte en grundig kliniske undersøgelse. Associationer og dissociationer eksisterer mellem seneforandringer på RTUS og symptomer forbundet med PT (Drew et al., 2014; McAuliffe et al., 2016). RTUS bør derfor, kun anses som et delelement i diagnosticering, klassificering og evaluering. En mulig årsag til en manglende entydig association kan skyldes en manglende gylden standard for diagnosticering af PT (J L Cook et al., 2016). Både kliniske men også tilknyttet målinger af seneforandringer som fremkommer på RTUS. Der er således brug for mere forskning med fokus på diagnosticeringen og evalueringen af PT.

6. Konklusion

Hensigten med dette studie var den, at bestemme intra -og intertester reliabiliteten ved målinger af patellarsenens posterior-anterior tykkelse på ultralydsbilleder hos en gruppe individer med symptomatiske patellarsener og en gruppe individer med asymptomatiske patellarsener.

Fysioterapeuter uden lang erfaring kan opnå høj intra -og intertester reliabilitet ved målinger af patellarsenens tykkelse med RTUS. Resultatet viste, at intra -og intertester reliabiliteten tilknyttet indeværende testere var meget god og bedre for posterior-anterior målinger af patellarsenen longitudinal end transversal på symptomatiske og asymptomatiske sener. Dog var det mulig for en erfaren tester at opnå meget høj intratester reliabilitet på tværsnitsbilleder af asymptomatiske sener. Generelt var intra -og intertester reliabiliteten imidlertid lavere for målinger tilknyttet tværsnitsbilleder og befængt med større målefejl.

Ud fra denne undersøgelse bør den posterior-anterior tykkelse til symptomatisk og asymptomatiske patellarsener måles med senen i et longitudinal snit, hvor afstanden fra apex patella til det første målepunkt er fastlagt. Her opnås de mest pålidelige og reproducerbare målinger af patellarsenens tykkelse. Udføres målinger af patellarsenens tykkelse på tværsnitsbilleder, bør dette udføres af den samme tester, specielt hvis målingerne skal sammenlignes over tid. Individer med patellar tendinopati havde i tillæg signifikant tykkere sener end friske individer. Dette skal dog tolkes med forbehold om mulig heterogenitet mellem grupperne.

Resultaterne i indeværende studie skal alligevel tolkes med den betingelse, at resultatet er baseret på intra -og intertester reliabiliteten for kun tre fysioterapeuter med forskellig niveau af erfaring. Studiets generaliserbarhed er derfor begrænset til indeværende fysioterapeuter.

Yderlig forskningen af intra -og intertester reliabiliteten er nødvendig tilknyttet RTUS og målinger af patellarsenen for at øge generaliserbarheden. Fremtidige studier bør inkludere betydelig flere testere og gentagende målinger end hidtil anvendt i litteraturen. Der kan med fordel indgå grupper af fysioterapeuter med forskellig grad af erfaring, for at favne hele populationen af fysioterapeuter som anvender RTUS. Samtidig bør flere fysioterapeuter med forskellig grad af erfaring indsamle sonografiske billeder uafhængig af testerne. Således indlemmes hele processen tilhørende ultralydsskanning af patellarsenen i reliabilitetsanalysen.

7. Litteraturliste

- Abate, M., Silbernagel, K. G., Siljeholm, C., Di Iorio, A., De Amicis, D., Salini, V., . . . Paganelli, R. (2009). Pathogenesis of tendinopathies: inflammation or degeneration? *Arthritis Res Ther*, *11*(3), 235. doi:10.1186/ar2723
- Albers, I. S., Zwerver, J., Diercks, R. L., Dekker, J. H., & Van den Akker-Scheek, I. (2016). Incidence and prevalence of lower extremity tendinopathy in a Dutch general practice population: a cross sectional study. *BMC Musculoskeletal Disorders*, *17*(1), 16. doi:10.1186/s12891-016-0885-2
- Association, W. M. (1964, 2013). WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. Retrieved from <http://www.wma.net/en/30publications/10policies/b3/>
- Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med*, *26*(4), 217-238.
- Beyer, N., Magnusson, P., & Thorborg, K. (2012). *Målemetoder i forebyggelse, behandling og rehabilitering : teori og anvendelse* (2.udg. ed.). København: Munksgaard.
- Bianchi, S., & Martinoli, C. (2007). *Ultrasound of the Musculoskeletal System*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Birkler, J. (2005). *Videnskabsteori: en grundbog*. København: Munksgaard Danmark.
- Bjordal, J. M., Demmink, J. H., & Ljunggren, A. E. (2003). Tendon Thickness and Depth from Skin for Supraspinatus, Common Wrist and Finger Extensors, Patellar and Achilles Tendons. *Physiotherapy*, *89*(6), 375-383. doi:10.1016/S0031-9406(05)60030-2
- Black, J., Cook, J., Kiss, Z. S., & Smith, M. (2004). Intertester reliability of sonography in patellar tendinopathy. *J Ultrasound Med*, *23*(5), 671-675.
- Bojsen-Møller, F., Trantum-Jensen, J., & Simonsen, E. B. (2014). *Bevægeapparatets anatomi* (Vol. 13. udgave). København: Munksgaard.
- Bolvig, L., Fredberg, U., & Schifter Rasmussen, O. (2011). *Textbook on musculoskeletal ultrasound*. København: Munksgaard.
- Brukner, P., & Khan, K. (2012). *Brukner & Khan's clinical sports medicine* (Vol. 4 th). North Ryde, N.S.W: McGraw-Hill Education.
- Carter, R. E., & Lubinsky, J. (2016). *Rehabilitation research : principles and applications* (5th ed. ed.): Elsevier.

- Cieza, A., & Stucki, G. (2008). The International Classification of Functioning Disability and Health: its development process and content validity. *Eur J Phys Rehabil Med*, 44(3), 303-313.
- Clarsen, B., Bahr, R., Heymans, M. W., Engedahl, M., Midsundstad, G., Rosenlund, L., . . . Myklebust, G. (2015). The prevalence and impact of overuse injuries in five Norwegian sports: Application of a new surveillance method. *Scand J Med Sci Sports*, 25(3), 323-330. doi:10.1111/sms.12223
- Comin, J., Cook, J. L., Malliaras, P., McCormack, M., Calleja, M., Clarke, A., & Connell, D. (2013). The prevalence and clinical significance of sonographic tendon abnormalities in asymptomatic ballet dancers: a 24-month longitudinal study. *British Journal of Sports Medicine*, 47(2), 89-92. doi:10.1136/bjsports-2012-091303
- Cook, J. L., Khan, K. M., Kiss, Z. S., Purdam, C. R., & Griffiths, L. (2000). Prospective imaging study of asymptomatic patellar tendinopathy in elite junior basketball players. *J Ultrasound Med*, 19(7), 473-479.
- Cook, J. L., Khan, K. M., Kiss, Z. S., Purdam, C. R., & Griffiths, L. (2001). Reproducibility and clinical utility of tendon palpation to detect patellar tendinopathy in young basketball players. *British Journal of Sports Medicine*, 35(1), 65-69. doi:10.1136/bjms.35.1.65
- Cook, J. L., & Purdam, C. R. (2009). Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. *British Journal of Sports Medicine*, 43(6), 409-416. doi:10.1136/bjms.2008.051193
- Cook, J. L., Rio, E., Purdam, C. R., & Docking, S. I. (2016). Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research? *British Journal of Sports Medicine*. doi:10.1136/bjsports-2015-095422
- de Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, 59(10), 1033-1039. doi:<http://dx.doi.org/10.1016/j.jclinepi.2005.10.015>
- de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press.
- Dean, B. J. F., Gettings, P., Dakin, S. G., & Carr, A. J. (2016). Are inflammatory cells increased in painful human tendinopathy? A systematic review. *British Journal of Sports Medicine*, 50(4), 216-220. doi:10.1136/bjsports-2015-094754

- Docking, S. I., Ooi, C. C., & Connell, D. (2015). Tendinopathy: Is Imaging Telling Us the Entire Story? *J Orthop Sports Phys Ther*, 45(11), 842-852.
doi:10.2519/jospt.2015.5880
- Drew, B. T., Smith, T. O., Littlewood, C., & Sturrock, B. (2014). Do structural changes (eg, collagen/matrix) explain the response to therapeutic exercises in tendinopathy: a systematic review. *British Journal of Sports Medicine*, 48(12), 966-972. doi:10.1136/bjsports-2012-091285
- Fredberg, U., Bolvig, L., & Andersen, N. T. (2008). Prophylactic Training in Asymptomatic Soccer Players with Ultrasonographic Abnormalities in Achilles and Patellar Tendons. *The American Journal of Sports Medicine*, 36(3), 451-460. doi:doi:10.1177/0363546507310073
- Fredberg, U., Bolvig, L., Andersen, N. T., & Stengaard-Pedersen, K. (2008). Ultrasonography in Evaluation of Achilles and Patella Tendon Thickness. [Sonografische Bestimmung der Dicke von Achilles- und Patellarsehne]. *Ultraschall in Med*, 29(01), 60-65. doi:10.1055/s-2007-963027
- Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Med*, 30(1), 1-15.
- Kettunen, J. A., Kvist, M., Alanen, E., & Kujala, U. M. (2002). Long-Term Prognosis for Jumper's Knee in Male Athletes. *The American Journal of Sports Medicine*, 30(5), 689-692. doi:doi:10.1177/03635465020300051001
- Khan, K. M., Bonar, F., Desmond, P. M., Cook, J. L., Young, D. A., Visentini, P. J., . . . Wark, J. D. (1996). Patellar tendinosis (jumper's knee): findings at histopathologic examination, US, and MR imaging. Victorian Institute of Sport Tendon Study Group. *Radiology*, 200(3), 821-827.
doi:10.1148/radiology.200.3.8756939
- Kongsgaard, M., Kovanen, V., Aagaard, P., Doessing, S., Hansen, P., Laursen, A. H., . . . Magnusson, S. P. (2009). Corticosteroid injections, eccentric decline squat training and heavy slow resistance training in patellar tendinopathy. *Scand J Med Sci Sports*, 19(6), 790-802. doi:10.1111/j.1600-0838.2009.00949.x
- Kunnskapsdepartementet. (2006). Lov om behandling av etikk og redelighet i forskning [forskningsetikkløven]. Retrieved from <https://lovdata.no/dokument/NL/lov/2006-06-30-56>
- Lian, O., Holen, K. J., Engebretsen, L., & Bahr, R. (1996). Relationship between symptoms of jumper's knee and the ultrasound characteristics of the patellar

- tendon among high level male volleyball players. *Scand J Med Sci Sports*, 6(5), 291-296.
- Lian, Ø. B., Engebretsen, L., & Bahr, R. (2005). Prevalence of Jumper's Knee Among Elite Athletes From Different Sports. *The American Journal of Sports Medicine*, 33(4), 561-567. doi:doi:10.1177/0363546504270454
- Lindahl, M., & Juhl, C. (2010). *Den sundhedsvidenskabelige opgave - vejledning og værktøjskasse* (Vol. 2. udgave). Kbh: Munksgård Danmark.
- Lund, H., & Røgind, H. (2016). *Statistik i ord: en introduktion til statistiske metoder* (Vol. 2. udgave). Kbh: Munksgaard.
- Malliaras, P., Cook, J., Purdam, C., & Rio, E. (2015). Patellar Tendinopathy: Clinical Diagnosis, Load Management, and Advice for Challenging Case Presentations. *J Orthop Sports Phys Ther*, 45(11), 887-898. doi:10.2519/jospt.2015.5987
- Mc Auliffe, S., Mc Creesh, K., Purtill, H., & O'Sullivan, K. (2016). A systematic review of the reliability of diagnostic ultrasound imaging in measuring tendon size: Is the error clinically acceptable? *Physical Therapy in Sport*. doi:<https://doi.org/10.1016/j.ptsp.2016.12.002>
- McAuliffe, S., McCreesh, K., Culloty, F., Purtill, H., & O'Sullivan, K. (2016). Can ultrasound imaging predict the development of Achilles and patellar tendinopathy? A systematic review and meta-analysis. *British Journal of Sports Medicine*. doi:10.1136/bjsports-2016-096288
- McCreesh, K., & Lewis, J. (2013). Continuum model of tendon pathology - where are we now? *Int J Exp Pathol*, 94(4), 242-247. doi:10.1111/iep.12029
- McGraw, K. O., Wong, S. P., & Appelbaum, M. I. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*, 1(1), 30-46. doi:10.1037/1082-989X.1.1.30
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. W. (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. doi: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. W. (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. doi:<http://dx.doi.org/10.1016/j.jclinepi.2010.02.006>
- Norman, G., & Streiner, D. L. (2014). *Biostatistics : Bare Essentials (4th Edition)*. Shelton: Shelton, CT, USA: People's Medical Publishing House.
- Ohrndorf, S., Naumann, L., Grundey, J., Scheel, T., Scheel, A. K., Werner, C., & Backhaus, M. (2010). Is Musculoskeletal Ultrasonography an Operator-Dependent Method or a Fast and Reliably Teachable Diagnostic Tool? Interreader Agreements of Three Ultrasonographers with Different Training Levels. *International Journal of Rheumatology*, 2010. doi:10.1155/2010/164518
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*, 86(2), 420-428.
- Skou, S. T., & Aalkjaer, J. M. (2013). Ultrasonographic measurement of patellar tendon thickness--a study of intra- and interobserver reliability. *Clin Imaging*, 37(5), 934-937. doi:10.1016/j.clinimag.2013.01.007
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales : a practical guide to their development and use* (5th ed. ed.). Oxford: Oxford University Press.
- Sunding, K., Fahlström, M., Werner, S., Forssblad, M., & Willberg, L. (2016). Evaluation of Achilles and patellar tendinopathy with greyscale ultrasound and colour Doppler: using a four-grade scale. *Knee Surgery, Sports Traumatology, Arthroscopy*, 24(6), 1988-1996. doi:10.1007/s00167-014-3270-4
- Thornquist, E. (2003). *Vitenskapsfilosofi og vitenskapsteori : for helsefag*. Bergen: Fagbokforl.
- van der Worp, H., van Ark, M., Roerink, S., Pepping, G.-J., van den Akker-Scheek, I., & Zwerver, J. (2011). Risk factors for patellar tendinopathy: a systematic review of the literature. *British Journal of Sports Medicine*, 45(5), 446-452. doi:10.1136/bjism.2011.084079
- Van Genderen, F. R., De Bie, R. A., Helders, P. J. M., & Van Meeteren, N. L. U. (2003). Reliability Research: Towards a More Clinically Relevant Approach. *Physical Therapy Reviews*, 8(4), 169-176. doi:10.1179/108331903225003172
- van Middelkoop, M., van Linschoten, R., Berger, M. Y., Koes, B. W., & Bierma-Zeinstra, S. M. (2008). Knee complaints seen in general practice: active sport participants versus non-sport participants. *BMC Musculoskeletal Disorders*, 9(1), 36. doi:10.1186/1471-2474-9-36

- Visnes, H., Aandahl, H. A., & Bahr, R. (2013). Jumper's knee paradox--jumping ability is a risk factor for developing jumper's knee: a 5-year prospective study. *Br J Sports Med*, 47(8), 503-507. doi:10.1136/bjsports-2012-091385
- Visnes, H., & Bahr, R. (2013). Training volume and body composition as risk factors for developing jumper's knee among young elite volleyball players. *Scand J Med Sci Sports*, 23(5), 607-613. doi:10.1111/j.1600-0838.2011.01430.x
- Visnes, H., Tegnander, A., & Bahr, R. (2015). Ultrasound characteristics of the patellar and quadriceps tendons among young elite athletes. *Scand J Med Sci Sports*, 25(2), 205-215. doi:10.1111/sms.12191
- Warden, S. J., Kiss, Z. S., Malara, F. A., Ooi, A. B. T., Cook, J. L., & Crossley, K. M. (2007). Comparative Accuracy of Magnetic Resonance Imaging and Ultrasonography in Confirming Clinically Diagnosed Patellar Tendinopathy. *The American Journal of Sports Medicine*, 35(3), 427-436. doi:doi:10.1177/0363546506294858
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, 19(1), 231-240. doi:10.1519/15184.1
- Zwerver, J., Bredeweg, S. W., & Akker-Scheek, I. v. d. (2011). Prevalence of Jumper's Knee Among Nonelite Athletes From Different Sports. *The American Journal of Sports Medicine*, 39(9), 1984-1988. doi:doi:10.1177/0363546511413370

Bilag 1: REK godkendelse.



Region: REK nord	Saksbehandler:	Telefon:	Vår dato: 02.05.2016	Vår referanse: 2016/703/REK nord
			Deres dato: 30.03.2016	Deres referanse:

Vår referanse må oppgis ved alle henvendelser

Jon Joensen
IGS

2016/703 Reliabilitet undersøkelse av ultralyd diagnostikk på patellarsene tendinopati

Vi viser til søknad om forhåndsgodkjenning av ovennevnte forskningsprosjekt. Søknaden ble behandlet av Regional komité for medisinsk og helsefaglig forskningsetikk (REK nord) i møtet 21.04.2016. Vurderingen er gjort med hjemmel i helseforskningsloven § 10, jf. forskningsetikkloven § 4.

Forskningsansvarlig: UiB
Prosjektleder: Jon Joensen

Prosjektleders prosjekttomtale

Formålet med dette metodestudiet (reliabilitetsstudiet) er å dokumentere reliabiliteten ved måling av patellarsenens tykkelse med ultralyd. I nyere tid er ultralydskanning blitt vanligere til diagnostisering av Jumpers knee. Tilstanden er en av de mest vanlig senelidelser i kneet og kjennetegnet ved aktivitetsrelaterte smerter på fremsiden av kneet. På ultralyd kan en fortykket patellarsene medregnes som et tegn på Jumpers knee. Det er derfor viktig å vite om målinger av patellarsenens tykkelse er pålitelige når to eller flere måler samme ultralydbilde. Denne studien ønsker å analysere hvorvidt to fysioterapeuter oppnår høy pålitelighet ved måling av patellarsenens tykkelse med ultralyd. Samt hvorvidt en fysioterapeut oppnår høy pålitelighet ved flere målinger av patellarsenens tykkelse.

Vurdering

Framleggingsplikt

De prosjektene som skal framlegges for REK er prosjekt som dreier seg om "medisinsk og helsefaglig forskning på mennesker, humant biologisk materiale eller helseopplysninger", jf. helseforskningsloven (h) § 2. "Medisinsk og helsefaglig forskning" er i h § 4 a) definert som "virksomhet som utføres med vitenskapelig metodikk for å skaffe til veie ny kunnskap om helse og sykdom". Det er altså formålet med studien som avgjør om et prosjekt skal anses som framleggelsespliktig for REK eller ikke.

I dette prosjektet er formålet beskrevet som at «Denne studien ønsker å analysere hvorvidt to fysioterapeuter oppnår samme resultat ved to eller flere målinger av patellarsenens tykkelse med ultralyd. I tillegg hvorvidt en fysioterapeut oppnår samme resultat ved flere målinger av patellarsenens tykkelse.»

I protokollen under avsnittet «hensikt» har prosjektleder beskrevet at «Det er derfor af stor betydning at kende reliabiliteten for ultralydsmålinger af patellarsenens tykkelse. Hensigten er den, at bestemme reliabiliteten ved måling af patellarsenens tykkelse med UL hos individer med klinisk diagnosticeret PT for to fysioterapeuter med basal opplæring i UL.»

I søknaden under punkt 4.1 står det at «Forsøgspersonerne vil ikke have direkte fordel af studiet, da der ikke er snak om et effektstudie. Imidlertid vil studiet være indirekte gunstig, grundet anvendelse af UL som

Besøksadresse:
MH-bygget UIT Norges arktiske
universitet 9037 Tromsø

Telefon: 77546140
E-post: rek-nord@asp.uit.no
Web: <http://helseforskning.etikkom.no/>

All post og e-post som inngår i
saksbehandlingen, bes adressert til REK
nord og ikke til enkelte personer

Kindly address all mail and e-mails to
the Regional Ethics Committee, REK
nord, not to individual staff

supplement i diagnosticeringen og behandling men foreløbig med begrenset dokumentation for reliabiliteten.»

Slik studien fremstår synes formålet å være å avdekke hvor treffsikker som metode ultralydsmålinger er.

Selv om dette er en helsefaglig studie og funnene i studien indirekte vil kunne gi en helsemessig gevinst faller ikke prosjektet inn under definisjonen av de prosjekt som skal vurderes etter helseforskningsloven.

Godkjenning fra andre instanser

Det påhviler prosjektleder å undersøke hvilke eventuelle godkjenninger som er nødvendige fra eksempelvis personvernombudet ved den aktuelle institusjon eller Norsk samfunnsvitenskapelig datatjeneste (NSD).

Vedtak

Etter søknaden fremstår prosjektet ikke som et medisinsk og helsefaglig forskningsprosjekt som faller innenfor helseforskningsloven. Prosjektet er ikke fremleggingspliktig, jf. hfl § 2.

Klageadgang

Du kan klage på komiteens vedtak, jf. forvaltningsloven § 28 flg. Klagen sendes til REK nord. Klagefristen er tre uker fra du mottar dette brevet. Dersom vedtaket opprettholdes av REK nord, sendes klagen videre til Den nasjonale forskningsetiske komité for medisin og helsefag for endelig vurdering.

Med vennlig hilsen

May Britt Rossvoll
sekretariatsleder

Kopi til: post@uib.no

Bilag 2: Forespørsel/samtykke erklæring for deltagelse i projektet

Forespørsel om deltagelse i forskningsprosjektet
Intra –og intertester reliabilitet for ultralydmåling av
patellarsenens tykkelse
En Tværsnitsstudie

Dette er et spørsmål til deg om å delta i en forskningsstudie.

Tilstand Jumpers knee (*patellar tendinopati*) er lokalisert til patellarsenen på fremsiden av kneet. Patellarsenen er senen til knestrekksmuskelen/fremre lårmuskel (*musculus quadriceps femoris*). Jumpers knee er en av de mest vanlig senelidelser i kneet og kjennetegnet ved aktivitetsrelaterte smerter på fremsiden av kneet. I nyere tid er ultralydskanning blitt vanligere til diagnostisering av Jumpers knee. Ultralydskanning kan inngå som et godt supplerende diagnostisk verktøy til den kliniske undersøkelse for Jumpers knee. Flere studier har vist at diagnostisk ultralyd er like presis som MR-skanning ved verifisering av klinisk diagnostiserte senelidelser. På ultralyd kan en patellarsene med hevelse (*fortykket sene*) være indikasjon på Jumpers knee. Det er derfor viktig å vite om ultralydmålinger av patellarsenens tykkelse er pålitelige, altså i hvilket omfang det er mulig å få samme resultat ved to eller flere målinger.

Denne studien ønsker å analysere hvorvidt tre fysioterapeuter oppnår samme resultat ved to eller flere målinger av patellarsenens tykkelse med ultralyd. I tillegg hvorvidt en fysioterapeut oppnår samme resultat ved flere målinger av patellarsenens tykkelse.

Du er blitt spurt om å delta i denne studien fordi du har kontaktet fastlegen/fysioterapeut grunnet smerter i kneet eller du mottaker behandling for Jumpers knee av fysioterapeut, fastlege eller annet helsepersonell. Du driver med idrett og er plaget med knesmerter eller har Jumpers knee, eller så er du frisk og uten Jumpers knee.

Hva innebærer prosjektet

Det er ønskelig å inkludere 2 grupper i studien for sammenligning. En gruppe med Jumpers knee og en gruppe friske personer.

Prosjektet innebærer informasjon om alder, kjønn, sykehistorie og utførelse av en klinisk undersøkelse av kneet ditt. Den kliniske undersøkelse medfører registrering av smerteterskel, som uttrykkes med et algometer (*i kilogram*) som trykkes mot senen til en kjenner ubehag. Ytterlig vil det bli foretatt en måling av isometrisk styrke for fremre lårmuskel. Videre utføres bildediagnostikk med ultralyd av kneet. Med ultralydskanneren vil det bli tatt bilde av patellarsenen på begge knærne for senere måling av tykkelsen. Har du jumpers Knee vil du bli bedt om å svare på et kort spørreskjema.

Hele undersøkelsen vil ha en varighet på cirka en time. Telefonnummer registreres i forkant av undersøkelsen for avtale av time og slettes ved oppmøte. Du vil bli spurt om å ta av på benene, undertøy kan beholdes på. Deltagelse i dette prosjekt vil ikke kreve at du avviker/opphører fra/med et eventuelt igangsatt behandlingsforløp.

Når undersøkelsen av alle deltagerne er gjennomført, vil forskerne måle tykkelsen av patellarsenen ved bruk av ultralydbildene. Deretter vil forskerne analysere hvor pålitelige disse målingene er, altså hvor forskjellig forskerne måler.

Mulige fordele og ulemper

Ultralydskanningen er uten kjente bivirkninger og uten helserisiko. Ultralyd fremstår som en smerte -og strålefri målemetode. Den kliniske undersøkelse er likeledes uten kjente bivirkninger og risiko. Registrering av smerteterskelen utføres ved et trykk på patellarsenen, trykket stoppes ved ubehag. Testen er risikofri og du vil ikke oppleve ubehag/smerte etter testen.

Du vil ikke direkte oppleve en bedring av din tilstand ved deltagelse i denne studien, da studien ikke inneholder analyse av en behandlingseffekt. Du vil få utført en klinisk undersøkelse av kneet samt ultralydskanning, som du kan dra fordel av ved eventuell videre behandlingstiltak.

Frivillig deltagelse og mulighet for å trekke sitt samtykke

Det er frivillig å delta i prosjektet. Dersom du ønsker å delta, undertegner du samtykkeerklæringen på siste side. Du kan når som helst og uten å oppgi noen grunn trekke ditt samtykke. Dette vil ikke få konsekvenser for din videre behandling. Dersom

du trekker deg fra prosjektet, kan du kreve å få slettet innsamlede prøver og opplysninger, med mindre opplysningene allerede er kodet/anonymisert, inngått i analyser eller brukt i vitenskapelige publikasjoner. Dersom du senere ønsker å trekke deg eller har spørsmål til prosjektet, kan du kontakte Niels Bjerre Thygesen, telefon 45 11 74 41, email Niels.Thygesen@student.uib.no .

Hva skjer med informasjonen om deg?

Informasjonen som registreres om deg skal kun brukes slik som beskrevet under hensikten med studien. Du har rett til innsyn i hvilke opplysninger som er registrert om deg og rett til å få korrigert eventuelle feil i de opplysningene som er registrert.

Alle opplysningene vil bli behandlet uten navn og fødselsnummer eller andre direkte eller indirekte personidentifiserbare opplysninger. Dine data vil bli kodet og det vil ikke være mulig å identifisere deg av dataene eller resultatene. Telefonnummer for avtale av time slettes ved oppmøte og kan derfor ikke kobles til dataene.

Prosjektleder har ansvar for den daglige driften av forskningsprosjektet og at opplysninger om deg blir behandlet på en sikker måte. Det vil ikke være mulig å identifisere deg av resultatene når disse publiseres.

Din deltagelse i dette prosjekt gir deg rett til informasjon om resultatet av studien. For informasjon om studieutfallet ta kontakt til Niels B. Thygesen, email Niels.Thygesen@student.uib.no

Forsikring

Ultralydskanningen er en smertefri metode uten kjente bivirkninger og en velkjent og anvendt metode. Dette samme gjør seg gjeldende for den kliniske undersøkelse. Derfor vurderes det unødvendig med en egen forsikring i studien.

Økonomi

Ved deltagelse i prosjektet vil eventuelt tapt arbeidsfortjeneste eller omkostninger i forbindelse med transport ikke bli dekket.

Godkjenning

Prosjektet er godkjent av Regional komite for medisinsk og helsefaglig forskningsetikk.
Reference nummer til REK: 2016/703/REK nord. Prosjektet er vurdert ikke
meldepliktig av Norsk Samfunnsvitenskapelig datatjeneste AS.

Samtykke til deltagelse i prosjektet

Jeg er villig til å delta i prosjektet

Sted og dato

Deltakers signatur

Deltakers navn med trykte bokstaver

Stedfortredende samtykke når berettiget, enten i tillegg eller istedenfor

Som nærmeste pårørende til _____ (Fullt navn)
samtykker jeg til at hun/han kan delta i prosjektet.

Sted og dato

Pårørendes signatur

Pårørendes navn med trykte bokstaver

Jeg bekrefter å ha gitt informasjon om studien.

Sted og dato

Signatur

Rolle i prosjektet

Bilag 3: Dataindsamlingskema klinisk undersøgelse

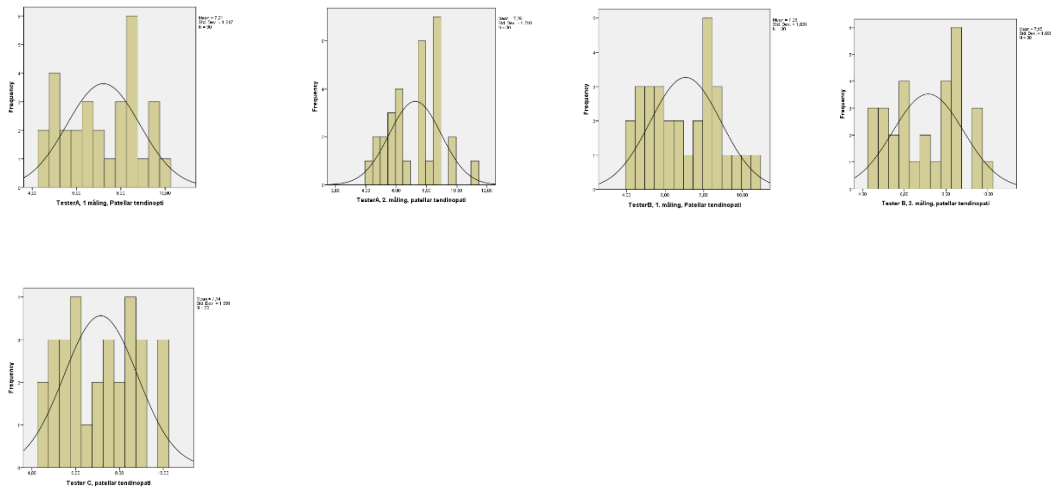
ID	Alder	Kjøn M: mann K: Kvinne	Historie med patellarsene smerter under og etter aktivitet (Ve/Hø)	Smerte/ømhed ved palpation av patellarsenen (Ve/Hø)	idrettsutøver (Ja/Nei)	Varighet av symptomer >1 måned >3 måneder >6 måneder >1 år	Sekundær diagnose	Operasjon i kneet (år siden) (Ve/Hø)	I behandling (hvilken/når?)	Injeksjoner/NSAID (Når) (Ve/Hø)
Niels 01									Nei	Nei
Niels 02									Nei	Nei
Niels 03									Nei	Nei
Niels 04									Nei	Nei

Bilag 4: Dataskema til registrering af målinger

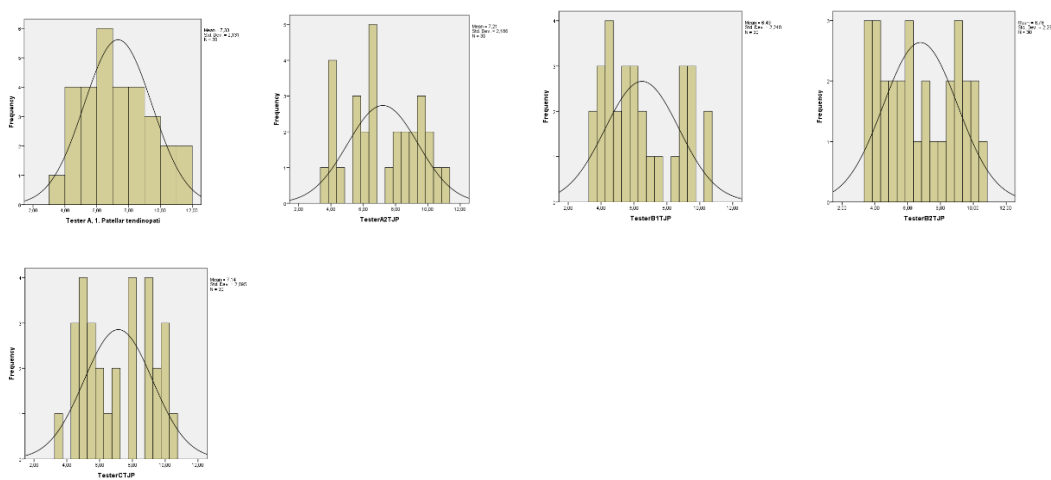
ID	Longitudinale målinger			
	Venstre, bilde nr ..	Data i cm.	Kommentar	Høyre, bilde nr..
Niels 01	NR 1			NR 5
Niels 02	NR 1			NR 5
Niels 03	NR 1			NR 5
Niels 04	NR 1			NR 5
Niels 05	NR 1			NR 5
Niels 06	NR 1			NR 5
Niels 07	NR 1			NR 5
Niels 09	NR 1			NR 5
Niels 10	NR 1			NR 4

Bilag 5: Histogrammer til illustration af fordelingen af data.

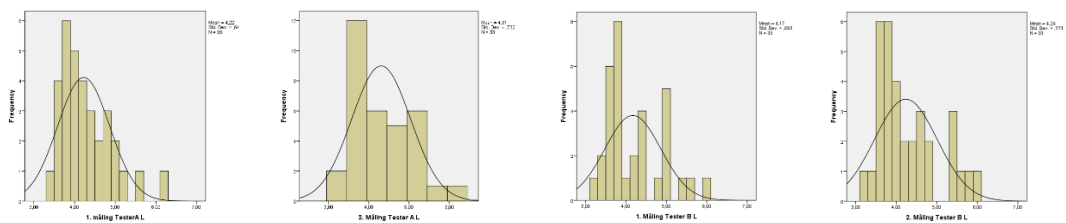
Longitudinal, symptomatisk: fra venstre tester A måling 1. og 2., Tester B måling 1. og 2. og tester C

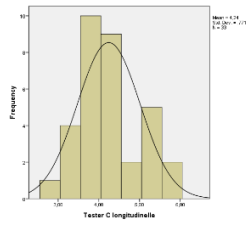


Transversal, symptomatisk: fra venstre tester A måling 1. og 2., Tester B måling 1. og 2. og tester C

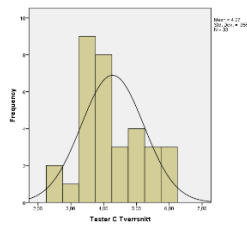
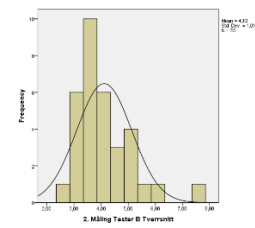
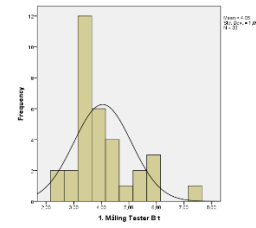
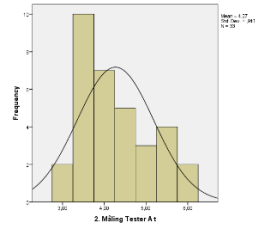
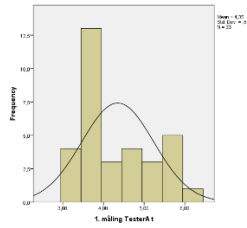


Longitudinal, asymptomatic: fra venstre tester A måling 1. og 2., Tester B måling 1. og 2. og tester C



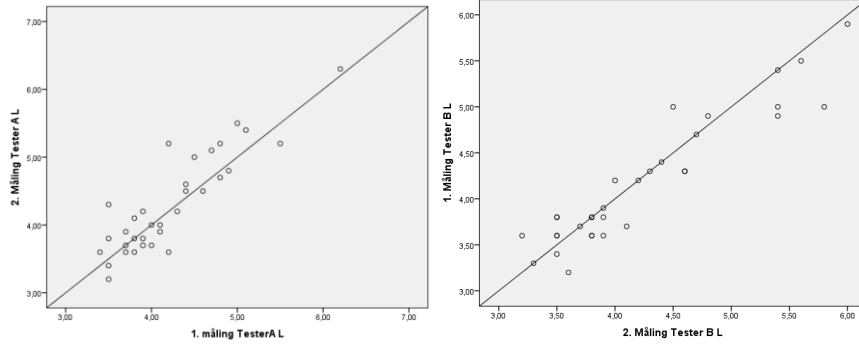


Transversal, asymptotisk: fra venstre tester A måling 1. og 2., Tester B måling 1. og 2. og tester C

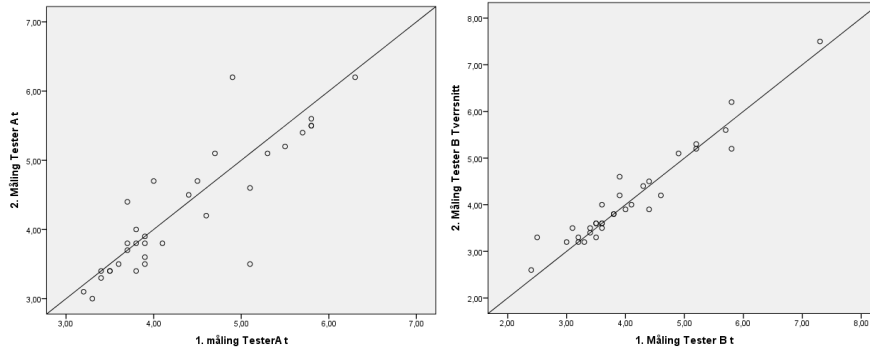


Bilag 6: Scatterplots for illustration af forholdet mellem målingerne.

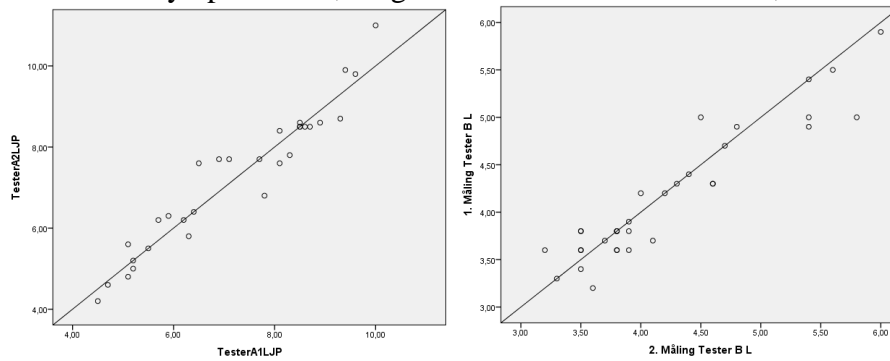
Intratester asymptotisk, longitudinal: Fra venstre tester A, tester B.



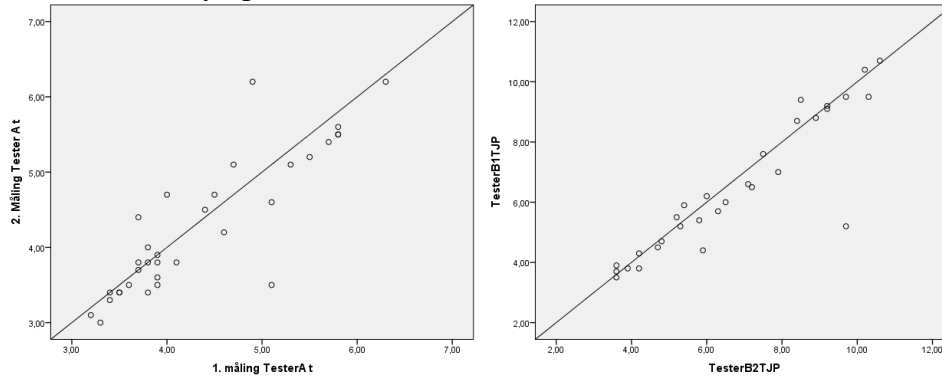
Intratester asymptotisk, transversal: Fra venstre tester A, tester B.



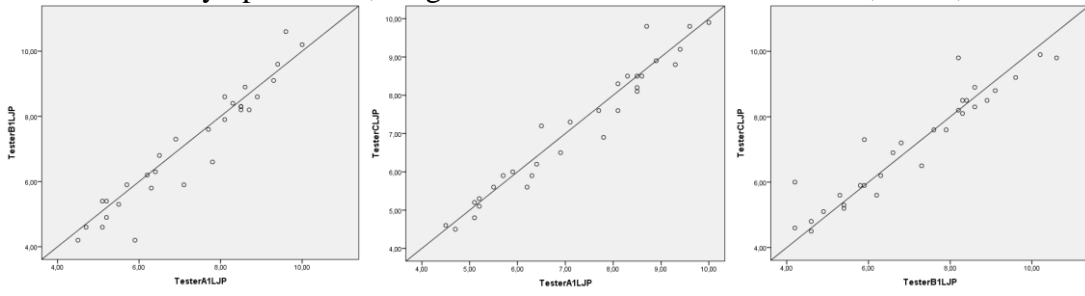
Intratester symptomatisk, longitudinal: Fra venstre tester A, tester B



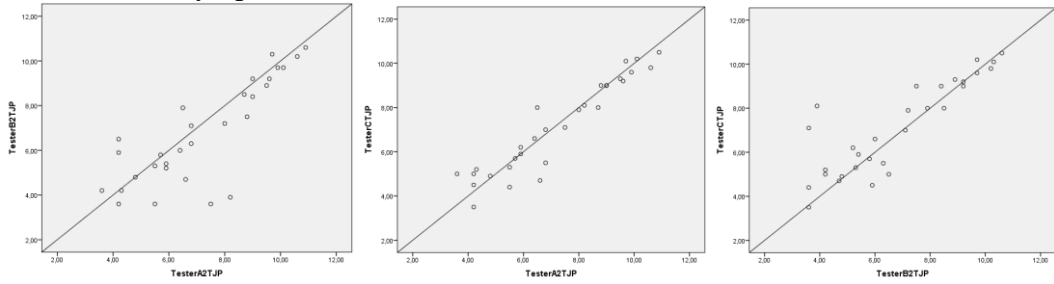
Intratester symptomatisk, transversal: Fra venstre tester A, tester B



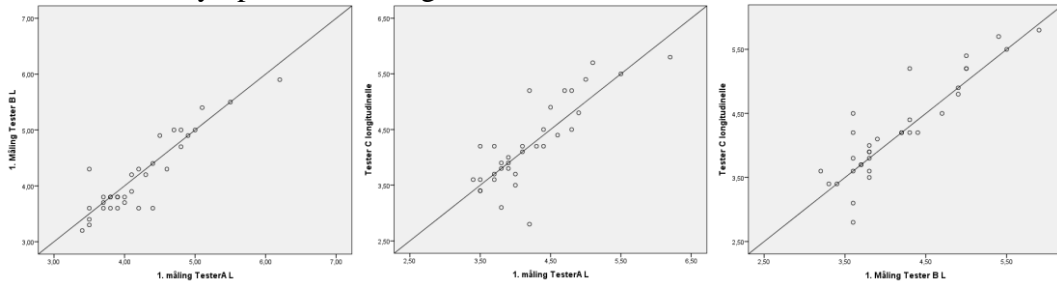
Intertester symptomatisk, longitudinal: Fra venstre tester A1/B1, A1/C, B1/C



Intertester symptomatisk, transversal: Fra venstre tester A2/B2, A2/C, B2/C



Intertester asymptomatisk, longitudinal: Fra venstre tester A1/B1, A1/C, B1/C



Intertester asymptomatisk, transversal: Fra venstre tester A1/B1, A1/C, B1/C

