

Knut Magnus Bauge Helle

Uke-til-uke reliabilitet for målinger av eksentrisk hamstringsstyrke og isometrisk adduksjons-, abduksjons-, innadrotasjons-, utadrotasjons- og fleksjonsstyrke i hoften blant kvinnelige Toppseriespillere i fotball

-En metodestudie

Masteroppgave i idrettsvitenskap
Institutt for idrettsmedisinske fag
Norges idrettshøgskole, 2021

Forord

Denne masteroppgaven markerer slutten på et femårig studieløp på Norges idrettshøyskole (NIH), og derfor også slutten på studentlivet for denne gang. Jeg har økt min kunnskap og er mange erfaringer, opplevelser og venner rikere. Tiden på NIH kommer jeg aldri til å glemme!

Arbeidet med oppgaven har vært utfordrende, spennende og ikke minst lærerikt. Spesielt har det vært gøy å komme tett på Toppseriespillerne, å få testet de både før og etter fotballsesongen 2020. Det mest utfordrende med oppgaven var å arbeide med et stort datamateriale og gjennomføre de statistiske analysene. Dette arbeidet tok mye lengre tid enn jeg hadde sett for meg, og fikk meg til å innse hvor mye arbeid som ligger bak tall som blir framstilt i tabeller og figurer i forskning.

Jeg ønsker å rette en stor takk til mine to veiledere Roald Bahr og Roar Amundsen. Dere har guidet meg gjennom denne oppgaven, tross utfordringer med planlegging og gjennomføring av studien grunnet korona-restriksjoner. Takk til Roald fordi du deler av din kunnskap, framstiller deg på en forståelig måte, og gir raske tilbakemeldinger på arbeidet mitt. En ekstra takk til deg Roar for alle timene vi tilbrakte sammen med innsamling av datamaterialet og for tett oppfølging i ukene inn mot innlevering.

Takk, Sanna og Sindre for korrekturlesing.

Og til slutt, takk til alle spillerne i Toppserien som tok del i denne reliabilitetsstudien.

Knut Magnus Bauge Helle

Oslo, juni 2021

Sammendrag

Bakgrunn: Styrketester blir brukt for å undersøke risikofaktorer for skader, for å måle framgang etter en treningsperiode og for å vurdere retur til idrett for skadde spillere. En pågående kohortstudie ved Norges idrettshøgskole kartlegger skader og sykdommer for norske kvinnelige elitefotballspillere, og undersøker risikofaktorer for hamstring-, hoft- og lyskeskader. I denne studien gjennomfører spillerne i Toppserien en standardisert helseundersøkelse før og etter hver sesong, som inneholder blant annet testing av styrke i hamstrings, lyske og hoftefleksorer. Reliabiliteten til disse testene er ikke undersøkt blant kvinnelige fotballspillere. Formålet med denne masteroppgaven var derfor å undersøke uke-til-uke reliabilitet til seks standardiserte styrketester, som alle inngår i den standardiserte helseundersøkelsen.

Metode: Tjueåtte skadefrie kvinnelige fotballspillere i Toppserien (21 ± 3 år) ble inkludert i denne test-retest reliabilitetsstudien. Vi undersøkte eksentrisk hamstringsstyrke testet med NordBord, isometrisk adduksjons-, abduksjons-, utadrotasjon-, og innadrotasjonsstyrke i hoften testet med ForceFrame og isometrisk hoftefleksjonsstyrke testet med hånd-holdt dynamometer. Spillerne gjennomførte tre maksimale repetisjoner på alle testene, og ble testet av samme testleder, ved to anledninger, med én ukes mellomrom. Test-retest reliabiliteten ble beregnet med LOA, ICC, SEM/SEM% og MDC/MDC%. Resultatene fra testene er oppgitt i maksimalkraft og som gjennomsnittlig maksimalkraft målt i Newton. Høyre og venstre bein er analysert hver for seg i alle testene.

Resultater: For beregning av maksimalkraft viste testene en $ICC_{3,1}$ på moderat til høy reliabilitet, og varierte fra 0,75 til 0,95 (95% KI: 0,51-0,98). SEM% viste en variasjon fra 3,2% til 8,4%. Samlet for alle styrketestene varierte den gjennomsnittlige differansen (\bar{d}) mellom testdagene fra -7,7N til 9,2N, og MDC95% varierte fra 9,0% til 23,2%. Test-retest reliabiliteten ble noe bedre ved å beregne gjennomsnittsverdier fra hver testdag.

Konklusjon: Målefeilene for samtlige tester skyldes hovedsakelig tilfeldig feil. Testene viste generelt moderat til høy relativ reliabilitet (ICC) og samtlige tester viste akseptable SEM%-verdier $\leq 10\%$. MDC95% for de fleste testene viste akseptable verdier, hvor en tenkt intervensjon kan tenke seg å gi en større framgang på individnivå enn MDC95%. Det må allikevel påpekes at MDC95% for innadrotasjons- og hoftefleksjonstesten var over 20% for maksimalkraft, noe som indikerer at det trengs store endringer på individnivå for å kunne fastslå en endring av klinisk betydning. De sistnevnte testene er derfor best egnet til å vurdere styrke på gruppenivå, eller brukes i forbindelse med skader og retur til idrett for kvinnelige fotballspillere i Toppserien. På grunn av metodiske svakheter med studien bør resultatene bli tolket med forsiktighet og kun anvendes for utvalg med tilsvarende variabilitet som i denne studien.

Forkortelser

ABD	Abduksjon av hoftelddet
ADD	Adduksjon av hoftelddet
CV	Variasjonskoeffisient
HHD	Hånd-heldt dynamometer
ICC	Intraklasse korrelasjonskoeffisient
IKD	Isokinetisk dynamometer
IR	Innadrotasjon av hoftelddet
KI	Konfidensintervall
IHS	Idrettens Helsecenter
LOA	Limits of agreement
MDC	Minste reelle endring
ME	Metodefeil
NH	Nordic hamstring
SD	Standardavvik
SEM	Standardfeil til målingene
UR	Utadrotasjon av hoftelddet

Innholdsfortegnelse

1. Introduksjon	8
1.1 Formål.....	10
1.2 Problemstilling	10
2. Teori.....	11
2.1 Idrettsskader	11
2.1.1 Hamstringsskader.....	11
2.1.2 Hofte- og lyskeskader	12
2.1.3 Oppsummering hamstringsskader- og hofte/lyskeskader.....	14
2.2 Risikofaktorer og forebygging	14
2.2.1 Risikofaktorer for hamstringsskader.....	16
2.2.2 Forebygging av hamstringsskader	17
2.2.3 Risikofaktorer for hofte- og lyskeskader	18
2.2.4 Forebygging av hofte- og lyskeskader.....	18
2.2.5 Praktiske implikasjoner hofte- og lyske.....	19
2.3 Verdien til en standardisert helseundersøkelse.....	19
2.4 Målemetodens egenskaper	21
2.4.1 Reliabilitet.....	21
2.4.2 Kilder til målefeil.....	22
2.4.3 Kvantifisering av reliabilitet	24
2.4.4 Testens evne til å oppdage endring	28
2.4.5 Validitet	29
2.5 Målemetoder for muskelstyrke.....	29
2.5.1 Isometrisk testing.....	29
2.5.2 Isotonisk testing	30
2.5.3 Isokinetisk testing	30
2.5.4 Valg av tester	30
2.5.5 Reliabilitet for de valgte målemetodene	31
3. Metode	38
3.1 Rekruttering	38
3.2 Testere.....	39
3.3 Datainnsamling.....	40
3.4 Måleinstrumenter.....	41
3.5 Testprotokoll	42
3.6 Databehandling	46
3.7 Statistiske analyser.....	47
3.8 Etikk.....	49

4.	Resultater	50
4.1	Utvalg	50
4.2	Forskjell mellom utvalget og øvrige spillere	50
4.3	Test-retest reliabilitet for maksimalkraft (N)	51
4.4	Test-retest reliabilitet for gjennomsnittlig maksimalkraft (N)	53
4.5	Bland-Altman plott	55
4.5.1	Maksimalkraft (N)	56
4.5.2	Gjennomsnittlig maksimalkraft (N)	58
5.	Diskusjon.....	60
5.1	Hovedfunn resultater	60
5.2	Reliabilitet for styrkemålingene.....	61
5.2.1	Hvorfor det er utfordrende å sammenligne med tidligere studier	61
5.2.2	Eksentrisk hamstringsstyrke	62
5.2.3	Isometrisk adduksjons-, abduksjons-, utadrotasjons og innadrotasjonsstyrke i hoften	64
5.2.4	Isometrisk hoftefleksjonsstyrke	66
5.2.5	Oppsummering.....	68
6.	Metodediskusjon.....	69
6.1	Testere.....	69
6.2	Utvalg	69
6.3	Datainnsamling.....	71
6.4	Målemetoder.....	74
6.5	Statistiske analyser.....	77
6.5.1	Relativ reliabilitet	78
6.5.2	Absolutt reliabilitet	79
6.6	Oppsummering metodediskusjon.....	80
6.7	Praktiske implikasjoner	80
7.	Konklusjon	82
	Referanseliste	83
	Tabelloversikt.....	98
	Figuroversikt	99
	Vedlegg.....	100

1. Introduksjon

I 2018 var det 144 059 aktive kvinnelige fotballspillere i Norge (NFF, 2018). Kvinnefotballen er i stadig utvikling, og antall spillere og ressurser som brukes på utvikling av kvinnefotball har økt de siste årene (UEFA, 2015). Prestasjonsnivået har også forbedret seg og elitespillere har større treningsvolum og konkurransekrav enn tidligere (Datson et al., 2014). En høyere totalbelastning kan ha konsekvenser for skaderisikoen spillerne blir utsatt for, og har blitt assosiert med en høyere forekomst av skader og sykdom (Drew & Finch, 2016). Selv om populariteten til kvinnefotball øker, ligger skadestudier et stykke bak herrefotballen.

Fotball er en idrett som er karakterisert med maksimale spurter, spark, sidebevegelser og hurtige akselerasjoner (Arnason et al., 2008; Tyler et al., 2010). Som et resultat av dette er hamstrings-, hofte- og lyskeskader blant de vanligste skadene i kvinnefotball (Hägglund et al., 2009; Larruskain et al., 2018; Crossley et al., 2020; Waldén et al., 2015). Hamstringsskader står for rundt 12% av skadene for mannlige fotballspillere (Ekstrand et al., 2011). Blant studier som har undersøkt skadefrekvensen i kvinnefotball, stod hofte- og lyskeskader for 2-16% av skadeforekomsten (Larruskain et al., 2018; Nilstad et al., 2014; Hägglund et al., 2009; Faude et al., 2005; Tegnander et al., 2008). En svakhet med tidligere studier på kvinnefotball er at de fleste har undersøkt fraværsskader (Nilstad et al., 2014; Hägglund et al., 2009; Jacobson & Tegner, 2007; Tegnander et al., 2008; Faude et al., 2005; Gaulrapp et al., 2010). Mye tyder på at insidensen av hamstringsskader og hofte- og lyskeskader er langt større for kvinnelige fotballspillere når man også tar ikke-fraværsskader i betraktning (Harøy et al., 2017a; Langhout et al., 2019). Det er derfor mulig at tidligere studier kun har avdekket toppen av isfjellet når det gjelder helseproblem blant kvinnelige fotballspillere. Ikke-fraværsskader kan redusere prestasjon og funksjonsnivå til spillere, uten å føre til fravær fra fotball. Det er derfor viktig å få en oversikt over disse skadene for å danne et totalbilde av skadeproblematikken i kvinnefotball (Bahr et al., 2020; Harøy et al., 2017a).

Målet med skadeforebyggende arbeid består av å tidlig oppdage risikofaktorer, for deretter å minimere disse før en skade oppstår (Meeuwisse, 1994). Grunnleggende forståelse for både risikofaktorene og mekanismene bak skader er derfor viktig for å kunne identifisere spillere med høyere skaderisiko, og for å kunne utvikle effektive skadeforebyggende programmer (Bahr & Krosshaug, 2005). Muskelstyrke har blitt identifisert som en modifiserbar risikofaktor for skade i hamstrings og lysken for fotballspillere (Orchard et al., 1997; Timmins

et al., 2016; Light & Thorborg, 2016; Thorborg et al., 2011; Engebretsen et al., 2010). For å redusere forekomsten av hamstringsskader og hoft- og lyskeskader i fotball er det utviklet ulike skadeforebyggende program som baserer seg på styrketrening. Regelmessig trening av eksentrisk hamstringsstyrke gjennom øvelsen *Nordic hamstrings* har vist over 50% reduksjon i hamstringsskaderisiko for mannlige fotballspillere (Petersen et al., 2011; Arnason et al., 2008; van Horst et al., 2015). Det er likeledes evidens for at lav eksentrisk hamstringsstyrke er en modifierbar risikofaktor for mannlige fotballspillere (Timmins et al., 2016).

Regelmessig trening av hoftadduktorene gjennom øvelsen *Copenhagen adduction* har vist en reduksjon på 40% av hoft- og lyskeskader i en stor epidemiologisk studie av mannlige fotballspillere i Norge (Harøy et al., 2019). Foreløpig er det mangler i litteraturen av store epidemiologiske studier som undersøker effekten av skadeforebyggende programmer for hamstrings og hoft- lyske i kvinnelig elitefotball.

Som en del av det skadeforebyggende arbeidet har standardiserte helseundersøkelser blitt populært å bruke på idrettsutøvere. Slike helseundersøkelser har som mål å kartlegge risikofaktorer for skader og sykdom, identifisere spillere med smerter og skaffe baselinemålinger for styrkeverdier i blant annet hamstrings og lysken for fotballspillere (IOC, 2009). I en pågående kohortstudie ved Norges idrettshøgskole kartlegges skader og sykdommer for norske kvinnelige Toppseriespillere. Studien har også et mål om å undersøke risikofaktorer for hamstringsskader og hoft- og lyskeskader. I forbindelse med studien gjennomfører spillerne en standardisert helseundersøkelse før og etter hver sesong. Testene i den standardiserte helseundersøkelsen er valgt ut på bakgrunn av relevans for prestasjon og for å fungere som målestokk for retur til idrett under rehabilitering ved en eventuell skade. Testene består blant annet av testing av eksentrisk hamstringsstyrke med NordBord, testing av adduksjons-, abduksjons-, utadrotasjons- og innadrotasjonsstyrke i hoften med ForceFrame og testing av hoftefleksjonsstyrke med hånd-heldt dynamometer (HHD).

For å vurdere om disse testene og målemetodene er gyldige er vi avhengig av at de er reliable, valide og at de kan identifisere endring over tid (Carter et al., 2011; Portney & Watkins, 2009; Atkinson & Nevill, 1998; Weir, 2005). Gullstandarden for testing av muskelstyrke i underekstremitetene er isokinetisk dynamometer (IKD) (Timmins et al., 2016; van Dyk et al., 2018). IKD kan vurdere flere variabler samtidig (hastighet, kraft, vinkel for maks kraftutvikling og kraftkurver), og under ulikt muskelarbeid (isometrisk og isokinetisk). Derimot er IKD sin begrensning at apparatet er dyrt, ikke egnet for å fraktes og trenger omfattende opplæring av testere og deltakere ved bruk (Opar et al., 2013). Apparatene som

ble brukt i denne studien (NordBord, ForceFrame, HHD) er alle rimeligere, mindre og bærbarere. Dette gjør disse apparatene mer tilgjengelig enn IKD under klinisk testing. Reliabiliteten til flere av måleapparatene som blir brukt i den pågående kohortstudien er allerede undersøkt (Opar et al., 2013; Wiesinger et al., 2020; Desmyttere et al., 2019; Ryan et al., 2019; Thorborg et al., 2010a; Ishøi et al., 2019). De fleste studiene som undersøker reliabiliteten, har derimot et lite utvalg ($n < 25$) og undersøkte grupper med andre karakteristika enn kvinnelige fotballspillere.

Denne studien er den første som undersøker test-retest reliabilitet for kvinner med bruk av NordBord og ForceFrame, og er den første som vurderer hoftefleksjonsstyrke for HHD med den valgte metoden. Det er sentralt å vurdere reliabiliteten til styrketester og apparater som blir brukt i skadeforebyggende program, kartlegging og retur til idrett også for kvinnelige fotballspillere.

1.1 Formål

Formålet med denne masteroppgaven var å undersøke uke-til-uke reliabiliteten for seks standardiserte styrketester gjennomført i forbindelse med den standardiserte helseundersøkelsen for kvinnelige fotballspillere i Toppserien. Testene som ble vurdert var eksentrisk hamstringsstyrke målt med NordBord (Vald Performance, Albion, Australia), isometrisk hofteabduksjon-, hofteadduksjon-, innadrotasjons- og utadrotasjonsstyrke målt med ForceFrame (Vald Performance, Albion, Australia), og isometrisk hoftefleksjonsstyrke målt med HHD (Hoggan Health Industries Inc. West Draper, UT, USA).

1.2 Problemstilling

Hva er uke-til-uke reliabiliteten for målinger av eksentrisk hamstringsstyrke og isometrisk adduksjons-, abduksjons-, innadrotasjons-, utadrotasjons- og fleksjonsstyrke i hoften hos skadefrie kvinnelige fotballspillere i Toppserien?

2. Teori

2.1 Idrettsskader

Idrettsskader blir definert som vevsskader eller andre forstyrrelser av normal fysisk funksjon som følge av rask eller repetert overføring av kinetisk energi under deltakelse i idrett (Bahr et al., 2020). Muskel- og skjelettproblemer eller skader er vanlige i mange idretter, og er den hyppigste årsaken til fravær for fotballspillere (Ekstrand et al., 2011). Skader kan ha kort- eller langvarig negativ virkning for fotballspillere og kan redusere lagets prestasjoner (Hägglund et al., 2013). Det er da viktig for alle involverte parter (trener, spiller, klubbeier) å begrense skadeomfanget til et minimum (Ekstrand, 2013).

Skader blir i hovedsak kategorisert på to måter: (1) Akutte skader forekommer plutselig, hvor årsaken til skaden er kjent, eller (2) belastningsskader med en gradvis utvikling og uten kjente traumer (Bahr et al., 2020). I noen tilfeller er etiologien blandet. Tidligere studier på kvinnefotball viser at 80-90% av skadene er akutte og 10-20% er belastnings-relaterte (Tegnander et al., 2008; Jacobson & Tegner., 2007; Faude et al., 2005; Nilstad et al., 2014; Gaulrapp et al., 2010). På kvinnelig elitenivå varierer skadeforekomsten fra 12,6 til 30,6 skader per 1000 time med kampeksponering, og fra 1,4 til 8,4 skader per 1000 time med trening (Nilstad et al., 2014; Hägglund, et al., 2009; Jacobson & Tegner, 2007; Tegnander et al., 2008; Faude et al., 2005; Gaulrapp et al., 2010; Larruskain et al., 2018).

I kvinnefotball er skader i underekstremitetene den vanligste årsaken til fravær, hvor en kohortstudie utført på norske fotballspillere rapporterte at disse sto for 60%-85% av alle fraværsskadene (Nilstad et al., 2014). Gjennom kartlegging av kvinne- og herrefotball viser det seg at gruppene har en ulik skadeprofil (Crossley et al., 2020). Risikoen for fremre korsbåndsskade er over dobbelt så stor for kvinner (Montalvo et al., 2017). Kvinner har i tillegg en økt risiko for hjernerystelse, ankelskader og kneskader, sammenlignet med menn (Dick., 2009; Boling et al., 2010; Murphy et al., 2003). På den andre siden kan det se ut til at menn har større risiko for hamstrings- og lyskeskader enn kvinner (Cross et al., 2013; Waldén et al., 2015). Hamstrings- og lyskeskader er likevel blant de vanligste skadene i kvinnefotball, så disse bør undersøkes videre (Crossley et al., 2020; Waldén et al., 2015).

2.1.1 Hamstringsskader

Hamstrings beskriver muskulaturen som er lokalisert på baksiden av låret. Muskelgruppen består av m. biceps femoris, m. semimembranosus og m. semitendinosus og har som oppgave

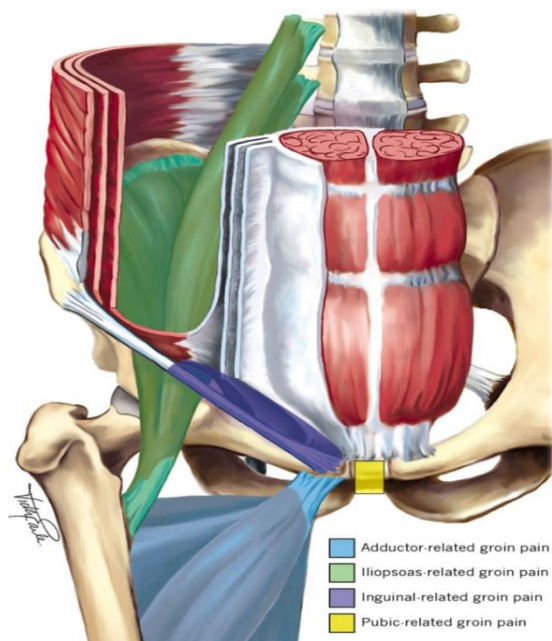
å strekke i hoftelrådet og bøye i kneleddet (Dahl & Rinvik, 2010). Majoriteten av hamstringstrekksskadene i fotball for menn er lokalisert i det lange hodet til m. biceps femoris (Ekstrand et al., 2011).

Skader i hamstring er vanlig i idretter som er karakterisert med maksimale spurter, spark og hurtige akselerasjoner (Arnason et al., 2008). De fleste hamstringsskader for menn er ikke-kontakt skader, forekommer ofte akutt og står for omtrent 12% av alle fotballrelaterte skader (Ekstrand et al., 2011). Det er også en tendens til at hamstringsskader oppstår oftere i kamp enn under trening, og er vanligere i månedene før sesong enn under og etter (Feeley et al., 2008; Ekstrand et al., 2011). Insidensen av akutte hamstringsskader er anslått til å være 60% lavere blant kvinnelige enn mannlige fotballspillere (Crossley et al., 2020; Cross et al., 2013; Opar et al., 2014). En nylig publisert oversiktsartikkel som undersøkte skader i kvinnefotball anslo at insidensen av hamstringsskader er 0,22 (0,15-0,32) per 1000 time med trening/kamp (Crossley et al., 2020). I samme studie ble det også funnet en tendens til at hamstringsskader hadde en høyere insidens i elitefotball for kvinner kontra fotball på lavere nivå (Crossley et al., 2020).

Selv om risikoen for hamstringsskader virker til å være lavere for kvinner enn menn, er skaden fremdeles et stort problem i kvinnefotball. Hägglund og medarbeidere (2009) undersøkte fotballskader i Sverige gjennom 2005-sesongen, og konkluderte med at skader i låret, særlig relatert til hamstrings, var den vanligste skaden hos begge kjønn. Forfatterne spekulerte i at forekomsten av hamstringsskader i kvinnefotball øker som en konsekvens av større treningsvolum, samt generelt høyere tempo under eksponering for fotball (Hägglund et al., 2009). Det er imidlertid gjennomført lite forskning på kartlegging og forebygging av hamstringsskader i kvinnefotball.

2.1.2 Hofte- og lyskeskader

Lyskeområdet kan beskrives som området fra de nedre mage-musklene til den antero-mediale delen av låret (Weir et al., 2015). Det har tidligere blitt brukt ulike begreper for å beskrive de samme symptomene på lyskerelaterte-skader, derfor ble det i 2014 gjennomført et møte i Doha for å bli enige om en standard terminologi (Weir et al., 2015). For å diagnostisere lyskeskader anbefalte klinikerne å dele opp skaden i seks grupper: (1) adduktor-relaterte, (2) iliopsoas-relaterte, (3) inguinal-relaterte, (4) pubis-relaterte, (5) hofte-relaterte, og (6) andre årsaker til lyskesmerter (Weir et al., 2015).



Figur 1: Beskriver lokasjon på de definerte områdene for lyskesmerter av Weir og medarbeidere (2015 s. 771). (Gjengitt med tillatelse)

Få studier har undersøkt kliniske diagnoser for lyskeskader. Werner og medarbeidere (2019) registrerte skader for europeiske mannlige toppklubber over en 15års-periode. Resultatene viste at adduktor-relaterte lyskeskader stod for 2/3 av skadene, etterfulgt av iliopsoas-relaterte og pubis-relaterte skader. Lignende resultater ble observert i en toårig kohort-studie av mannlige fotballspillere i Qatar (Mosler et al., 2018).

Hofte- og lyskeskader er blant de vanligste skadene i seniorfotball, og står for 12,8% av alle skader for menn og 6,9% av alle skader for kvinner (Waldén et al., 2015). Blant studier som har undersøkt skadefrekvensen i kvinnefotball stod hofte- og lyskeskader for 2-16% av det totale antall skader (Larruskain et al., 2018; Nilstad et al., 2014; Hägglund et al., 2009; Faude et al., 2005; Tegnander et al., 2008). Insidensen av hofte- og lyskeskader blant kvinnelige fotballspillere er 0,15 (0,10 – 0,21) skader per 1000 time med trening/kamp (Crossley et al., 2020). De akutte hofte- og lyskeskadene oppstår ofte under raske akselerasjoner, raske vendinger, retningsforandringer og når spillere sparker ballen (Tyler et al., 2010). Skaden kan blant annet påvirke ferdigheter og resultere i smerter ved deltakelse i fotball (Mosler et al., 2018). For kvinner er risikoen for å pådra seg lyskerelaterte-skader noe større under kamp enn under trening, mens herrespillere har en over dobbelt så stor risiko under kamp (Tyler et al., 2010). Sammenlignet med menn virker det som om kvinnelige fotballspillere har en lavere insidens av lyskeskader, men årsaken til dette er fremdeles uklar. Anatomiske forskjeller i

lyske og bekkene har dog blitt foreslått som mulige årsaker til lavere skadeforekomst (Schache et al., 2017).

De fleste studier på kvinnefotball har studert skader som setter spillere ut av trening/konkurransen, omtalt som fraværsskader (Nilstad et al., 2014; Hägglund et al., 2009; Jacobson & Tegner., 2007; Tegnander et al., 2008; Faude et al., 2005; Gaulrapp et al., 2010). Hofte- og lyskeskader faller derimot ofte utenfor definisjonen fraværsskader. En studie som undersøkte både mannlige og kvinnelige elitespillere i Norge over en seks ukers periode med skaderegistrering, fant en fire ganger høyere insidens av hofte- og lyskeskader av ikke-fraværsskader (36%), enn for fraværsskader (9%), blant de kvinnelige fotballspillerne i studien (Harøy et al., 2017a). Hofte- og lyskeskader har også blitt rapportert som den vanligste ikke-fraværsskaden blant kvinnelige fotballspillere (Langhout et al., 2019). Svakheten med å kun rapportere fraværsskader er at metoden kun fanger opp skader som fører til fravær fra trening og kamp. Metoden tar ikke hensyn til om en fotballspiller spiller med småskader, som kan resultere i redusert prestasjon (Bahr et al., 2020). Mange hofte- og lyskeproblemer er belastningsrelaterte og gir gradvis større smerte og/eller redusert prestasjon uten å nødvendigvis føre til fravær fra idretten. Omfanget av hofte- og lyskeproblemer i kvinnefotball kan derfor ha blitt underrapportert i flere studier, som bruker tradisjonelle metoder og kun undersøker fraværsskader for å vurdere skadeomfanget (Harøy et al., 2017a; Langhout et al., 2019).

2.1.3 Oppsummering hamstrings- og hofte/lyskeskader

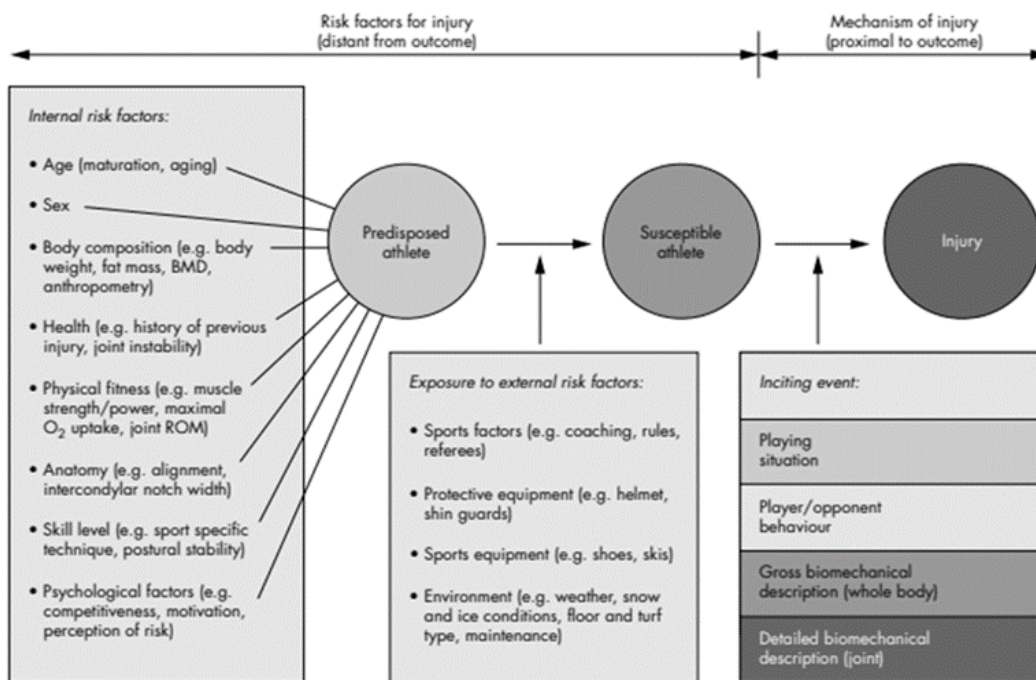
Både hamstrings- og lyskeskader er et stort problem i kvinnefotball. Informasjonen vi har om skadene i dag, viser en lavere skadeinsidens i kvinnefotball enn i herrefotball (Cross et al., 2013; Waldén, et al., 2015; Crossley et al., 2020). Høyere tempo og høyere totalbelastning har blitt sett i sammenheng med økt skaderisiko i kvinnefotball (Drew & Finch, 2016; Hägglund et al., 2009). Anvendelse av bedre verktøy for å vurdere samtlige skader, fremfor å studere fravær fra trening og/eller konkurranse, kan gi et bedre bilde over skadeproblematikken i kvinnefotball (Harøy et al., 2017a; Bahr, 2009).

2.2 Risikofaktorer og forebygging

Den klassiske tilnærmingen til forskning på skadeforebyggende arbeid er beskrevet som en fire-stegs modell (Mechelen et al., 1992). Det første steget i modellen går ut på å bestemme omfanget av problemet, inkludert alvorlighetsgraden og insidensen av skaden. Steg nummer to omhandler identifisering av risikofaktorene og skademekanismene som er forbundet med

skaden, herunder årsaks-sammenhengen til hvordan og hvorfor skaden oppstår (Figur 2). Det tredje steget består av å utvikle og implementere skadeforebyggende tiltak basert på risikofaktorene som ble identifisert i steg to. Det siste steget består av å evaluere effekten skadeforebyggende programmer har på reduksjon av skader (Mechelen et al., 1992). Denne modellen viser tydelig viktigheten av å få kartlagt risikofaktorer for å kunne utvikle effektive skadeforebyggende programmer.

Risikofaktorer for skader blir ofte delt inn i ytre og indre faktorer. De ytre risikofaktorene er knyttet til omgivelsene, mens de indre er knyttet til spilleren selv (Bahr & Krosshaug, 2005). Risikofaktorer kan også bli kategorisert som modifiserbare og ikke-modifiserbare. De modifiserbare faktorene, som for eksempel muskelstyrke, er viktig å finne for å utvikle gode skadeforebyggende programmer. De ikke-modifiserbare faktorene kan fortelle oss hvem som har økt risiko for skade, basert på faktorer en spiller ikke kan påvirke, som for eksempel alder og kjønn (Bahr, 2016). Årsaken til idrettsskader er ofte kompleks og sammensatt av flere faktorer (figur 2).



Figur 2: Modell som beskriver årsakssammenhenger for skader av Meeuwisse et al. (1994) modifisert av Bahr & Krosshaug (2005). (Gjengitt med tillatelse)

En vanlig måte å undersøke risikofaktorer på er ved bruk av standardiserte helseundersøkelser (IOC, 2009). Standardiserte helseundersøkelser kan variere fra korte helsevurderinger til mer omfattende vurderinger, som kan inkludere test av respirasjon, kardiovaskulær og nevrologisk funksjon, i tillegg til undersøkelser av generell medisinsk og muskel- og skjelett status (IOC,

2009; Dvorak et al., 2009; Fuller et al., 2007). I 2009 utga «The International Olympic Committee» (IOC) og «the Fédération Internationale de Football Association» (FIFA) retningslinjer for standardisert helseundersøkelse, med den hensikt å sette en standard for effektiv testing, samt bistå i å oppdage mulige helse- og skaderisikoer tidlig (IOC, 2009; Dvorak et al., 2009). Alle idretter har et skademønster og dette mønsteret må vurderes nøye når skadeforebyggende tiltak blir iverksatt. For å sikre at samtlige undersøkelser dekker de viktige aspektene ved idrettens egenart bør derfor standardiserte helseundersøkelser tilpasses de spesifikke idrettene (IOC, 2009). FIFA anbefaler undersøkelse av muskel- og skjelett i underekstremitetene, spesielt rettet mot hamstrings-, lyske-, kne- og ankelskader (Dvorak et al., 2009).

2.2.1 Risikofaktorer for hamstringsskader

Til tross for en stor vitenskapelig innsats gjennom flere år er ikke forekomsten av hamstringsskader redusert i fotball, men ser heller ut til å kunne øke for mannlige fotballspillere (Ekstrand et al., 2013). Forskning har lokalisert en rekke ikke-modifiserbare risikofaktorer, inkludert økt alder, spillerposisjon og tidligere skader, som øker risikoen for fremtidige hamstringsskader (Maniar et al., 2016; van Dyk et al., 2017; Engebretsen et al., 2010). Nylig har større oppmerksomhet blitt rettet mot modifiserbare faktorer som kan påvirkes for å redusere risikoen for hamstringsskader. Kroppsvekt, muskelstyrke, utilstrekkelig oppvarming, muskeltretthet og dårlig muskel fleksibilitet har blitt foreslått som modifiserbare riskfaktorer for hamstringsskader for mannlige fotballspillere (van Dyk et al., 2017; Timmins et al., 2016).

Risikofaktorene knyttet til styrke inkluderer forskjell i hamstringsstyrke mellom høyre og venstre bein, muskulær ubalanse mellom hamstrings og quadriceps, og lav eksentrisk hamstringsstyrke (Croisier et al., 2008; Opar et al., 2015; Timmins et al., 2016; Orchard et al., 1997; Fousekis et al., 2011). Timmins og medarbeidere (2016) undersøkte mannlige fotballspillere (n= 152) og deres risiko for hamstringsskade. Spillere som presterte under 337N i en test av eksentrisk hamstringsstyrke, hadde over fire ganger høyere relativ risiko for hamstringsskade gjennom en sesong, sammenlignet med de som skåret bedre enn 337N. Høy eksentrisk hamstringsstyrke ble også forbundet med lavere risiko for hamstringsskade for eldre fotballspillere i studien (Timmins et al., 2016). Studien kan gi en indikasjon på hvilke spillere som har en økt risiko for skade. For at slike studier skal ha en nytteverdi for kvinnelige fotballspillere, er man imidlertid avhengig av å undersøke verdier for kvinnelige kohorter i større epidemiologisk studier.

Det er også andre store epidemiologiske studier som har undersøkt styrkes betydning for hamstringsskader, som motstrider funnene til Timmins og medarbeidere (2016). van Dyk og medarbeidere (2016) undersøkte 614 profesjonelle mannlige fotballspillere over en fire-års periode. Hamstrings: quadricepsratio ble i denne studien vurdert til å være en svak risikofaktor for skade, men ikke tilstrekkelig til å kunne vurdere ratioen som en individuell risikofaktor for hamstringsskade. En annen studie av van Dyk og medarbeidere (2017) undersøkte hamstringsstyrkes betydning for skaderisiko i en prospektiv kohortstudie med 413 mannlige fotballspillere. De rapporterte at verken isokinetisk eller isotonisk hamstringsstyrke var forbundet med risiko for hamstringsskade. Heller ikke hamstrings: quadricepsratioen hadde betydning for skaderisiko. Forskerne konkluderer med at den kliniske betydningen av hamstringsstyrke som risikofaktor for hamstringsskader er ubegrunnet.

2.2.2 Forebygging av hamstringsskader

Det finnes en rekke studier som rapporterer at eksentrisk styrketrening av hamstrings reduserer risikoen for hamstringsskader. En øvelse som er mye brukt, er *Nordic hamstrings*. Øvelsen baserer seg på at en spiller starter knestående med en knevinkel på 90° og med strake hofter, mens en annen spiller holder beina igjen. Spilleren lener seg deretter sakte framover, mens hun motstår tyngdekraften med hamstringsmuskulaturen til hun treffer bakken. Øvelsen er nærmere forklart og illustrert av Mjøltnes og medarbeidere (2004). Studier som har undersøkt effekten av *Nordic hamstring* i et skadeforebyggende program, viser at insidensen av nye hamstringsskader kan halveres (Petersen et al., 2011; Arnason et al., 2008; Horst et al., 2015). Dette gjelder både for spillere med tidligere skader i hamstrings, spillere uten en kjent patologi og spillere i forskjellige aldersgrupper (van Dyk et al., 2019; Attar et al., 2017). Paradoksalt nok har ikke effekten av *Nordic hamstrings* vist signifikant reduksjon i hamstringsskader for to kvinnelige kohorter (Espinosa et al., 2015; Soligard et al., 2008). Espinosa og medarbeidere (2015) undersøkte *Nordic hamstrings* på én kohort av kvinnelige fotballspillere i første og andre divisjon i Spania. Resultatet ble imidlertid påvirket av et lite utvalg i kontrollgruppen (n=21) og i intervensjonsgruppen (n=22). Soligard og medarbeidere (2008) undersøkte unge norske kvinnelige fotballspillere (n=1892) i alderen 13-17 år over 8 måneder. FIFA 11+, som inkluderer trening av *Nordic hamstrings* ble gjennomført av 1055 spillere i intervensjonsgruppen, mens 837 spiller var i kontrollgruppen. Forskerne observerte bare åtte hamstringsrelaterte skader i intervensjonsgruppen og fem skader i kontrollgruppen. Lav skadeinsidens i studien kan dermed være en medvirkende årsak til at forskerne ikke fant noe signifikant reduksjon i hamstringsskader for yngre kvinnelige fotballspillere. På bakgrunn

av metodiske svakheter med tidligere studier anbefales også *Nordic hamstring*s for kvinnelige fotballspillere (van Dyk et al., 2019).

Litteraturen er enig i at *Nordic hamstring*s reduserer risikoen for hamstringsskader, men mekanismene bak den drastiske reduksjonen i hamstringsskader er fremdeles uklare (van Dyk et al., 2019). En mulig forklaring er at regelmessig trening av *Nordic hamstring*s gir økt fasikkel-lengde i biceps femoris, som videre kan beskytte hamstringsmuskulatur mot framtidig skade (Timmins et al., 2016). Økt eksentrisk hamstringstyrke er en annen mulig forklaring (Harøy et al., 2017b; van Dyk et al., 2019; Timmins et al., 2016). Mekanismene bak den skadeforebyggende effekten av øvelsen bør dog undersøkes videre.

2.2.3 Risikofaktorer for hofte- og lyskeskader

For hofte- og lyskeskader er tidligere skader i lysken den største risikofaktoren for skade (Arnason et al., 2004). Økt alder, spillerposisjon og lav styrke er andre faktorer som har blitt assosiert med hofte- og lyskeskader for fotballspillere (Weir et al., 2015). Engebretsen og medarbeidere (2010) rapporterte at mannlige fotballspillere med svak hofteadduktorstyrke hadde fire ganger høyere risiko for å pådra seg lyskeskader enn spillere med normal hofteadduktorstyrke. Mosler og medarbeidere (2018) rapporterte økt risiko for mannlige fotballspillere når de studerte eksentrisk hofteadduktorstyrke med HHD. Derimot fremhevet Mosler og medarbeidere (2018) et viktig problem under måling av muskelstyrke med HHD. De konkluderte med at assosiasjonen ikke er sterk nok til å skille mellom individuelle spillere i risiko for skade, på bakgrunn av at forskjellen potensielt er mindre enn målefeilen til HHD. Det er også evidens for at en sideforskjell i styrke mellom høyre og venstre bein, og en stor forskjell i adduksjons- og abduksjonsstyrke kan øke risikoen for skade i en kohort med mannlige ishockeyspillere (Tyler et al., 2010).

2.2.4 Forebygging av hofte- og lyskeskader

Det er få studier som har undersøkt skadeforebyggende program rettet mot å redusere hofte- og lyskeskader for kvinnelige fotballspillere. Studier som har studert intervensjonsprogrammer viser at eksentrisk styrketrening av hofteabduktorene og adduktorene, og generelle skadeforebyggende programmer kan bistå til å redusere risikoen for hofte- og lyskeskader (Jensen et al., 2014; Soligard et al., 2008). For mannlige fotballspillere har regelmessig trening av hofteabduktorene vist signifikant reduksjon for hofte- og lyskeskader. Harøy og medarbeidere (2019) gjennomførte en stor epidemiologisk studie med over 600 mannlige fotballspillere i Norge. Resultatene viser en reduksjon på over 40% i hofte-

og lyskeskader for spillere som trente en progressiv variant av *Copenhagen adduction exercise* tre ganger i uken før sesong, og én gang i uken gjennom sesongen. Øvelsen blir på bakgrunn av den skadeforebyggende effekten anbefalt til fotballspillere på alle nivå.

2.2.5 Praktiske implikasjoner hofte- og lyske

Testing av hofte- og lyskestyrke er en del av skadeforebyggende arbeid og rehabilitering etter skader (Desmyttere et al., 2019). Testing kan bistå til å identifisere spillere med økt risiko for skade og for å tidlig oppdage spillere som er i ferd med å få en lyskerelatert skade (Light & Thorborg, 2016). Videre foreslo Wollin og medarbeidere (2018) en kombinasjon av regelmessig overvåking av hofte- og lyskestyrke, samt bruk av *hip and groin hagos score* (HAGOS), for å identifisere fotballspillere med lyskesmerter. En slik tilnærming til hofte- og lyskeskader er sentral, da vi vet at mange av skadene er belastingsrelaterte. I behandling av hofte- og lyskeskader har en forskjell på under 10% mellom maksimalkraft i det skadde og det friske beinet blitt foreslått som en klinisk milepæl før en spiller returnerer til idretten etter en skade (Thorborg et al., 2011). Mer spesifikt har det blitt anbefalt en hofte adduksjon/abduksjon ratio på mer enn 90% og en adduksjonsstyrke lik kontralateral side før man returnerer til idretten etter en adduktor-relatert skade (Nicholas & Tyler, 2002). Tolking av utad- og innadrotasjonsstyrke bør undersøkes videre for å vurdere om sideforskjeller eller lav styrke er en risikofaktor for skade i lyskere regionen (Desmyttere., 2019; Thorborg et al., 2010b).

2.3 Verdien til en standardisert helseundersøkelse

Bahr (2016) beskriver at tre steg er viktig for å validere screeningstester som blir brukt i standardiserte helseundersøkelser, knyttet til deres egnethet til å predikere og forhindre skader; (1) Et sterkt forhold mellom screeningstesten og skaderisiko må bli observert i større prospektive studier, (2) testen må valideres i relevante populasjoner ved hjelp av passende statistiske verktøy, og (3) et intervensjonsprogram rettet mot spillere identifisert med høy risiko må være mer effektivt enn intervensjonsprogrammer som blir gitt til alle spillerne. Videre påpeker Bahr (2016) at formålet med screening for skader ofte blir misforstått. Fram til nå er det ingen screeningstester som klarer å predikere hvem som blir skadet med tilstrekkelig nøyaktighet. Dette skyldes i hovedsak screeningstestene sin sensitivitet og spesifisitet (Bahr, 2016).

Sensitivitet og spesifisitet er to sentrale faktorer ved en diagnostisk test. Sensitivitet er sannsynligheten for at en syk person får et positivt svar. Spesifisitet er sannsynligheten for at

en frisk person får negativt svar (Bahr, 2016). En stor forskjell når en screener for sykdom, sammenlignet med å screene for skade, er at utfallsvariabelen er annerledes. Enten har du en sykdom eller så har du ikke en sykdom, alternativene er derfor *ja* eller *nei*. Når man screener for skade, er det mer komplisert. Spørsmålet blir da om testen klarer å predikere hvem som får en skade utfra eksempelvis den totale styrken, forholdet mellom høyre og venstre bein eller forholdet mellom agonister og antagonister (Bahr, 2016). Bahr (2016) illustrerte forholdet mellom sensitivitet og spesifisitet med å vise til en større studie av van Dyk og medarbeidere (2016). Studien gikk over fire år med 614 deltakere, hvor forskerne observerte at hamstringsstyrke på 60°/sek var uavhengig assosiert med skaderisiko i hamstrings (odds ratio 1,37 per 1 NM/kg forskjell). Bahr (2016) viste til at flere av spillerne som ble beregnet til å ha god styrke i studien, ble skadet i hamstrings. På den andre siden var det flere som var beregnet til å ha høyere risiko i form av lavere styrke som ikke ble skadet (Bahr, 2016). En stor utfordring knyttet til å vurdere skaderisiko er at resultatene ofte er kontinuerlige variabler. Selv om resultater i en muskelstyrketest viser en assosiasjon mellom eksempelvis lav eksentrisk hamstringsstyrke og skaderisiko, er det ikke tilstrekkelig å bruke en test til å predikere hvem som får en skade. Årsaken er at det ofte er overlapp mellom individene som blir predikert til å ha en økt skaderisiko og de som ikke er det i hvem som blir skadet.

Som beskrevet tidligere i oppgaven er det mange risikofaktorer for skade i hamstring og hoftelyske for fotballspillere. Bahr (2016) foreslo at vi må vurdere skaderisikoen utfra flere forhold enn kun de modifiserbare risikofaktorene, som styrken oppnådd på en test. Spillere med ikke modifiserbare risikofaktorer som tidligere skader, aldersforskjeller, samt kjønnsforskjeller for ulike skader, må også tas med i vurderingen når man setter inn skadeforebyggende tiltak rettet mot den gruppen som blir beregnet til å ha en økt skaderisiko (Bahr, 2016). Med andre ord kan ikke en muskelstyrketest i seg selv, brukes som eneste virkemiddel til å predikere skade.

Testing av muskelstyrke er per dags dato dårlig egnet til å predikere hvem som blir skadet, så da kan man rette spørsmål om hvorfor klubber skal bruke tid på å kartlegge styrken til spillerne? Testene kan heldigvis ha andre formål. Resultater fra muskelstyrketester kan avdekke prestasjonsframgang til spillere og/eller fungere som en målestokk under rehabilitering ved en skade, for å vurdere når det er hensiktsmessig med retur til idretten (Bakken, 2018). Videre kan testene undersøke styrkeforholdet mellom beina og forholdet mellom agonister og antagonister (Raastad et al., 2010). Testene kan også brukes for å oppdage endring over tid, eksempelvis fra før sesong til etter sesong eller over lengre tid.

Testene kan da vurdere om den skadeforebyggende treningen spillerne har gjennomført har hatt innvirkning på maksimalstyrken i forskjellige muskelgrupper i underekstremitetene eller om styrken endrer seg gjennom sesongen. Testene kan også brukes til å kartlegge smerter, og sørge for at spillere som oppgir smerter får oppfølging av medisinsk personell (Bakken, 2018). Det er derfor viktig å ha med testing av hamstrings og hofte- og lyskestyrke som en del av den standardiserte helseundersøkelsen for kvinnelige fotballspillere, ettersom insidensen av skadene er stor. Det er også sentralt å undersøke reliabiliteten til testene som blir brukt i standardiserte helseundersøkelser. Dette for å kunne si noe om presisjonen til målingene. En lite reliabel test med stor målevariasjon vil ha liten betydning i klinikken, fordi resultatene som blir oppnådd i en test er knyttet til tilfeldigheter og ikke gjenspeiler den faktiske styrken (Carter et al., 2011).

2.4 Målemetodens egenskaper

En målemetode er en måte å evaluere, forstå og differensiere karakteristika hos personer eller objekter (Portney & Watkins, 2009). De to grunnleggende forutsetningene for målemetoder er validitet og reliabilitet (Carter et al., 2011; Portney & Watkins, 2009; Atkinson & Nevill, 1998). Validitet refererer til om målemetoden måler det den har som hensikt å måle (O'Donoghue, 2012; Atkinson & Nevill, 1998). Reliabilitet refererer til graden en metode eller test er fri fra målefeil (Portney & Watkins, 2009; Carter et al., 2011). Målemetoden må også være responsiv nok til å fange opp små, men klinisk relevante endringer over tid (Carter et al., 2011).

2.4.1 Reliabilitet

Noe av det første som undersøkes før man benytter seg av en ny målemetode er reliabiliteten. Dersom målemetoden ikke gir samsvarende resultat ved repeterte målinger, er den ikke reliabel (Atkinson & Nevill, 1998; O'Donoghue, 2012). Videre kan det også være hensiktsmessig å undersøke reliabiliteten til et måleinstrument på forskjellige grupper i befolkningen (slik som kvinner og menn), for å undersøke for systematiske forskjeller (Carter et al., 2011). Litteraturen nevner ulike termer som brukes i forbindelse med begrepet reliabilitet; repeterbarhet, nøyaktighet, presisjon, reproduserbarhet, stabilitet, test-retest reliabilitet og konsistens er blant begrepene som knyttes til reliabiliteten til en målemetode (Atkinson & Nevill, 1998). I denne oppgaven blir begrepene reproduserbarhet og test-retest reliabilitet anvendt for å beskrive hvor reliabel en målemetode er. Hvilken type reliabilitet som blir undersøkt er knyttet til hensikten og metoden med studien. De to vanligste måtene å

vurdere reliabilitet til en ny målemetode på er å undersøke intra-rater reliabilitet og inter-rater reliabilitet (Carter et al., 2011).

Intra-rater reliabilitet: undersøker graden av samsvar eller korrelasjon av gjentatt testing under samme betingelse. Subjekter, instrument og tester forblir det samme. Det er urimelig å tenke at reproduserbarheten til en test kan bli perfekt. De fleste test-retest studier inneholder kombinasjoner av instrumentfeil, testerfeil og deltakervariasjon (Carter et al., 2011; O'Donoghue, 2012). Test-retest intervall må velges med omhu, for å unngå at muskeltretthet, læringseffekt eller hukommelseffekt skal påvirke resultatet. Samtidig må ikke intervallet bli for langt slik at en reelle endringer har forekommet i målevariabelen (Portney & Watkins, 2009).

Inter-rater reliabilitet: refererer til når de samme subjektene blir uavhengig målt av forskjellige testere (O'Donoghue, 2012). Ved undersøkelse av inter-rater reliabilitet vil en høyere reliabilitet bli oppnådd desto større samsvar det er mellom testerne (Carter et al., 2011).

2.4.2 Kilder til målefeil

Ved test-retest studier er det et mål å redusere målefeilen til et minimum, for å vurdere måle-metodens reliabilitet. Variasjonen til deltakerne (within-subjects) blir beskrevet som det viktigste reliabilitetsmålet da den sier noe om nøyaktigheten til målemetoden (Hopkins, 2000). En standardisert testprotokoll med testere som har fått god opplæring i bruk av måleinstrumenter, og deltakere som har fått en nøye innføring i øvelser som skal testes er sentralt for å minimere risikoen for målefeil (Portney & Watkins, 2009).

Uavhengig av hvilken type reliabilitet som undersøkes, er det to komponenter av varians som assosieres med målefeil. Dette er systematiske og tilfeldige feil. Summen av disse variablene beskriver den totale målefeilen (Atkinson & Nevill, 1998). Klassisk reliabilitetsteori tar utgangspunkt i at alle observerte målinger består av en sann skår og en feilkomponent, hvor den sanne skårens verdi er uavhengig av målingene (Haugen et al., 2018). Differansen mellom den observerte og den sanne skåren blir betraktet som summen av systematisk og tilfeldige målefeil. For å estimere målefeilen til en målemetode kan vi gjennomføre repeterte målinger, hvor variasjonen til en deltakers skår blir vurdert som målefeil (Carter et al., 2011). Gjennom å identifisere graden av målefeil blir forutsetningen bedre for å forutsi variasjonen som oppstår ved repeterte målinger (Haugen et al., 2018).

Systematisk feil

Systematisk feil refererer til en trend hvor målingene går i en bestemt retning fra test til retest. Resultatene kan både være negativt eller positivt rettet (Atkinson & Nevill, 1998).

Systematiske feil kan skyldes feil i måleinstrumenter, målemetoder eller feil hos tester (Atkinson & Nevill, 1998). Feil kan for eksempel oppstå når deltakerne har for kort tid til å hente seg inn i en test-retest protokoll. Re-testen kan da vise dårligere resultater enn første test, som et resultat av muskeltretthet og mangel på restitusjon (Atkinson & Nevill, 1998; O'Donoghue, 2012). Det er heller ikke uvanlig med en trend hvor re-testen viser høyere verdier for deltakerne, som et resultat av læringseffekt. Deltakerne har da blitt kjent med testapparater og protokollen, og kan på den måten prestere bedre under re-testen (Atkinson & Nevill, 1998; O'Donoghue, 2012). Læringseffekten kan reduseres med tilvenning av øvelser/apparater som brukes i studien og med tilstrekkelig tid mellom test-retest (O'Donoghue, 2012). Hovedsakelig er det validiteten som blir truet av systematisk feil, ettersom målefeilen vil være konstant (Atkinson & Nevill, 1998). For å undersøke om et datasett inneholder systematisk feil kan en t-test eller ANOVA bli gjennomført. Slike tester blir brukt for å sammenligne gjennomsnittsverdiene i en test-retest studie (Atkinson & Nevill, 1998). Selv om t-testen kan være nyttig for å undersøke for systematisk feil, bør den ikke anvendes som det eneste statistiske verktøyet for å vurdere reliabilitet. Svakheten med t-testen er at den ikke tar hensyn til store tilfeldige variasjoner hos deltakerne i studien og kan på bakgrunn av dette vise en høy P-verdi (god reliabilitet), selv ved store individuelle forskjeller mellom testene (Atkinson & Nevill, 1998).

Tilfeldig feil

Den andre kilden til variasjon ved repeterte målinger er graden av tilfeldige feil (Atkinson & Nevill, 1998). Tilfeldig feilmåling påvirker resultatet av en test på en uforutsigbar måte (Haugen et al., 2018). Feilen kan oppstå som følge av variasjon knyttet til tester eller deltakere, omtalt som biologisk variasjon. Tilfeldig feil kan også relateres til mekanisk variasjon, som følge av feil med måleinstrumentet eller som en konsekvens av uoverensstemmelser med testprotokollen (Atkinson & Nevill, 1998). Variasjoner fra deltakere eller tester kan for eksempel være knyttet til motivasjon, læring, erfaringer, kompetanse eller helse (Portney & Watkins, 2009). Vanligvis kan undersøker gjøre lite for å redusere graden av tilfeldige feil, særlig om variasjonen er knyttet til måleinstrumentet som ble benyttet. Graden av tilfeldig feil er vanligvis større enn systematisk feil, grunnet at den er vanskeligere å kontrollere for i en test-retest protokoll (Atkinson & Nevill, 1998).

Det er også andre faktorer som kan påvirke reliabiliteten i en test-retest protokoll med én uke mellom testene. Ulik temperatur fra test til retest, varighet på testprotokoll, forskjellige klær og sko for deltakerne, ulik grad av motivasjon og tilbakemeldinger er alle faktorer som har blitt sett i sammenheng med målevariasjon (Currell & Jeukendrup, 2008). For kvinner er det heller ikke usannsynlig at menstruasjonssyklusen kan påvirke reliabiliteten til en test. En nylig publisert metaanalyse av McNulty og medarbeidere (2020) viser at fysisk prestasjon for noen kvinner er redusert i den første delen av menstruasjonssyklusen som et resultat av lavere østrogennivå.

2.4.3 Kvantifisering av reliabilitet

Reliabilitet kan kvantifiseres på to måter: relativ og absolutt reliabilitet (Carter et al., 2011). Relativ reliabilitet undersøker forholdet mellom to eller flere repeterte målinger (between-subjects), mens absolutt reliabilitet undersøker variasjonen i resultater fra måling til måling (within-subjects) (Carter et al., 2011).

Relativ reliabilitet

Relativ reliabilitet er basert på en idé om at deltakerne i en gruppe vil beholde sin posisjon (rangering) ved repeterte målinger om en målemetode er reliabel (Carter et al., 2011). En deltaker vil muligens ikke få den samme skåren ved første og andre test, men det er forventet at den som presterer best under første test også vil prestere best under re-testen, og motsatt (Atkinson & Nevill, 1998; Carter et al., 2011). Dette kan undersøkes med korrelasjonskoeffisienten, da denne indikerer grad av samsvar mellom repeterte målinger for den undersøkte variabelen (Koo & Li, 2016). Koeffisienten er et tall mellom 0 og 1, hvor en verdi nærmere 1 beskriver mindre variasjon og dermed høyere reliabilitet (Carter et al., 2011). Selv om korrelasjonskoeffisienten er et mye brukt verktøy for å beskrive reliabiliteten til en målemetode, inneholder den svakheter dersom den anvendes som det eneste statistiske verktøyet for å beregne reliabilitet. For det første er det mange måter å regne ut korrelasjonskoeffisienten på, noe som gjør at man kan få forskjellig korrelasjon ut fra det samme datasettet ved bruk av ulike formler (Shrout & Fleiss, 1979; Koo & Li, 2016). For det andre er koeffisienten dårlig egnet for et heterogent utvalg, som vil si at det er store forskjeller mellom studiens deltakere for den undersøkte variabelen. En sterk korrelasjon mellom test-retest er dermed å forvente selv om de individuelle variasjonene kan være store mellom testene (Weir, 2005).

Intraklasse korrelasjonskoeffisient (ICC) er den mest brukte metoden for å vurdere relativ reliabilitet i intra-rater og inter-rater studier (Koo & Li, 2016). ICC er en univariat reliabilitetskoeffisient og kan derfor anvendes ved to eller flere målinger av den samme variabelen. Hvilken ICC som blir benyttet er relatert til *modell, definisjon og type* (Koo & Li, 2016). Totalt er det utarbeidet 10 måter å estimere ICC. De første seks metodene ble introdusert av Shrout & Fleiss (1979), og det ble senere utarbeidet fire nye metoder av McGraw & Wong (1996).

Det er tre ICC-modeller som er ledende når man skal vurdere reliabiliteten til en målemetode. Modell 1 er en enveis tilfeldig variasjonsanalyse som benyttes når målinger utføres av et tilfeldig utvalg testere. Denne metoden blir sjeldent anvendt i kliniske reliabilitets-studier, fordi modellen kun tar hensyn til en uavhengig variabel (Koo & Li, 2016). Modell 2 er en toveis tilfeldig variasjonsanalyse, her blir alle deltakerne målt av de samme testerne, utvalgt fra en større gruppe testere. Denne modellen er best når resultatene skal generaliseres (Koo & Li, 2016). Modell 3 er en toveis blandet variasjonsanalyse som brukes når vi er interessert i de utvalgte testerne og resultatene ikke skal generaliseres (Koo & Li, 2016). Shrout & Fleiss (1979) anbefalte å bruke modell 3 i test-retest studier fordi det ikke er rimelig å generalisere én testers resultater til en større populasjon. Modellen anbefales også for test-retest studier med bakgrunn i at utvalgt ikke kan bli beskrevet som randomisert ved repeterte målinger (Portney & Watkins, 2009).

Definisjon omhandler hva beregningsmetoden fra korrelasjonen baserer seg på. Det er i hovedsak to valg, enten enighet eller konsistens. Hovedforskjellen mellom metodene er at enighet tar hensyn til både systematisk og tilfeldig målefeil i sin beregning, mens konsistens kun tar hensyn til tilfeldig feil. Korrelasjonen blir normalt sett høyere (bedre) ved bruk av konsistens enn enighet (Koo & Li, 2016). Enighet som «definisjon» bør alltid velges for test-retest studier fordi målingene er meningsløse dersom det ikke er enighet under repeterte målinger (Koo & Li, 2016). Hvilken *type* ICC som velges avhenger av verdiene som ligger til grunn for estimatet. Dersom det er én måling som skal vurderes, for eksempel maksimalt dreiemoment eller maksimalkraft, anvendes «single measures». Om det er flere målinger å ta hensyn til, eksempel tre målinger fra en styrketest, anvendes «average measures» (Ko & Li, 2016).

ICC oppgis normalt sammen med et 95% konfidensintervall (KI). Ettersom ICC-verdien som blir oppnådd i reliabilitetsstudier er den forventede verdien av den sanne ICC, er det mer

relevant å vurdere ICC basert på 95% KI, og ikke bare den oppgitte ICC verdien (Koo & Li, 2016).

Det finnes ingen vitenskapelig enighet om hvordan korrelasjonskoeffisienten skal tolkes, eller hva som kjennetegner en akseptabel verdi for ICC (Carter et al., 2011; Koo & Ki, 2016).

Reliabilitet kjennetegner hvor god en målemetode er til å reproducere svar, og det er opp til forskeren selv å vurdere hvor god reliabiliteten til den valgte målemetoden bør være for at den skal kunne forsvares (Portney & Watkins, 2009).

Det er blitt satt ulike grenseverdier for hva som tolkes som god relativ reliabilitet. Disse verdiene blir ofte delt inn i lav, moderat, god og høy reliabilitet. Det er foreslått at ICC < 0,5 tilsvarer svak reliabilitet, 0,5 – 0,75 moderat reliabilitet, 0,75 – 0,90 god reliabilitet, og ICC > 0,90 høy reliabilitet (Koo & Li, 2016). Usikkerhet rundt tolkningen av ICC-verdier og testens svakheter ved heterogene utvalg, gjør at ICC aldri bør brukes som det eneste målet på reliabilitet i en test-retest studie (Atkinson & Nevill, 1998). ICC-verdien er i stor grad påvirket av distribueringen av styrkeverdiene til et utvalg. En homogen gruppe med lite variabilitet i testresultatene vil ofte føre til lav ICC ved repeterte målinger og høy reliabilitet for et heterogent utvalg. ICC egner seg derfor best til å illustrere forholdet (korrelasjonen) mellom repeterte målinger, men verdien sier lite om de individuelle variasjonene (Atkinson & Nevill, 1998; Hopkins, 2000). I litteraturen blir det anbefalt å oppgi ICC i reliabilitetsstudier, men kun for å komplementere utregninger av absolutt reliabilitet (Atkinson & Nevill, 1998).

Absolutt reliabilitet

Absolutt reliabilitet undersøker variasjonen av testen fra måling til måling (within subject). Desto mindre variasjonen er, desto høyere er reliabiliteten (Carter et al., 2011). Metoder for å vurdere absolutt reliabilitet er standardfeil til målingen (SEM), variasjonskoeffisient (CV), limits of agreement (LOA) og metodefeil (ME) (Atkinson & Nevill, 1998; Portney & Watkins, 2009).

SEM er en måte å kvantifisere målefeilen ved repeterte målinger. Jo mindre SEM er, desto bedre er reliabiliteten til målingene (Weir, 2005). SEM kan bli beregnet ved å anvende standardavviket (SD) og ICC i formelen $SEM = SD \cdot \sqrt{1-ICC}$ (Atkinson & Nevill, 1998; Weir, 2005), eller uttrykkes i forhold til gjennomsnittsmålingene (\bar{X}) som prosent $SEM\% = \frac{SEM}{\bar{X}} * 100$. Ved å benytte SEM% kan resultater sammenlignes på tvers av studier. Fra ligningen kan man se at verdien av SEM blir påvirket av både ICC og SD. Høy ICC-verdi og et smalt SD vil dermed gi en lavere SEM-verdi, og dermed bedre reliabilitet.

Variasjonskoeffisient (CV) er forholdet mellom standardavviket og gjennomsnittet til målingene presentert i prosent, og kan bli utregnet på følgende måte: $(CV = \frac{SD}{\bar{X}} * 100)$ (Atkinson & Nevill, 1998). CV blir anvendt på data som har flere repeterte målinger. Om man skal bruke SEM eller CV avhenger i stor grad om målingene inneholder heteroskedastiske verdier (Atkinson & Nevill, 1998). CV appellerer til situasjoner hvor grad av enighet mellom tester avhenger av størrelsen på de målte verdiene. Med andre ord antar CV at de største test-retest variasjonene finner sted hos de individene som skårer høyest på testene (Blant & Altman, 1995). Både SEM og CV er måter å kvantifisere målefeilen ved bruk av standardavviket til målingene og inneholder mange likheter (Atkinson & Nevill, 1998; Weir, 2005).

En svakhet med SEM og CV ligger i den statistiske beregningen. Metodene baserer seg kun på 68% sannsynlighet for at den sanne skåren finnes innen ± 1 SEM av den observerte skåren. Tradisjonelt sett ønsker man å operere med 95% sannsynlighet (Atkinson & Nevill, 1998). En annen kritikk mot CV og SEM er at metodene gir en enkelt skår, som skal beskrive en gjennomsnittlig feil. Beregningene skjuler da viktig informasjon om størrelsen og stabiliteten av de individuelle målefeilene (O'Donoghue, 2012).

LOA ble først introdusert av Bland og Altman i 1986 (Bland & Altman, 1986). LOA er en metode som viser et 95% referanseintervall for målefeilen (Atkinson & Nevill, 1998). LOA baserer seg på å finne den gjennomsnittlige differansen (\bar{d}) mellom testrundene, for så å multiplisere standardavviket til differansen med 1,96 ($LOA = \bar{d} (1,96SD_{diff})$). Deretter trekkes denne summen fra \bar{d} for å finne nedre grense og \bar{d} legges til for å finne øvre grense (Bland & Altman, 1986). LOA kan framstilles grafisk i et Bland-Altman plott ved å legge til LOA og \bar{d} og vi finner da deltakernes individuelle forskjeller mellom testrundene (Bland & Altman, 1986). Plottet kan gi en indikasjon på de systematiske målefeilene (\bar{d}) og de tilfeldige målefeilene ved å se på spredningen og retningen rundt null-linjen (Atkinson & Nevill, 1998). Svakheter med metoden er at med et lite utvalg, vil SD til differansen mellom målingene i større grad bli påvirket av ekstremverdier i den ene eller andre retningen. Målefeilen kan da bli overvurdert med et lite utvalg. En utvalgsstørrelse på minimum 40 deltakere blir derfor anbefalt i litteraturen (Hopkins, 2000; Atkinson & Nevill, 1998).

Hva er akseptabel absolutt reliabilitet?

Det er opp til hver enkel forsker å bestemme hva som aksepteres av målefeil for at metoden skal være presis, ofte knyttet til formålet med testingen (Giaverina, 2015). Generelt kan man

tåle en lavere reliabilitet for målinger som anvendes for forskningsmessig bruk på gruppenivå, mens reliabiliteten for målinger som brukes for beslutningstaking eller diagnostisering på individnivå bør være så høy som mulig (Portney & Watkins, 2009). Signifikante eller ikke-signifikante funn må tolkes etter hvilken betydning de har i praksis. Kjennskap til målemetoden er derfor en forutsetning for å kunne vurdere om funnene er av betydning. Om målefeilen anses som reliabel, kommer derfor an på om målefeilen vurderes som akseptabel for den populasjonen som undersøkes under de gitte testforholdene.

2.4.4 Testens evne til å oppdage endring

Responsivitet er et begrep som beskriver i hvilken grad metoden er sensitiv for små endringer. Begrepet er relatert til både reliabiliteten og validitet til en målemetode. Det er viktig at en målemetode har evnen til å fange opp reelle endringer over tid, men også være stabil når endringer ikke har skjedd. Dette er en viktig egenskap når man vil måle effekten av en intervensjon, som for eksempel endring fra en pre- til en posttest (Carter et al., 2011). Det er flere statistiske termer som blir brukt for å beskrive en reell ending: minimal detectable change (MDC), smallest detectable change (SDC), smallest real change (SRC) og smallest real difference (SRD) (de Vet et al., 2001). Alle disse representerer den minste verdien som er over den forventede målefeilen (de Vet et al., 2001). Det blir i litteraturen hevdet at MDC er tett knyttet til LOA og vil tilsvare det øvre og nedre referanseintervallet til LOA (Atkinson & Nevill, 1998).

I beregningen av MDC er SEM den vanligste statistikken å bruke. Dersom en måling er lite reliabel, blir det indikert med en stor SEM-verdi. Da må det store endringer i deltakerens testverdier for at vi skal kunne si at det har forekommet en reell endring, ettersom denne endringen må være større enn målefeilen. På den andre siden vil en måling med en lav SEM kreve en mindre endring i resultatet for å finne en endring av betydning, ettersom målefeilen er mindre (Carter et al, 2011). MDC kan bli utregnet med 90% sannsynlighet med ($Z=1,69$) eller som 95% ($Z=1,96$) (Portney & Watkins, 2009). Vi kan regne ut MDC95 ved å bruke SEM i ligningen $MDC = 1,96 * SEM * \sqrt{2}$ (Weir, 2005), eller uttrykkes i forhold til gjennomsnittsmålingene i prosent ($MDC95\% = \frac{MDC}{\bar{x}} * 100$). MDC95% kan gjøre resultatene mer klinisk relevant, og kan gjøre det enklere å sammenligne resultatene med tidligere studier på området (Mokkink et al., 2010). MDC95% kan også uttrykkes basert på utvalgsstørrelsen i en gruppe $MDC_{gruppe} = \frac{MDC95\%}{\sqrt{n}}$. (de Vet et al., 2001). Denne verdien vil si noe om hvor mye

endring en gruppe bør ha, for at man kan konkludere med en reel endring mellom to tester med en 95% sannsynlighet.

2.4.5 Validitet

Validitet referer til om testen eller måleinstrumentet måler det den har som hensikt å måle (Carter et al., 2011). En forutsetning for en målemetodes validitet er reliabiliteten. En målemetode som ikke er reliabel er heller ikke valid. Validitet kan videre kategoriseres som intern og ekstern. Intern validitet referer til hvorvidt resultatene fra en studie er representative for utvalget og fenomenene som ble undersøkt. Ekstern validitet refererer til om resultatene fra utvalget kan generaliseres til en større populasjon (Carter et al., 2011). Intern og ekstern validitet kan derfor påvirker hverandre direkte. En strengt standardisert studie øker den interne validiteten, fordi man da har kontroll på ulike bias som kan føre til målefeil, eksempelvis treningsmengde, matinntak og søvn mellom to tester i en test-retest studie. På den andre siden reduserer en streng standardisering den eksterne validiteten og etterprøvbareheten av studien. Det er også andre typer av validitet, som logisk-, kriterie-, innholds- og begrepsvaliditet, men dette blir ikke gått nærmere inn på i denne oppgaven. For å lese om disse henvises leseren til (Carter et al., 2011. s. 241-242).

2.5 Målemetoder for muskelstyrke

Muskelstyrke blir av Raastad og medarbeiderne (2010, s. 13) definert som «den maksimale kraften eller dreiemomentet en muskel eller muskelgruppe kan skape ved en spesifikk eller forutbestemt hastighet». En objektiv testmetode for maksimalstyrke bør være enkel å utføre, lett å repetere og resten av kroppen bør være fiksert under testing. Målemetoden bør også teste styrke så spesifikt som mulig, uten en stor påvirkning av tester (Raastad et al., 2010). For å måle muskelstyrke er det i hovedsak tre typer muskellarbeid som kan måles: isometrisk, isotonisk og isokinetisk (Winter, 2007).

2.5.1 Isometrisk testing

Isometrisk testing krever at det skapes en maksimalkraft eller dreiemoment mot en ubevegelig motstand (Winter, 2007). Isometriske tester kan enkelt standardiseres og har normalt sett en høy reproducerbarhet. Isometrisk testing krever liten grad av tilvenning for deltakerne. Metoden kan blant annet brukes til å måle en muskels maksimalkraft og kraftutvikling (Winter, 2007). En fordel med isometrisk testing kontra isotonisk og isokinetisk testing, er at metoden fører til mindre belastning for muskler- og skjelett. Dette kan ha en stor betydning dersom en undersøger spillere med en kjent patologi (Thorborg et al., 2010b). Selv om

isometrisk testing inneholder en rekke fordeler knyttet til reproduserbarhet, er det spørsmål rundt den praktiske nytten i forhold til idrettsprestasjon, da det er vist forskjeller i aktivering av muskler under isometriske tester i forhold til dynamiske bevegelser (Winter, 2007). Det faktum at mange dynamiske bevegelser er utført med en betydelig forlengelse av elastiske elementer i muskel-sene struktur, kan svekke en isometrisk tests overførbarhet til idretten. Testen er dermed begrenset til å måle muskelstyrken med et bestemt sett av betingelser, uten å ta hensyn til effekten av elastiske elementer under en bevegelse (Winter, 2007).

2.5.2 Isotonisk testing

Isotoniske tester består av å bevege en masse med akselerasjon og deselerasjon. Akselerasjonen endres her ved ulike leddvinkler under en bevegelse (Winter, 2007). Typiske isotoniske tester er målinger av eksentrisk og konsentrisk muskelstyrke. Isotoniske tester kan blant annet brukes til å få verdier for maksimalkraft og kraftutvikling gjennom en bevegelse. Overførbarhetsverdien mellom isotoniske tester og idrettsprestasjon er normalt sett større enn under isometriske testing. Siden isotoniske tester kan gi en større mengde informasjon (eksempelvis potensiale for muskelrekruttering, elastisk energilager, maksimal muskelkontraksjonskraft og indikasjoner på utmattelse) og ofte er tettere knyttet til prestasjon i idretten, kan det være nødvendig å bruke slike tester under klinisk testing, selv om målemetoden er mindre reliabel enn isometriske tester (Winter, 2007).

2.5.3 Isokinetisk testing

Målemetoden som omhandler måling av kraft og dreiemoment under en konstant hastighet er isokinetisk testing (Winter, 2007). Typisk blir slike tester gjennomført i spesifikke maskiner som IKD. Under isokinetisk testing kan muskelens kraft, hastighetsutvikling og/eller dreiemoment måles under en forutbestemt hastighet. Testing med IKD holder ofte en høy reproduserbarhet, forutsatt at testere har fått god opplæring og deltakerne har gått igjennom flere tilvenningsøkter til apparatet i forkant av testingen (Winter, 2007).

2.5.4 Valg av tester

Design av det valgte testbatteriet krever god innsikt i idretten som undersøkes og i noen tilfeller kreative design (Winter, 2007). Spesifisitet til idretten som undersøkes er sentralt. Det er for eksempel unødvendig å teste fotballspillers overkroppsstyrke, da det ikke har stor innvirkning på en spillers skaderisiko eller prestasjon. Testene som velges bør gi så mye som

mulig informasjon om inter-individuelle forskjeller, rehabiliteringsfremgang eller måle en deltakers styrker og svakheter (Winter, 2007).

2.5.5 Reliabilitet for de valgte målemetodene

Eksentrisk hamstringsstyrke

Flere måleinstrumenter har blitt benyttet for å vurdere hamstringsstyrke, blant annet IKD og hånd-heldt dynamometer (HHD). Gullstandarden for testing av hamstringsstyrke er IKD (Timmins et al., 2016; van Dyk et al., 2018). Apparatet kan vurdere mange variabler samtidig (hastighet, kraft, vinkel for maks kraftutvikling, kraftkurver), og under ulikt muskelarbeid (isometrisk og isokinetisk). Test-retest reliabiliteten til IKD for måling av hamstringsstyrke er blitt vurdert til å være god i flere studier (Feiring et al., 1990; Pincivero et al., 1997; McCleary & Andersen, 1992; Impellizzeri et al., 2008). Derimot er IKD sin begrensning at apparatet er dyrt, det er ikke egnet for å fraktes og trenger omfattende opplæring av testere og deltakere ved bruk (Opar et al., 2013). IKD er likeledes en tidkrevende metode, med opp-til 25 minutter testtid per deltaker, som gjør det til en mindre egnet metode i klinisk setting hvor en større gruppe skal testes (Whiteley et al., 2012).

NordBord er et nytt apparat som er blitt utviklet for å teste hamstringsstyrke gjennom samme bevegelse som *Nordic hamstrings* øvelsen (Wiesinger et al., 2020). Det er lite, enkelt å frakte, rimeligere og mer tidseffektivt enn IKD, med rundt 2 minutters testtid (Opar et al., 2015). Dette har ført til at flere nyere studier har benyttet seg av NordBord for å måle fotballspilleres eksentriske hamstringsstyrke (Timmins et al., 2016; Pollard et al., 2019; Presland et al., 2018; Buchheit et al., 2016).

To studier har vurdert test-retest reliabiliteten til NordBord og disse er presentert i tabell 1 (Wiesinger et al., 2020; Opar et al., 2013). Opar og medarbeidere (2013) fant moderat til høy test-retest reliabilitet for målinger bilateralt, men svak reliabilitet under unilaterale målinger. Wiesinger og medarbeidere (2020) sammenlignet NordBord med IKD og fant en svak korrelasjon mellom apparatene. De konkluderte med at både NordBord og IKD har begrensinger for å måle styrkeverdier på individnivå i hamstrings basert på en MDC95%-verdi for NordBord på 19,1% og 15,0% for henholdsvis venstre og høyre bein.

Tabell 1: Reliabilitet for eksentrisk hamstringsstyrke med NordBord

Studie	Metode	Utvalg	Reliabilitet
Opar et al. (2013)	<p><i>Apparat:</i> Belastningsceller (Delphi Measurement Pty Ltd, Gold Coast, Australia)</p> <p><i>Design:</i> Tre testdager. Første tilvenning og de to andre gjeldende tester.</p> <p><i>Test:</i> Isotonisk NH</p> <p><i>Variabel:</i> N</p> <p><i>Testprosedyre:</i> 2 sett, 3 maks rep, 2min pause mellom sett.</p>	<p>Friske, n=30 godt trente mannlige idrettsutøvere (Amerikansk fotball, Rugby, Fotball, friidrett)</p>	<p>Bilateral testing av maksimalkraft (N)</p> <p><i>Venstre bein:</i> Effekt størrelse = -0,12, ICC = 0,83 (0,67-0,91) SEM = 27,5, SEM% = 8,5%, MDC = 76,2</p> <p><i>Høyre bein:</i> Effekt størrelse = -0,20, ICC = 0,90 (0,81-0,95), SEM = 21,7, SEM% = 5,8%, MDC = 60,1</p> <p>Bilateral testing av gjennomsnittlig kraft (N)</p> <p><i>Venstre bein:</i> Effekt størrelse = -0,13, ICC = 0,85 (0,71-0,93), SEM = 24,7, SEM% = 8,4%, MDC = 68,5</p> <p><i>Høyre bein:</i> Effekt størrelse = -0,18, ICC = 0,89 (0,78-0,95) SEM = 22,1, SEM% = 6,5%, MDC = 61,3</p>
Wiesinger et al. (2020)	<p><i>Apparat:</i> Belastningsceller (Megatron Elektronik GmbH & Co. KG)</p> <p><i>Design:</i> Tre testdager, 72timer mellom hver test. Tilvennings periode 1uke</p> <p><i>Test:</i> Isotonisk NH</p> <p><i>Variabel:</i> NM</p> <p><i>Testprosedyre:</i> Oppv:2 forsøk 80% av maks. Test: 2 maks rep, 1min pause mellom sett</p>	<p>Friske n=25 mannlige idrettsutøvere. 26 (3,0) år</p>	<p>Bilateral testing av maksimalt dreiemoment (NM)</p> <p><i>Venstre bein:</i> ICC_{3,1} = 0,90 (0,83-0,95), %CVme = 6,9%, MDC95% = 19,1%</p> <p><i>Høyre Bein:</i> ICC_{3,1} = 0,94 (0,89-0,97), %CVme = 5,5%, MDC95% = 15,0%</p>

N, newton; *NM*, newton meter; *ICC*, intraklasse korrelasjons koeffisient; *MDC*, minste reelle endring; *CV*, variasjonskoeffisient; *NH*, Nordic hamstrings; *SEM*, standardfeil til målingen; *oppv*, oppvarming; *rep*, repetisjon; *min*, minutter.

Isometrisk adduksjons-, abduksjons-, innadrotasjons-, utadrotasjons- og fleksjonsstyrke i hoften

Flere måleinstrumenter har blitt brukt for å vurdere hofte- og lyskestyrke. Blant de vanligste er IKD, HHD og ForceFrame. IKD har blitt vurdert til å være valid og reliabel og er gullstandarden for testing av hofte- og lyskestyrke (Stark et al., 2011). Svakheter ved å anvende IKD som målemetode er de samme som ble nevnt under testing av eksentrisk hamstringstyrke (Desmyttere et al., 2019). Som et resultat av dette har HHD og ForceFrame blitt populært å bruke. HHD er rimelig, bærbart og enkelt å bruke. Apparatet har vist en god test-retest reliabilitet for å måle isometrisk hoftestyrke i flere studier (Ishøi et al., 2019; Charlton et al., 2017; Thorborg et al., 2010a). På den andre siden er det usikkerhet rundt apparatets inter-rater reliabilitet, da det er funnet systematisk feil forbundet til tester sin styrke (Thorborg et al., 2013; Kemp et al., 2013).

ForceFrame er et nytt apparat som er blitt utviklet for å teste isometrisk styrke av en rekke muskelgrupper i underekstremitetene og det kan måle styrke bilateralt. I likhet med IKD, eliminerer ForceFrame systematisk feil forbundet til testers styrke. Apparatet er også bærbart og trenger ingen omfattende opplæring eller kompetanse for å brukes (O'Brian et al., 2018). Test-retest reliabiliteten til ForceFrame er blitt vurdert i to studier (Ryan et al., 2019; Desmyttere et al., 2019). ForceFrame har som HHD blitt anvendt for å måle hoftefleksjonsstyrke, men med svakere reliabilitet (Desmyttere et al., 2019).

Studier som har undersøkt reliabiliteten for abduksjons- og adduksjonsstyrke med ForceFrame er presentert i tabell 2. Begge studiene ble gjennomført på menn og hadde få deltakere (n=18 og n=20). Studiene gjennomførte test-retest samme dag og resultatene viser akseptabel test-retest reliabilitet for adduksjonsstyrke med en ICC fra 0,85 til 0,95 (95% KI: 0,74-0,97) og en SEM% fra 6,3% til 8,2% (Desmyttere et al., 2019; Ryan et al., 2019). For abduksjonsstyrke viser målingene en ICC fra 0,82 til 0,90 (95% KI: 0,67-0,95) og en SEM% fra 4,7% til 6,2%. Bedre ICC og SEM% verdier ble oppnådd ved å beregne gjennomsnitt av maksimalkraften fra tre repetisjoner for både adduksjons- og abduksjonsstyrke (Desmyttere et al., 2019). Studiene regnet ikke ut MDC. Begge studiene konkluderte med at ForceFrame er et reliabelt apparat for å måle styrken i hoften, og er et alternativ til gullstandarden IKD for måling av mannlige fotballspillere.

Det er én studie som har undersøkt test-retest reliabiliteten for utad- og innadrotasjonsstyrke med ForceFrame (tabell 3) (Desmyttere et al., 2019). Studien ble gjennomført på menn og hadde 20 deltakere. For utadrotasjonsstyrke viser resultatene en ICC fra 0,88 til 0,95 (95%

KI: 0,77-0,97) og en SEM% fra 4,1% til 5,7%. For innadrotasjonsstyrke viser resultatene en ICC fra 0,77 til 0,90 (95% KI: 0,60-0,95) og en SEM% fra 8,8% til 13,0%. Bedre ICC og SEM% verdier ble oppnådd ved å beregne gjennomsnitt av maksimalkraften fra tre repetisjoner for både innad- og utadrotasjonsstyrke. Studien regnet ikke ut MDC. Forskerne konkluderte med at ForceFrame har god reliabilitet for å vurdere utad- og innadrotasjonsstyrke når gjennomsnittet av maksimalkraften for tre repetisjoner blir brukt i analysen.

Studier som har undersøkt test-retest reliabiliteten for hoftefleksjonsstyrke med HHD er presentert i tabell 4 (Thorborg et al., 2010a; Ishøi et al., 2019). Begge studiene er gjennomført på både menn og kvinner og hadde få deltakere (n=17 og n=9) (Thorborg et al., 2010a; Ishøi et al., 2019). Resultatene viser en ICC fra 0,91 til 0,95 (95% KI: 0,66-0,98) og en SEM% fra 4,7% til 5,9%. Ishøi og medarbeidere (2019) fant en MDC 95% på 16,4% i sin studie. Med hensyn til MDC95% verdien konkluderte Ishøi og medarbeidere (2019) med at resultatene hadde liten klinisk betydning på individnivå med mindre testen ble brukt i forbindelse med skader, hvor det ofte blir vist store styrkeforskjeller fra før skade til etter skade. De inkluderte også en utregning av MDC95% basert på utvalgsstørrelsen i gruppen og fant da en MDC95% på 4,0%, illustrerer at testen var godt egnet til å vurdere styrke på gruppenivå.

Tabell 2: Reliabilitet for isometrisk abduksjons- og adduksjonsstyrke med ForceFrame/Groinbar

Studie	Metode	Utvalg	Reliabilitet
Ryan et al (2019)	<p><i>Apparat:</i> Groinbar (Vald Performance, Albion, Australia).</p> <p><i>Design:</i> Test-retest samme dag.</p> <p><i>Test:</i> Isometrisk styrketest</p> <p><i>Variabler:</i> N</p> <p><i>Testprosedyre:</i> ryggliggende under Groinbar, knær 60°. Oppv: 1 submaks (80%) + 1maks rep med hofteadd i 5sek</p>	<p>Friske, n= 18</p> <p>Profesjonelle Australske fotballspillere</p> <p>23 (2) år</p>	<p>Bilateral testing av maksimal kraft hofteadduksjon (N)</p> <p><i>Venstre bein:</i></p> <p>ICC = 0,95 (0,89-0,97)</p> <p>CV = 6,3%</p> <p><i>Høyre bein:</i></p> <p>ICC = 0,94 (0,88-0,97)</p> <p>CV = 6,7%</p>
Desmyttere et al (2019)	<p><i>Apparat:</i> Groinbar (Vald Performance, Albion, Australia)</p> <p><i>Design:</i> test-retest etter 15min med pause</p> <p><i>Test:</i> Isometrisk styrketest</p> <p><i>Variabler:</i> N·kg⁻¹</p> <p><i>Testprosedyre:</i> ryggliggende knær og hofte 45°. Sensorer på mediale femor condyles for adduksjon og på de laterale femor condylene for abduksjon. 3 maks rep hofteadd. i 5sek + 3 maks rep hofteabd. 10sek hvile mellom hver rep.</p>	<p>Friske, n= 20 mannlige fotballspillere.</p> <p>21 (2) år</p>	<p>Bilateral testing av maksimalkraft hofteadduksjon (N·kg⁻¹)</p> <p>ICC_{3,1} = 0,85 (0,74-0,92), SEM = 0,39, SEM% = 8,2%, 95% LOA = (lower -0,9, upper 1,2)</p> <p>Bilateral testing av gjennomsnittligkraft for hofteadduksjon av tre tester (N·kg⁻¹)</p> <p>ICC_{3,3} = 0,92 (0,85-0,93), SEM = 0,29, SEM% = 6,3%, 95% LOA = (lower -0,9, upper 1,2).</p> <p>Bilateral testing av maksimalkraft hofteabduksjon (N·kg⁻¹)</p> <p>ICC_{3,1} = 0,82 (0,67-0,90), SEM = 0,29, SEM% = 6,2%, 95% LOA = (Lower -0,6, upper 0,9).</p> <p>Bilateral testing av gjennomsnittligkraft for hofteabduksjon av tre tester (N·kg⁻¹)</p> <p>ICC_{3,3} = 0,90 (0,80-0,95), SEM = 0,21, SEM% = 4,7%, 95% LOA = (Lower -0,6, upper 0,9).</p>

KI, konfidensintervall; ICC, intraklasse korrelasjons koeffisient; N, newton; SD, standardavvik; SEM, standardfeil til målingene; LOA, limits of agreement; N.kg⁻¹, newton/kg⁻¹ oppv, oppvarming; rep, repetisjon; min, minutter

Tabell 3: Reliabilitet for isometrisk utad- og innadrotasjonsstyrke med ForceFrame/Groinbar

Studie	Type test	Utvalg	Reliabilitet
Desmyttere et al (2019)	<p><i>Apparat:</i> Groinbar (Vald Performance, Albion, Australia)</p> <p><i>Design:</i> test-retest etter 15min pause</p> <p><i>Test:</i> Isometrisk styrketest</p> <p><i>Variabel:</i> N·kg⁻¹</p> <p><i>Testprosedyre:</i> Ryggliggende med knær og hofte på 90°. Sensorer ble plassert på mediale malleol for UR. Sensorene ble plassert på laterale malleol for IR. 3 maks rep UR. i 5sek + 3 maks rep IR. 10sek hvile mellom hver rep.</p>	<p>Friske, n=20 mannlige fotballspillere</p> <p>21 (2) år</p>	<p>Bilateral testing av maksimal kraft UR (N·kg⁻¹) ICC_{3,1} 0,88 (0,79-0,94), SEM = 0,10, SEM% = 5,7%, 95% LOA (Lower -0,3, upper 0,3)</p> <p>Bilateral testing av gjennomsnittlig kraft for UR av tre tester (N·kg⁻¹) ICC_{3,3} 0,95 (0,90-0,97), SEM = 0,07, SEM% = 4,1%, 95% LOA (Lower -0,3, upper 0,3).</p> <p>Bilateral testing av maksimalkraft IR (N·kg⁻¹) ICC_{3,1} 0,77 (0,60-0,87), SEM = 0,28, SEM% = 13,0%, 95% LOA (lower -0,9, upper 0,6)</p> <p>Bilateral testing av gjennomsnittlig kraft for IR av tre tester (N·kg⁻¹) ICC_{3,3} 0,90 (0,80-0,95), SEM =0,17, SEM% = 8,8%, 95% LOA, (lower -0,7, upper 0,5)</p>

IR, innadrotasjon; UR; utadrotasjon, KI, konfidensintervall; ICC, intraklasse korrelasjons koeffisient; N, newton; SD, standardavvik; SEM, standardfeil til målingene; LOA, limits of agreement,; N.kg⁻¹, newton/vekt⁻¹; oppv, oppvarming; rep, repetisjon; min, minutter.

Tabell 4: Reliabilitet for isometrisk hoftefleksjonsstyrke målt med hånd-heldt dynamometer

Studie	Metode	Utvalg	Reliabilitet
Ishøi et al (2019)	<p><i>Apparat:</i> Hånd-Holdt dynamometer (Hoggan MicoFET2, Scientific L.L.C, Salt Lake City, USA)</p> <p><i>Design:</i> Test-retest med 30min mellomrom</p> <p><i>Test:</i> Isometrisk</p> <p><i>Variabel:</i> N</p> <p><i>Testprosedyre:</i> Deltaker på kanten av en undersøkelsesbenk. Hoften 90° flektert. Dynamometer 5cm over patella festet med et fiksert belte. Oppv: 1 submaks og 1 maksimalt forsøkt. Test: 3 maks rep av hoftefleksjon i 3-4sek. 60 sek pause mellom hver rep.</p>	<p>Friske, n= 17 (9menn, 8kvinner)</p> <p>25 (4) år</p>	<p>Testing av maksimalkraft (N)</p> <p>ICC_{2,1} 0,95 (0,88-0,98), SEM = 18,8, SEM% = 5,9%, MDC individuell =16,4%, MDC gruppe = 4,0%.</p>
Thorborg et al (2010)	<p><i>Apparat:</i> Hånd-heldt dynamometer (Power track II commander, Melbourne, Australia)</p> <p><i>Design:</i> test-retest med 1 uke mellom</p> <p><i>Test:</i> Isometrisk</p> <p><i>Variabel:</i> N</p> <p><i>Testprosedyre:</i> Deltakerne på kanten av en undersøkelsesbenk. Hoften 90° flektert, overkroppen strak. Dynamometer plassert 5cm ovenfor patella. Oppv: 1 submaks rep, Test: 4 maks rep av hoftefleksjon i 3-4 sek. 30 sek pause mellom reps.</p>	<p>Friske, n=9 (4kvinner, 5menn).</p> <p>Menn 27 (5) år</p> <p>Kvinner 25 (4) år</p>	<p>Test av maksimalkraft (N)</p> <p>ICC_{2,1} 0,91 (0,66-0,98), SEM = 13,5, SEM% = 4,9%, MDC 95 = 37,3.</p> <p>Test av gjennomsnittlig kraft (N) av tre beste forsøk.</p> <p>ICC_{2,3} 0,92 (0,70-0,98), SEM = 12,5, SEM% = 4,7%, MDC 95 = 34,5.</p>

HF, hoftefleksjon; *KI*, konfidensintervall; *ICC*, intraklasse korrelasjons koeffisient; *N*, newton; *SD*, standardavvik; *SEM*, standardfeil til målingene; *MDC*, minste reelle endring; *oppv*, oppvarming; *rep*, repetisjon; *sek*, sekunder

3. Metode

Denne masteroppgaven er en del av en større pågående prospektiv kohortstudie. Norges idrettshøgskole og Senter for idrettsskedeforskning er ansvarlig for prosjektet, mens Norges fotballforbund og Toppfotball kvinner er med som samarbeidspartnere. Hensikten med hovedprosjektet er å kartlegge skader og sykdommer for norske kvinnelige elitefotballspillere. Sekundærhensikten er å undersøke risikofaktor for underkstremitetsskader, med et spesielt fokus på lyske- og hamstringsskader. Målet med denne masteroppgaven var å undersøke uke-til-uke reliabiliteten for seks standardiserte tester av maksimalkraft i hamstrings, lyske og hofta hos kvinnelige fotballspillere i Toppserien, som brukes for å undersøke styrke som risikofaktor for hamstrings- og hofta-/lyskeskader. Studien ble utført som en test-retest reliabilitetsstudie, basert på data innsamlet i den prospektive kohortstudien og data fra en retest. Denne oppgaven følger retningslinjene som ble presentert av Kottner og medarbeidere (2011) for rapportering i reliabilitetsstudier (GRRAS).

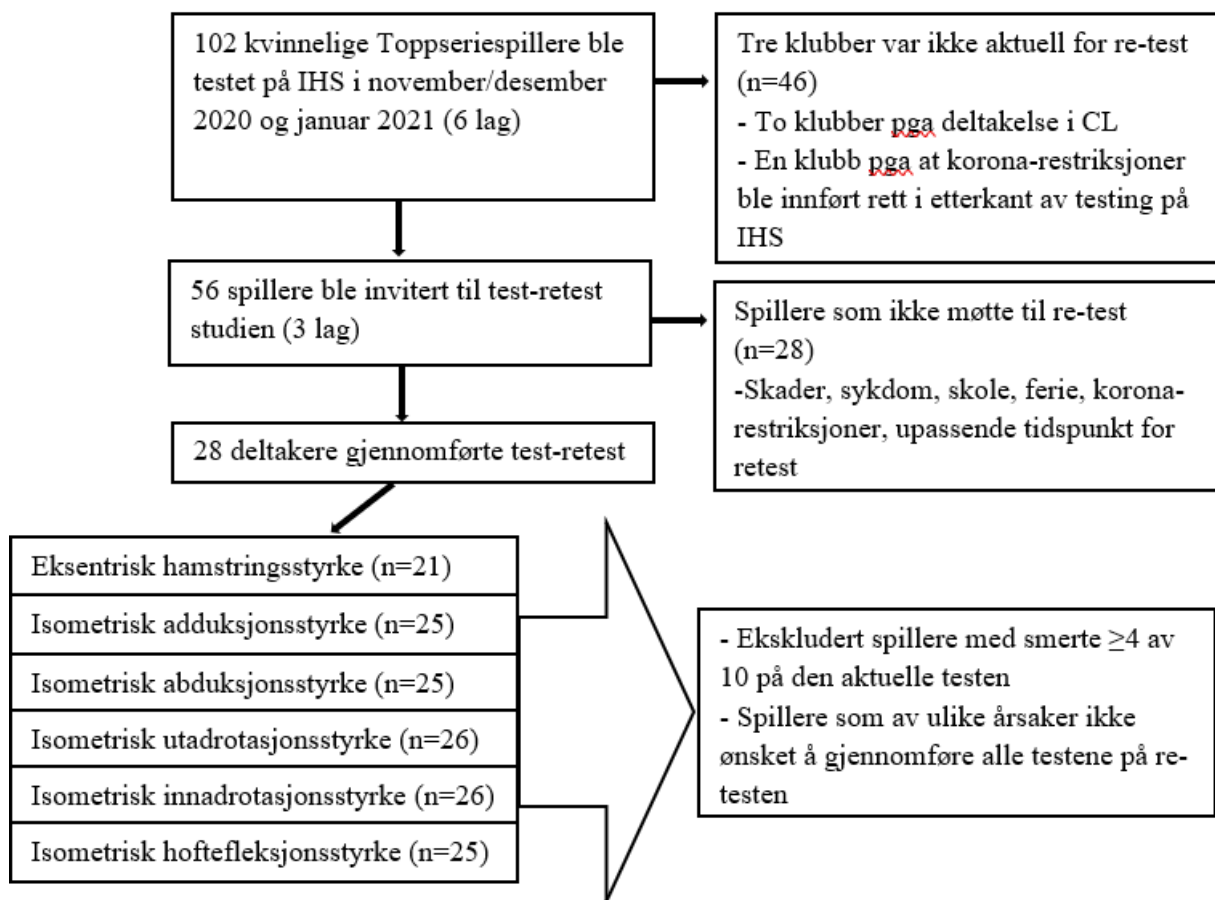
Før og etter hver sesong gjennomfører samtlige spillere i Toppserien et testbatteri som er utviklet av medisinsk personell på Idrettens helsesenter (IHS). Testene er valgt ut på bakgrunn av relevans for prestasjon og for å fungere som en målestokk for retur til idrett under rehabilitering ved en eventuell skade. Testene hvor test-retest reliabilitet ble beregnet var: 1) eksentrisk hamstringstyrke målt med NordBord, 2) isometrisk adduksjons-, 3) isometrisk abduksjons-, 4) isometrisk utadrotasjons-, 5) isometrisk innadrotasjonsstyrke i hoften målt med ForceFrame, og 6) isometrisk hoftefleksjonsstyrke målt med HHD.

3.1 Rekruttering

De 10 klubbene i Toppserien skulle etter planen møte opp til testing på Idrettens helsesenter (IHS) etter sesongslutt 2020. På grunn av korona-epidemien var det kun seks av lagene som møtte opp for fysisk testing. Vi inviterte tre av lagene basert i Oslo-området til å gjennomføre re-testing en uke etter testing på IHS i november/desember 2020 og januar 2021 for å undersøke reproduserbarheten til de nevnte testene.

Rekrutteringen til studien skjedde via invitasjon som ble sendt ut til lagenes fysiske trenere på e-post i forkant av testing på IHS. Deltakerne fikk informasjon om prosjektet og informert samtykke ble hentet inn elektronisk via AthleteMonitoring (FITSTATS Technologies, Inc, New Brunswick, Canada).

Inklusjonskriteriene for deltakelse i studien var kvinner som var i spillertropen til et Toppserie-lag ved sesongslutt 2020. Eksklusjonskriterier var selvrapporterte muskulære, kardiovaskulære eller nevrologiske problem, som hindret maksimal muskelkontrasjon på den første eller andre testdagen. Resultater fra hofteadduksjons- og fleksjonstesten ble også ekskludert fra analysen dersom deltaker rapporterte smerte ≥ 4 , på en skala fra 0-10, under den første eller andre testdagen. Dersom det var usikkert om en deltaker kunne gjennomføre en spesifikk test, ble det tatt en vurdering av fysisk trener og deltaker selv.



Figur 3: Flytdiagram for rekruttering

3.2 Testere

Det var to stipendiater og én mastergradsstudent ved Norges idrettshøgskole som gjennomførte testene. Testerne hadde øvd på hverandre og fått opplæring i bruk av apparater av personell på IHS i forkant av testingen. Det var samme tester som gjennomførte de samme målingene under den første og den andre testdagen. Alle testerne hadde gjennomført en

tilsvarende testprotokoll med alle 10 lagene i Toppserien i februar/mars 2020 og kjente derfor testprosedyrene godt.

3.3 Datainnsamling

Den første testdagen foregikk i lokalene til IHS på Ekeberg. Totalt brukte deltakerne rundt 1,5 timer på første testdag. De inkluderte lagene i studien fikk sin egen dag hvor de ble testet på IHS. Spillere gjennomførte testene i puljer bestående av 10-12 personer. Når deltakerne ankom IHS, ble først høyde målt med et bevegelig antropometer, deretter varmet deltakerne opp i 15 min før testene startet (figur 4).



Figur 4: Tidslinje over datainnsamlingen

Testbatteriet på IHS bestod av fem stasjoner (tabell 5), hvor spillerne ble fordelt tilfeldig ut i grupper på 2-3 personer på de ulike stasjonene av testlederne. Testene ble gjennomført parallelt. Når en gruppe var ferdig på en stasjon gikk de videre til neste stasjon. Deltakeren som startet å teste på en stasjon var først på neste stasjon, slik at hun fikk minimum 5 minutters pause mellom de ulike stasjonene.

Testdag to ble gjennomført syv dager etter første testdag i lagenes egne klubbhus. Protokollen her var lik som første testdag, bortsett fra at her ble bare stasjon 1-3 gjennomført, da det var disse reproduserbarheten ble undersøkt på. Totalt brukte deltakerne rundt 1 time på andre testdag.

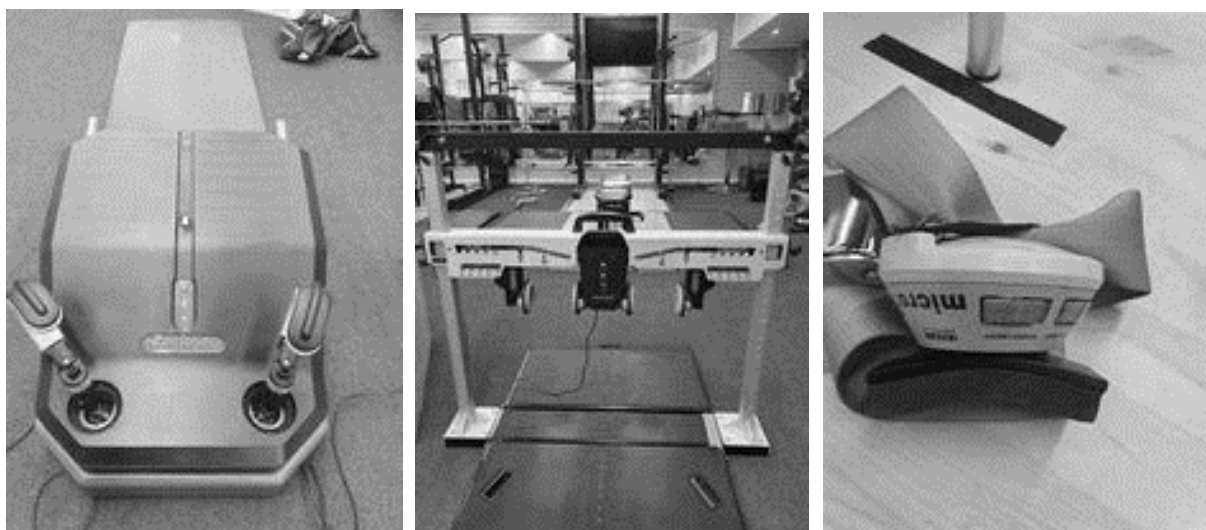
Hverken deltakerne eller testerne var blindet for resultatene. Testresultatene fra hver testdag ble ikke analysert eller brukt for å sammenligne resultater under testperioden.

Tabell 5: Oversikt over testbatteriet

Stasjon	Test
1	Eksentrisk hamstringsstyrke
2	ADD [*] -, ABD [^] -, IR [#] - og UR [⊠] -styrke i hoften
3	Hoftefleksjonsstyrke
4	Beinpress
5	Countermovement jump test (CMJ test).

* Adduksjon, ^ Abduksjon, #Innadrotasjon, ⊠ utadrotasjon

3.4 Måleinstrumenter



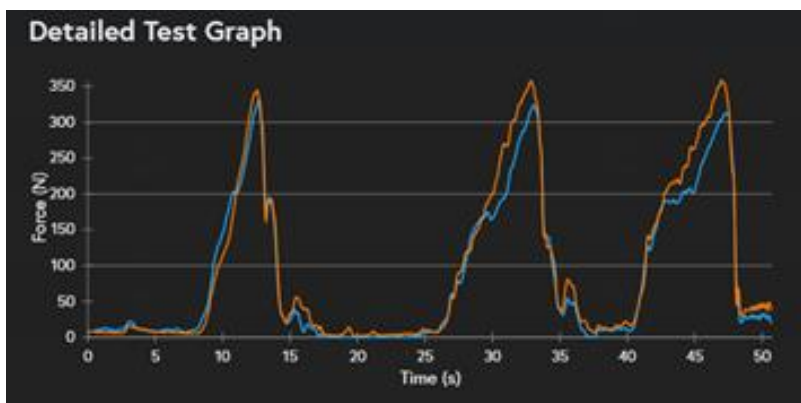
Figur 5: Illustrerer måleinstrumentene benyttet i denne studien., NordBord, ForceFrame og Hånd-heldt dynamometer sett fra venstre til høyre.

Eksentrisk hamstringsstyrke ble målt ved å benytte systemet NordBord (Vald Performance, Albion, Australia). NordBord er et 90 cm langt og 60 cm bredt polsrett Brett med to ankelkroker. Krokene er koblet til to kraftceller som måler kraftproduksjonen i venstre og høyre bein i Newton (N) under utførelse av *Nordic hamstrings* øvelsen. Med bruk av en USB-kabel koblet til en bærbar PC ble resultatene lastet opp til programmet *Vald Hub*.

ForceFrame (Vald Performance, Albion, Australia) ble i denne studien brukt til å måle isometrisk adduksjon-, abduksjon-, innadrotasjon- og utadrotasjonsstyrke i hoften. Dette er et

system som kan brukes til å måle isometrisk styrke for en rekke muskelgrupper i både under- og overekstremitetene. Systemet er designet slik at begge bein kan testes samtidig. Det består av en justerbar rigg som kan justeres både vertikalt og horisontalt slik at tester kan gjennomføres med de rette vinklene for deltakere med ulik kroppsstørrelse. ForceFrame har fire kraftceller, to på innsiden og to på utsiden som er festet til riggen, noe som gjør det mulig for deltakerne og presse både innover og utover under måling av muskelstyrke. Med bruk av en USB-kabel koblet til en bærbar PC, ble resultatene lastet opp til programmet *Vald Hub*.

Hoftefleksjonsstyrken ble målt med microFET2 (Hoggan Health Industries Inc. West Draper, UT, USA), som er et batteridrevet hånd-heldt apparat som måler maksimalkraft i Newton (N), for verdier opp til 890N (Pfister et al., 2018). Utgiveren oppgir selv at apparatet har en feilmargin innen 1%.



Figur 6: Illustrerer utdrag fra *Vald Hub* etter 3 repetisjoner med Nordic hamstrings testet med NordBord

Vald Hub er en sentralisert dataanalyse- og kontoadministrasjonsplattform for alle Vald Performance sine apparater. Systemet er passordbeskyttet, slik at det kun var de som analyserte resultatene som hadde tilgang. Tester fra deltakerne ble lastet opp til *huben* for videre analyse. Systemet gir informasjon om blant annet maksimalkraft (N) for begge bein, gjennomsnittlig kraftutvikling, sideforskjeller mellom bein og antall repetisjoner.

3.5 Testprotokoll

Etter at deltakerne hadde gjennomført en felles oppvarming i regi av lagets fysiske trener i 15 minutter startet testene. Alle testene hadde en øvelsesespesifikk standardisert oppvarming, med gradvis mer intensive repetisjoner på 50%, 75% og 95% av maksimal opplevd innsats. For hoftefleksjonstesten var det to oppvarmingsforsøk på 50% og 95% av maksimal opplevd innsats. Hensikten med prøvoforsøkene var å gjøre deltakerne kjent med testene.

Testposisjonen fra første testdag ble notert ned, slik at deltakerne kunne gjennomføre den samme testen med samme standardisert testposisjonen på re-testen.

Bilateral eksentrisk hamstringsstyrke

Eksentrisk hamstringsstyrke ble testet gjennom å utføre tre maksimale repetisjoner av *Nordic hamstrings* øvelsen i NordBord. Deltagerne knelte på NordBordet og anklene ble sikret superior for lateral malleol av individuelle låsbare seler festet til kraftceller som måler kraftutviklingen. Selene virket vinkelrett på leggen. Øvelsen startet med en kneposisjon på 90° og hoften og overkroppen måtte holdes i nøytral posisjon, med armene foldet på brystet (figur 7). Deltaker måtte deretter gradvis lene seg framover med lavest mulig hastighet mens hun motsto bevegelsen med begge beina.

Utførelsen av bevegelsen ble observert visuelt av testleder, og en repetisjon ble underkjent om deltakeren hadde en stor hoftefleksjon eller ikke kontrollerte nedstigningen fra starten av bevegelsen. En repetisjon ble akseptert når kraftutviklingen nådde en tydelig topp (indikerer maksimal eksentrisk styrke) og ble fulgt av en rask kraftnedgang. Denne kraftnedgangen oppstår når spillerne ikke lenger er i stand til å motstå effekten av tyngdekraften som virker på segmenter over kneleddet (Opar et al, 2013). For de deltakerne som klarte å motstå bevegelsen hele veien ble det muligens ikke en tydelig topp i kraftutvikling, men repetisjonen ble godkjent dersom hun kontrollerte bevegelsen hele veien. Metoden vi brukte for å teste eksentrisk hamstringsstyrke er beskrevet i tidligere studier (Opar et al., 2013; Wiesinger et al., 2020).

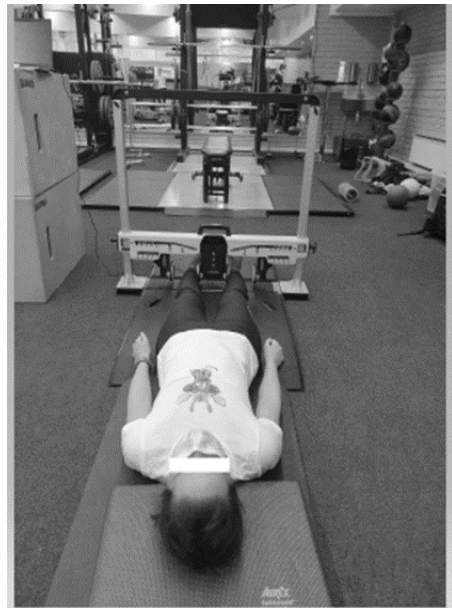
Bilateral isometrisk hofteadduksjonsstyrke

Testing av isometrisk hofteadduksjonsstyrke ble utført ryggliggende, med armene plassert på siden av kroppen. Hoftene og knærne var i nøytral posisjon 0°, med hodet hvilende på en balansepute (figur 8). Deltakeren ble instruert til å presse så hardt hun kunne i 3-4 sekunder mot de indre sensorene på ForceFrame. Testleder ga verbale instruksjoner for å motivere deltakeren gjennom hele testen. Godkjent test etter tre maksimale repetisjoner, med 5



Figur 7: Illustrerer eksentrisk hamstringstest med NordBord (Vald Performance, Albion, Australia)

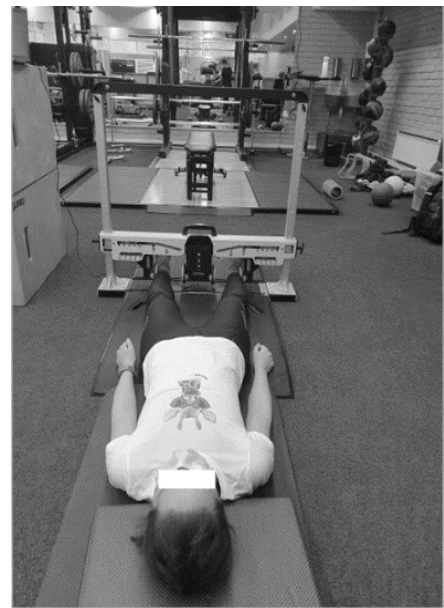
sekunder pause mellom hver repetisjon. Mediale malleol (innsiden av ankene) var i kontakt med de indre sensorene på ForceFrame på alle testforsøkene. Utførelsen av bevegelsen ble observert visuelt av testleder og dersom en deltaker løftet setet ble repetisjonen underkjent. Metoden som ble brukt for å måle hofteadduktorstyrke er beskrevet i tidligere studier (O'Brien et al., 2018; Desmyttere et al., 2019). Deltakerne ble også bedt om å rapportere symptomer på smerte ved bruk av en subjektiv skala fra 0-10 etter at de hadde gjennomført testen.



Figur 8: Illustrerer isometrisk hofteadduksjonstest med ForceFrame (Vald Performance, Albion, Australia)

Bilateral isometrisk hofteabduksjonsstyrke

Øvelsen ble utført rygliggende, med armene plassert ved siden av kroppen. Hoften og knærne var i nøytral posisjon 0° , med hodet hvilende på en balansepute (figur 9). Deltakeren ble instruert til å presse så hardt hun kunne i 3-4 sekunder mot de ytre sensorene på ForceFrame. Testleder ga verbale instruksjoner for å motivere deltakeren gjennom hele testen. Godkjent test etter tre maksimale repetisjoner, med 5 sekunder pause mellom hver repetisjon. Lateral malleol (utsiden av ankene) var i kontakt med de ytre sensorene på ForceFrame på alle testforsøkene. Utførelsen av bevegelsen ble observert visuelt av testleder og dersom en deltaker løftet setet ble repetisjonen underkjent. Metoden vi brukte for å måle hofteabduktorstyrke er beskrevet i tidligere studier (O'Brien et al., 2018; Desmyttere et al., 2019).



Figur 9: Illustrerer isometrisk hofteabduksjonstest med ForceFrame (Vald Performance, Albion, Australia)

Bilateral isometrisk utadrotasjonsstyrke i hoften

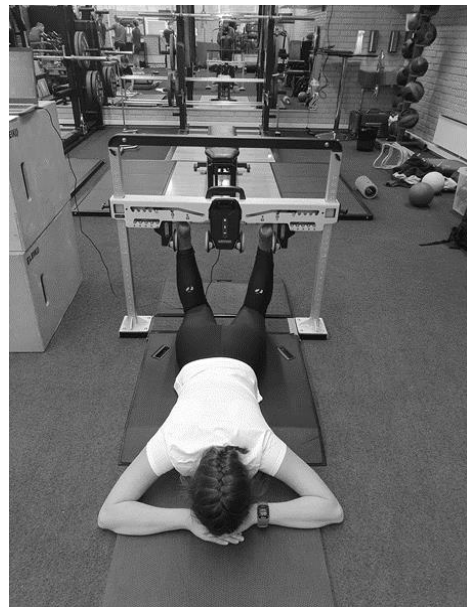
Øvelsen ble utført mageliggende, med armene foldet under hodet. Hofteposisjon var nøytral 0° og knærne på 90° (figur 10). Deltakeren ble instruert til å presse så hardt hun kunne i 3-4 sekunder mot de indre sensorene på ForceFrame. Testleder ga verbale instruksjoner for å motivere deltakeren gjennom hele testen. Godkjent test etter tre maksimale repetisjoner, med 5 sekunder pause mellom hver repetisjon. Mediale malleol (innsiden av anklene) var i kontakt med de indre sensorene på ForceFrame på alle testforsøkene. Utførelsen av bevegelsen ble observert visuelt av testleder og dersom en deltaker løftet setet ble repetisjonen underkjent. Metoden vi brukte for å måle utadrotasjonsstyrke i hoften er beskrevet i tidligere studier (O'Brien et al., 2019; Desmyttere et al., 2019).



Figur 10: Illustrerer isometrisk utadrotasjonstest med ForceFrame (Vald Performance, Albion, Australia)

Bilateral isometrisk innadrotasjonsstyrke i hoften

Øvelsen ble utført mageliggende, med armene foldet under hodet. Hofteposisjon var nøytral 0° og knærne på 90° (figur 11). Deltakeren ble instruert til å presse så hardt hun kunne i 3-4 sekunder mot de ytre sensorene på ForceFrame. Testleder ga verbale instruksjoner for å motivere deltakeren gjennom hele testen. Godkjent test etter tre maksimale repetisjoner, med 5 sekunders pause mellom hver repetisjon. Lateral malleol (utsiden av anklene) var i kontakt med de indre sensorene på ForceFrame på alle testforsøkene. Utførelsen av bevegelsen ble observert visuelt av testleder og dersom en deltaker løftet setet ble repetisjonen underkjent. Metoden vi brukte for å måle innadrotasjonsstyrke i hoften er beskrevet i tidligere studier (Desmyttere et al., 2019).



Figur 11: Illustrerer innadrotasjonstest med ForceFrame (Vald Performance, Albion, Australia)

Isometrisk hoftefleksjonsstyrke

Øvelsen ble utført ryggliggende med armene langs siden av en undersøkelsesbenk. Beinet som ble testet hadde en nøytral hofteposisjon med kneet hengende ned fra bordet på 90° (figur 12). Deltakeren ble instruert til å presse så hardt hun kunne i 3-4 sekunder mot dynamometeret.

Testleder ga verbale instruksjoner for å motivere deltakeren gjennom hele testen. Godkjent test etter tre maksimale repetisjoner på hvert bein, med 10 sekunders pause mellom hver repetisjon. Dersom høyeste resultat (N) ble oppnådd på tredje forsøk fikk deltakeren et fjerde forsøk og et femte forsøk dersom høyeste resultat ble oppnådd på det fjerde. Dynamometeret ble på alle testforsøkene plassert 5cm ovenfor patella. Det ble i forkant av testen brukt målebånd og sportsteip for å

markere posisjonen hvor dynamometeret ble plassert. Utførelsen av testen ble observert visuelt av testleder. Testen ble underkjent om deltakerne bøyde ryggen framover eller bakover under gjennomføringen for å skape kraft. For å motvirke bevegelse i overkroppen og hoften under testing ble det brukt fikserte belter illustrert på figur 12. Fiksert belte ble også brukt på dynamometeret for å forhindre at testleder sin styrke skulle påvirke resultatet (Thorborg et al., 2013). Resultater fra hver repetisjon ble notert ned på skjema og ført inn i Excel.

3.6 Databehandling

Resultatene fra samtlige tester er oppgitt i maksimalkraft (peak newton) og som gjennomsnittlig maksimalkraft (average peak newton). Førstnevnte representerer den høyeste kraften en deltaker oppnådde fra alle forsøkene under en test på en testdag. Sistnevnte representerer gjennomsnittlig maksimalkraft fra alle testforsøkene av adduksjons-, abduksjons-, innadrotasjons-, utadrotasjons- og eksentrisk hamstringsstyrke og de tre beste forsøkene fra hoftefleksjonstesten på en testdag. Høyre og venstre bein er analysert hver for seg i alle testene.

Resultatene fra *Nordic hamstrings* er i tillegg oppgitt som beste repetisjon totalt. Beste repetisjon totalt beskriver den av de tre forsøkene på en testdag som representerte den høyeste totalkraften (venstre bein + høyre bein)/2.



Figur 12: Illustrerer isometrisk hoftefleksjonstest med HDD (Hoggan Health Industries Inc. West Draper, UT, USA)

Det var kun testlederne som hadde tilgang til kontoen på ValdHub. For å sikre at data ikke gikk tapt, ble også maksimalkraften fra testene skrevet ned for hånd. Resultatene fra hofte-fleksjonstesten, høyde og vekt ble ført inn i Excel etter at deltakerne hadde blitt målt. All data fra prosjektet ble lagret på to krypterte minnepenner. Både minnepennene og test-arkene ble oppbevart i låsbart skap under hele prosjektet.

For å finne gjennomsnittlig maksimalkraft fra én test ble testrapporten fra deltakeren eksportert over i et Excel-dokument. Både NordBord og ForceFrame opererer med 50hz og dokumentet viser verdier for kraftutviklingen med 0,02 sekunders mellomrom. I dokumentet ble den høyeste verdien (N) som ble oppnådd for alle testforsøkene funnet, notert ned og ført inn i grovdata-skjemaet i Excel. For å redusere risikoen for systematisk feil i form av feilplotting, ble resultatene kontrollert to ganger, på ulike tidspunkt.

For at deltakerne skulle få noe utbytte av testingen fikk de vite resultatet etter testingen. Resultater fra testene ble også sendt ut til lagenes fysiske trenere i henhold til normal prosedyre på IHS. Resultatene kan være av betydning for fysiske trenere, som da kan avgjøre om det skal iverksettes spesifikk trening på individnivå eller gruppenivå, i forhold til resultatet deltakerne oppnådde.

3.7 Statistiske analyser

De statistiske analysene ble utført i Microsoft Excel (Microsoft Corp, Redmond, Washington, USA) og SPSS (IBM SPSS Statistics v. 24, Armonk, New York, USA). Shapiro-Wilk test ble benyttet for å undersøke normalfordelingen til testene. I tillegg ble histogrammer undersøkt, og forskjellen mellom median og gjennomsnitt studert for å undersøke normalfordelingen. For maksimalkraft var alle 24 variablene normalfordelte (shapiro-wilk test $\geq 0,05$). For gjennomsnittlig maksimalkraft var det 1 av 24 variabler som ikke var normalfordelt (innadrotasjon for høyre bein ved første test ($p=0,046$)). Det ble derfor besluttet å bruke parametriske analyser for alle variablene.

For å vurdere systematiske forskjeller mellom test-retest ble paret t-test utført for både maksimalkraft og gjennomsnittlig maksimalkraft for alle testene. For å undersøke om utvalget var representativt for populasjonen ble uavhengig t-test benyttet. Første testdagen til utvalget ble analysert i forhold til de øvrige Toppseriespillerne som ble testet i november/desember 2020. Variablene som ble undersøkt var alder, høyde, vekt, KMI og maksimalkraft fra alle testene. Undersøkelsen ble gjennomført for å studere om utvalget skilte seg ut fra store deler av populasjonen. For alle de statistiske analysene ble $P \leq 0,05$ satt som signifikansnivå.

Vi benyttet ICC for å utrykke relativ reliabilitet. For maksimalkraft ble en toveis blandet variasjonsanalyse brukt for å vurdere test-retest verdier ($ICC_{3,1}$). For gjennomsnittlig maksimalkraft ble ($ICC_{3,k}$) brukt for å ta hensyn til flere styrkevariabler. For alle ICC analysene ble enighet (agreement) brukt for å kunne vurdere både systematiske og tilfeldige målefeil (Weir, 2005). ICC verdiene ble oppgitt med tilhørende 95% konfidensintervall. Avgrensningene for ICC ble satt til $ICC < 0,5$ svak reliabilitet, $0,5$ til $0,75$ moderat reliabilitet, $0,75$ til $0,90$ god reliabilitet, og $ICC > 0,90$ høy reliabilitet (Koo & Li, 2016).

Vi benyttet SEM/SEM%, ME/CV_{ME} og LOA for å utrykke absolutt reliabilitet. SEM ble regnet ut som beskrevet av (Weir, 2005). SEM og LOA er også de mest brukte målene på absolutt reliabilitet i studiene vi sammenligner oss med. Det ble i denne studien valgt å prioritere SEM/SEM% og LOA i resultatdelen. ME/CV_{ME} er ikke direkte sammenlignbare med SEM/SEM%, men viste ganske like verdier. Leseren henvises til vedlegg 6 for å studere disse verdiene.

$$\textbf{Ligning 1: SEM} = SD_{\text{samlet}} * (\sqrt{1 - ICC})$$

$$\textbf{Ligning 2: SEM\%} = (SEM / \bar{X}) * 100$$

LOA blir fremstilt grafisk i Bland-Altman plott. Plottene viser \bar{d} mellom test-retest og de individuelle gjennomsnittsverdiene for hvert sett av målinger (Bland & Altman, 1986).

$$\textbf{Ligning 3: LOA} = \bar{d} \pm (1,96 * SD)$$

MDC er muligheten målemetoden har til å oppdage en reell endring. MDC kan gjøre tallene mer klinisk relevant, og gjør det enklere å sammenligne resultatene med tidligere studier på området (Mokkink et al., 2010). MDC kan bli utregnet med 90% med ($Z=1,69$) eller 95% ($Z=1,96$) (Portney & Watkins, 2009). Vi valgte å oppgi MDC95% for å danne et 95% konfidensintervall rundt målefeilen,

$$\textbf{Ligning 4: MDC 95} = 1,96 * SEM * \sqrt{2}$$

$$\textbf{Ligning 5: MDC 95\%} = (MDC / \bar{X}) * 100$$

3.8 Etikk

Deltakerne i reliabilitetsstudien ble inkludert via hovedprosjektet. Det var ikke nødvendig å søke om videre tillatelser for å teste, ettersom undersøkelse av reliabilitet var en del av dette prosjektet. Prosjektet ble godkjent av lokal etisk komité ved Norges idrettshøgskole (vedlegg 1), samt Norsk senter for forskningsdata (vedlegg 2). Prosjektet omfatter behandling av personvernsdata og helseopplysninger. I samsvar med Helsinkideklarasjonen har alle deltakerne mottatt skriftlig og muntlig informasjon om formålene og prosedyrene med prosjektet før de fikk utlevert informert samtykke til å delta (vedlegg 3). I samtykkeskjemaet kom det fram at deltakerne kunne trekke seg fra studien uten videre begrunnelse når de måtte ønske. Testledere og forskere som benyttet dataene var underlagt taushetsplikt. Studien er finansiert gjennom Senter for idrettsskedeforskning.

4. Resultater

4.1 Utvalg

Totalt 102 kvinnelige Toppseriespillere ble testet på IHS i november/desember 2020 og januar 2021. Fra de tre lagene som ble invitert til å gjennomføre test-retest var det 56 spillere. Totalt 28 av disse spillerne gjennomførte to tester. Disse 28 spillerne blir i denne oppgaven omtalt som «utvalget», mens de resterende spillerne som ble testet én gang blir omtalt som «øvrige spillere». Den yngste deltakeren fra utvalget var 16 år og den eldste 32 år.

Det var ingen forskjell mellom utvalget og øvrige spillere for deskriptive målinger (tabell 6).

Tabell 6: Deskriptive verdier for utvalget (n=28) og øvrige spillere (n=74) ved første test. Presentert med gjennomsnitt og standardavvik (SD)

	Utvalg	Øvrige spillere	P-verdi
Alder (år)	21 (3)	22 (4)	0,23
Vekt (kg)	63,4 (5,8)	65,9 (6,6)	0,67
Høyde (cm)	170 (6)	172 (6)	0,08
KMI* (kg/m²)	21,9 (1,5)	22,6 (1,7)	0,43

*KMI= kroppsmasseindeks

4.2 Forskjell mellom utvalget og øvrige spillere

Det ble vist en signifikant gruppeforskjell ($p \geq 0,01$) i abduksjonsstyrke for både venstre og høyre bein mellom utvalget og øvrige spillere (tabell 7). For alle de andre testene var det ingen gruppeforskjell.

Tabell 7: Forskjell fra styrketestene mellom utvalget og øvrige spillere. Presentert med gjennomsnitt og standardavvik (SD)

Test	Utvalg	Øvrige spillere	Differanse (95% KI)	n utvalg	n øvrige spillere	P-verdi
Nordic hamstrings						
- Venstre	320 (37)	331 (46)	11 (-11; 33)	21	66	0,31
- Høyre	335 (49)	349 (45)	14 (-9; 37)	21	66	0,22
Adduksjon						
- Venstre	155 (32)	165 (28)	10 (-3; 24)	25	72	0,14
- Høyre	154 (28)	166 (27)	12 (-1; 25)	25	72	0,06
Abduksjon						
- Venstre	140 (21)	156 (21)	16 (6; 26)	25	72	0,003*
- Høyre	145 (24)	159 (23)	14 (4; 25)	25	72	0,01*
Utadrotasjon						
- Venstre	130 (21)	138 (16)	8 (-1; 17)	26	44	0,07
- Høyre	128 (24)	136 (18)	8,3 (-2; 18)	26	44	0,10
Innadrotasjon						
- Venstre	134 (30)	137 (32)	3 (-12; 18)	26	44	0,70
- Høyre	141 (34)	137 (35)	-4 (-22; 13)	26	44	0,61
Hoftefleksjon						
- Venstre	525 (85)	549 (73)	24 (-11; 59)	25	73	0,18
- Høyre	549 (85)	563 (70)	15 (-19; 50)	25	73	0,37

* $P \leq 0,01$; KI= konfidensintervall

4.3 Test-retest reliabilitet for maksimalkraft (N)

Det var 1 av de 13 utfallsmålene som viste en systematisk forskjell ($p=0,02$) mellom testdagene; utadrotasjon for høyre bein (tabell 8). ICC_{3,1} viste moderat til høy reliabilitet og var fra 0,75 til 0,95 (95% KI: 0,51-0,98) for alle testene. SEM% viste en variasjon fra 3,2% til 8,4%. Samlet for alle styrketestene varierte den gjennomsnittlige differansen (\bar{d}) mellom testene fra -7,7N til 9,2N, og MDC95% varierte fra 9,0% til 23,2%.

Tabell 8: Test-retest reliabilitet for maksimalkraft (N)

Måling	Test (SD)	Retest (SD)	Diff (95% KI)	P- verdi	95% LOA	ICC _{3,1} (95% KI)	SEM (%)	MDC 95 (%)
Nordic hamstrings								
- Venstre	320 (37)	315 (40)	-5,1 (-13; 3)	0,21	-5,1 (-40; 30)	0,89 (0,75-0,95)	12,8 (4,0)	35,6 (11,2)
- Høyre	335 (49)	327 (44)	-7,7 (-16; 0)	0,06	-7,7 (-42; 26)	0,92 (0,80-0,97)	13,1 (4,0)	36,3 (11,0)
- Snitt av venstre+høyre	327 (40)	319 (39)	-7,6 (-13; 1)	0,06	-7,6 (-32; 17)	0,93 (0,80-0,97)	10,5 (3,2)	29,1 (9,0)
Adduksjon								
- Venstre	155 (32)	151 (32)	-3,6 (-9; 2)	0,19	-3,6 (-30; 22)	0,91 (0,80-0,96)	9,5 (6,2)	26,3 (17,2)
- Høyre	154 (28)	152 (31)	-2,1 (-6; 2)	0,31	-2,1 (-22; 18)	0,94 (0,87-0,97)	7,2 (4,7)	20,0 (13,1)
Abduksjon								
- Venstre	140 (21)	138 (20)	-2,2 (-5; 2)	0,10	-2,2 (-15; 11)	0,95 (0,88-0,98)	4,6 (3,3)	12,8 (9,2)
- Høyre	145 (24)	144 (25)	-1,1 (-5; 5)	0,53	-1,1 (-18; 15)	0,94 (0,88-0,98)	6,1 (4,2)	16,9 (11,7)
Utadrotasjon								
- Venstre	130 (21)	125 (21)	-4,3 (-9; 0)	0,06	-4,3 (-26; 17)	0,85 (0,68-0,93)	8,0 (6,3)	22,2 (17,4)
- Høyre	128 (24)	124 (21)	-3,5 (-7; -1)	0,02 *	-3,5 (-18; 11)	0,94 (0,84-0,97)	5,5 (4,4)	15,2 (12,1)
Innadrotasjon								
- Venstre	134 (30)	131 (35)	-3,0 (-9; 4)	0,34	-3,0 (-34; 28)	0,88 (0,75-0,94)	11,1 (8,4)	30,8 (23,2)
- Høyre	141 (34)	143 (38)	-2,3 (-8; 4)	0,44	-2,3 (-32; 27)	0,91 (0,82-0,96)	10,8 (7,6)	29,9 (21,0)
Hoftefleksjon								
- Venstre	525 (85)	526 (79)	1,5 (-20; 23)	0,89	1,5 (-101; 104)	0,80 (0,60-0,91)	36,7 (6,9)	101,7 (19,4)
- Høyre	548 (85)	557 (86)	9,2 (-16; 34)	0,45	9,2 (-110; 128)	0,75 (0,51-0,88)	42,7 (7,7)	118,4 (21,4)

* $P \leq 0,05$; SD, standardavvik; Diff, differanse; KI, konfidensintervall; LOA, limits of agreement = $\bar{d} \pm (1,96 * SD)$; ICC, intraklasse korrelasjonkoeffisient, (3,1) toveis-tilfeldig, enighet, enkel måling; SEM, standardfeil til målingen = $SD * \sqrt{(1-ICC)}$; SEM% = $(SEM / \bar{X}) * 100$; MDC 95, minste reelle endring = $1,96 * \sqrt{2} * SEM$; MDC 95% = $(MDC / \bar{X}) * 100$; snitt av venstre + høyre = Beste repetisjon totalt av tre repetisjoner med nordic hamstrings (Venstre bein + Høyre bein)/2

4.4 Test-retest reliabilitet for gjennomsnittlig maksimalkraft (N)

Det var ingen av testene for gjennomsnittlig maksimalkraft som viste systematisk forskjell mellom testdagene (tabell 9). ICC_{3,k} viste god til høy reliabilitet, og var fra 0,87 til 0,98 (95% KI: 0,72-0,99). SEM% viste en variasjon fra 2,2% til 6,1%. Samlet for alle styrketestene varierte den gjennomsnittlige differansen (\bar{d}) fra -7,2N til 13,2N og MDC95% varierte fra 6,0% til 17,8%.

Tabell 9: Test-retest reliabilitet for gjennomsnittlig maksimalkraft (N) fra målingene

Måling	Test (SD)	Retest (SD)	Diff (95% KI)	P-verdi	95% LOA	ICC _{3,k} (95% KI)	SEM (%)	MDC 95 (%)
Nordic hamstrings								
- Venstre	312 (38)	306 (37)	-6,3 (-13; 1)	0,09	-6,3 (-37; 26)	0,95 (0,87-0,98)	8,4 (2,7)	23,3 (7,6)
- Høyre	324 (48)	317 (43)	-7,2 (-15; 1)	0,07	-7,2 (-41; 26)	0,96 (0,90-0,98)	9,1 (2,8)	25,2 (7,9)
- Snitt av venstre+høyre	318 (42)	311 (38)	-6,7 (-13; 0)	0,07	-6,7 (-31; 18)	0,97 (0,90-0,99)	6,9 (2,2)	19,1 (6,0)
Adduksjon								
- Venstre	148 (29)	145 (32)	-3,2 (-8; 2)	0,22	-3,2 (-28; 22)	0,96 (0,90-0,98)	6,2 (4,2)	17,1 (11,7)
- Høyre	148 (27)	147 (31)	-1,2 (-5; 3)	0,57	-1,2 (-22; 19)	0,97 (0,93-0,99)	5,0 (3,4)	13,9 (9,4)
Abduksjon								
- Venstre	136 (21)	133 (20)	-2,2 (-5; 1)	0,10	-2,2 (-15 ;11)	0,97 (0,93-0,99)	3,6 (2,5)	10,0 (7,4)
- Høyre	140 (23)	139 (25)	-0,6 (-4; 3)	0,74	-0,6 (-15 ;14)	0,98 (0,95-0,99)	3,4 (2,4)	9,4 (6,7)
Utadrotasjon								
- Venstre	123 (21)	120 (19)	-3,1 (-8; 2)	0,17	-3,1 (-25 ;19)	0,91 (0,81-0,96)	6,0 (4,9)	16,6 (13,6)
- Høyre	122 (23)	119 (21)	-3,2 (-7; 1)	0,06	-3,2 (-20 ;13)	0,96 (0,91-0,98)	4,4 (3,7)	12,2 (10,1)
Innadrotasjon								
- Venstre	127 (29)	123 (32)	-4,0 (-6; 1)	0,18	-4,0 (-33 ;25)	0,93 (0,85-0,97)	8,0 (6,1)	22,2 (17,8)
- Høyre	133 (33)	135 (35)	2,4 (-3; 8)	0,40	2,4 (-25; 30)	0,96 (0,90-0,98)	6,9 (5,2)	19,1 (14,3)
Hoftefleksjon								
- Venstre	511 (83)	509 (79)	-1,9 (-21; 17)	0,84	-1,9 (-95 ;91)	0,91 (0,79-0,98)	24,6 (4,7)	68,2 (13,4)
- Høyre	524 (84)	538 (88)	13,2 (-10; 37)	0,26	13,2 (-99 ;125)	0,87 (0,72-0,94)	31,0 (5,9)	85,9 (16,2)

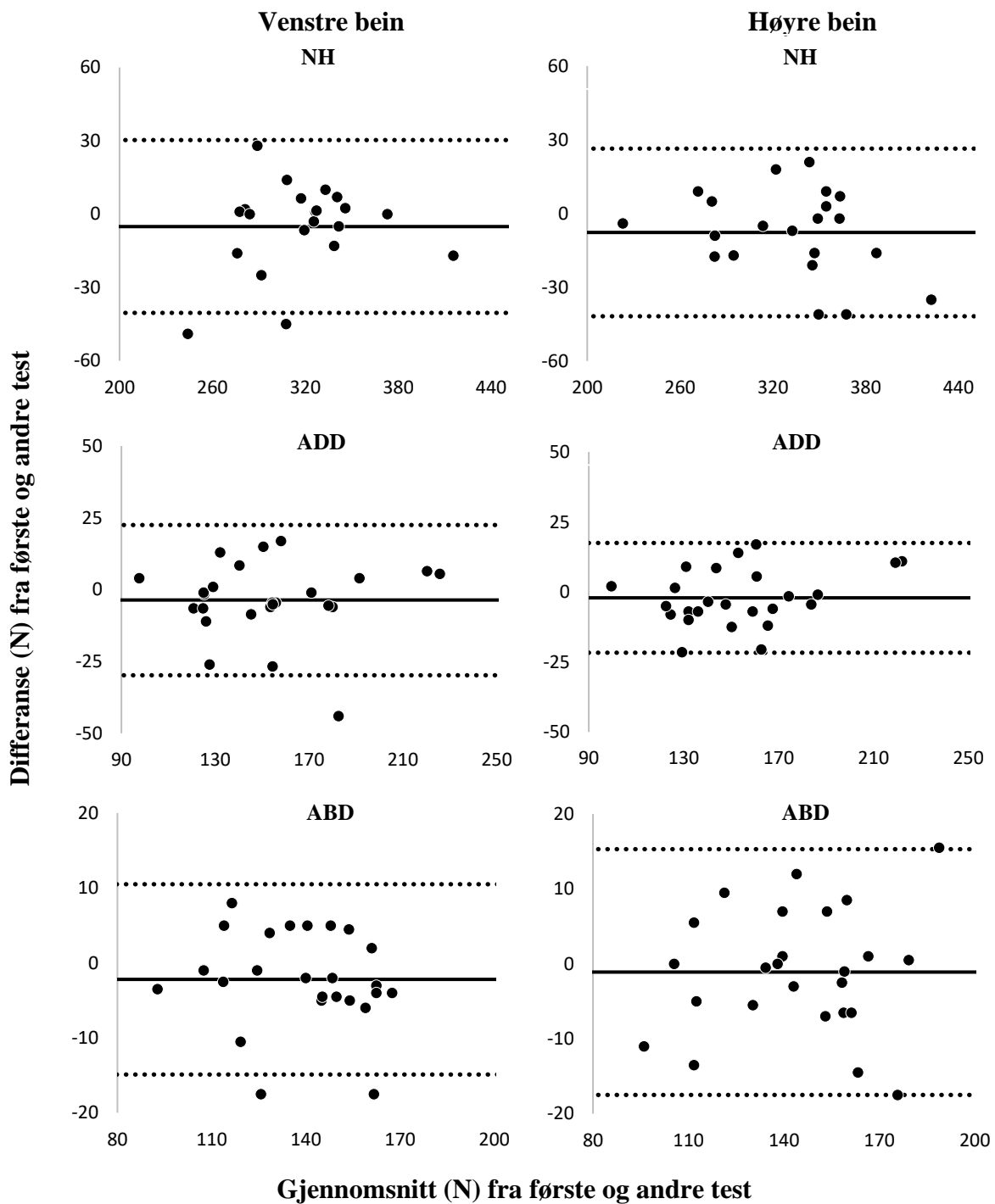
* $P \leq 0,05$; SD, standardavvik; KI, konfidensintervall; Diff, differanse; ICC_{3,k}, intraklasse korrelasjonkoeffisient, toveis-tilfeldig, enighet, gjennomsnittlig måling; KI, konfidensintervall; SEM, standardfeil til målingen = $SD \cdot \sqrt{1-ICC}$; SEM% = $(SEM/\bar{X}) \cdot 100$; MDC 95, minste reelle endring = $1.96 \cdot \sqrt{2} \cdot SEM$; MDC 95% = $(MDC/\bar{X}) \cdot 100$; LOA, limits of agreement = $\bar{d} \pm (1,96 \cdot SD)$; snitt av venstre+høyre = Alle repetisjonene for (høyre + venstre bein)/2

4.5 Bland-Altman plott

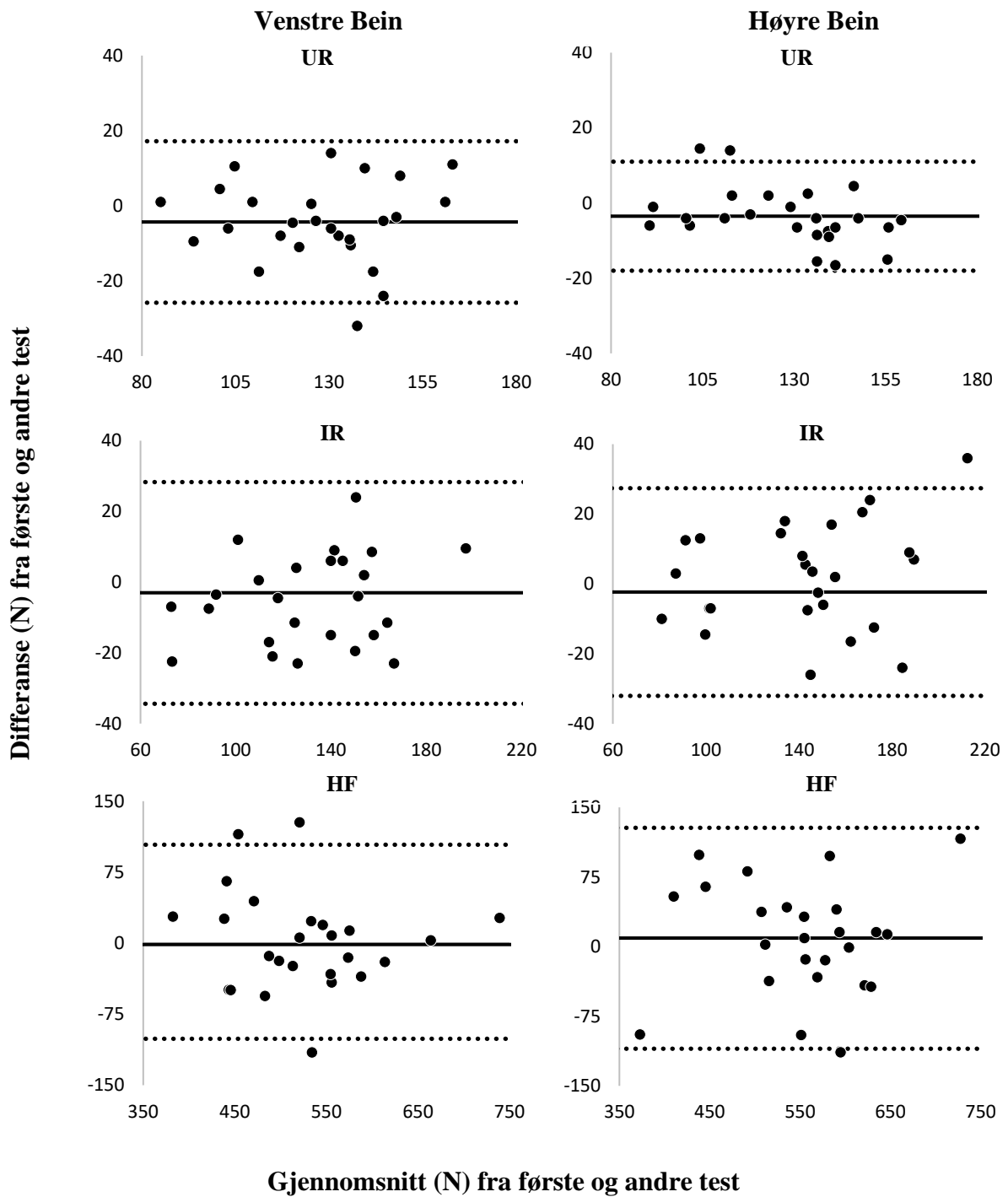
Det er laget 24 Bland-Altman plott fra testene. Tolv av plottene illustrer resultater for maksimalkraft (figur 13 og 14), mens de resterende 12 illustrer resultater for gjennomsnittlig maksimalkraft (figur 15 og 16). Plottene er basert på 95% LOA som er oppgitt i tabell 8 og 9. Felles for alle plottene var at den tilfeldige målefeilen var langt større enn den systematiske. De fleste plottene viser en gjennomsnittslinje under 0, indikerer at deltakerne skåret høyest resultat (N) på den første testdagen.

Heteroskedastisitet kan bli undersøkt visuelt av leseren i Bland-Altman plottene som er illustrert på de neste sidene.

4.5.1 Maksimalkraft (N)

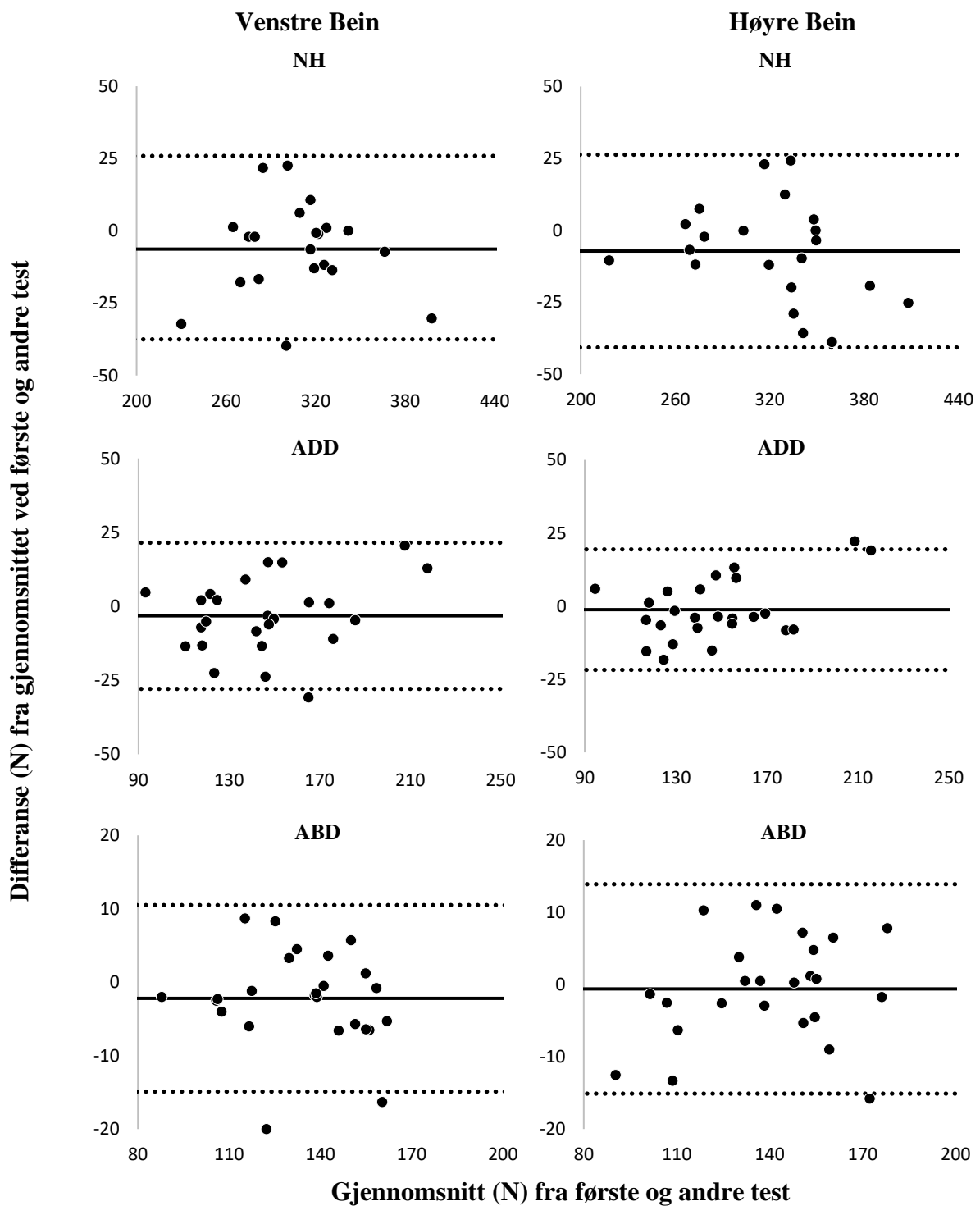


Figur 13 Bland-Altman plott som illustrerer forskjellen mellom test og retest målinger for maksimal eksentrisk hamstringstyrke (NH) ($n=21$), maksimal isometrisk adduksjonsstyrke (ADD) ($n=25$) og maksimal isometrisk abduksjonsstyrke (ABD) ($n=25$). Den solide linjen representerer den gjennomsnittlige forskjellen mellom testene (mean bias), mens de stiplede linjene representerer 95% referanseintervall for målingene. De svarte prikkene illustrer resultatet til deltakerne.

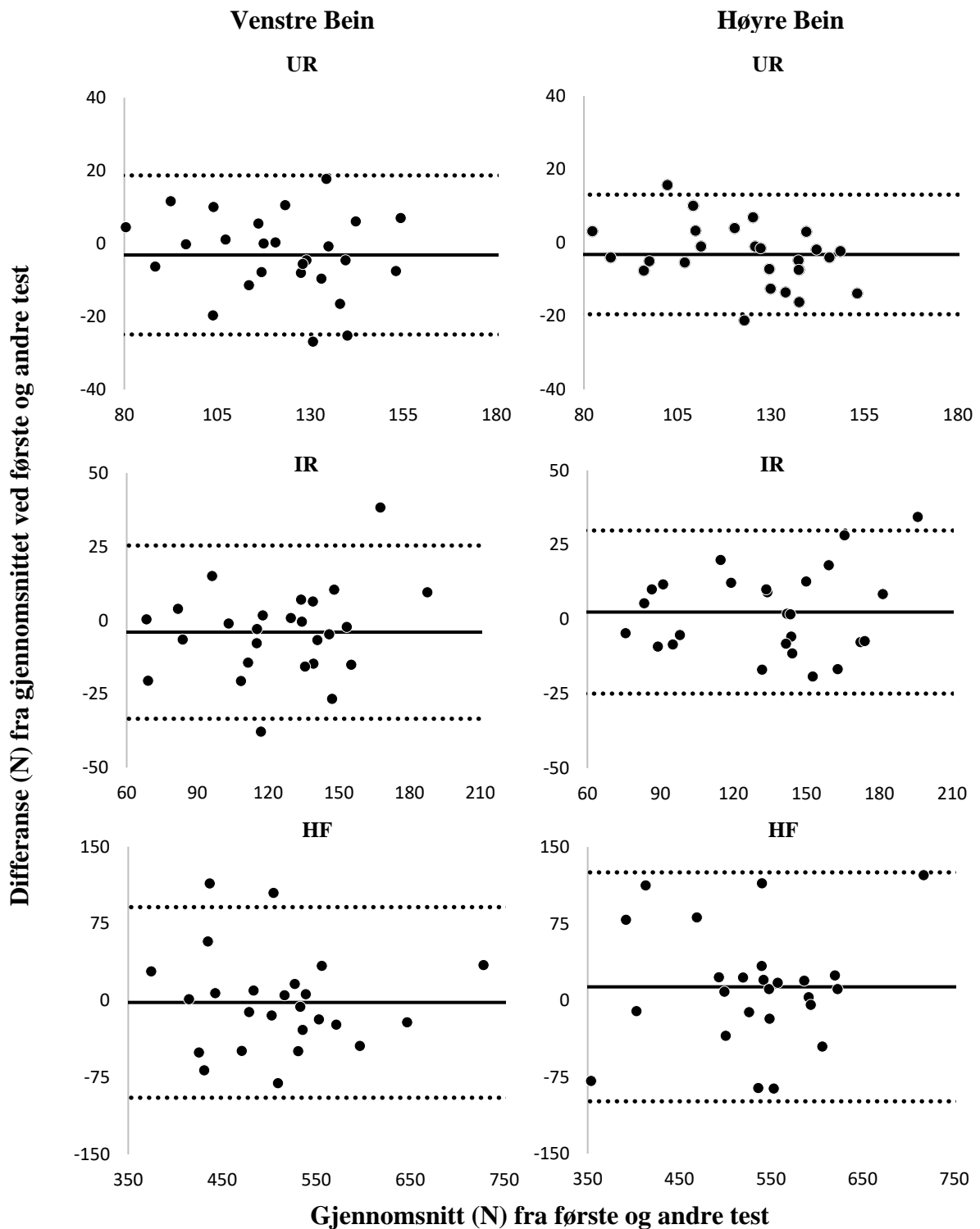


Figur 14: Bland-Altman plott som illustrerer forskjellen mellom test og retest målinger for maksimal isometrisk utadrotasjonsstyrke (UR) ($n=26$), maksimal isometrisk innadrotasjonsstyrke (IR) ($n=26$) og maksimal isometrisk hoftefleksjonsstyrke (HF) ($n=25$). Den solide linjen representerer den gjennomsnittlige forskjellen mellom testene (mean bias), mens de stiplede linjene representerer 95% referanseintervall for målingene. De svarte prikkene illustrer resultatet til deltakerne.

4.5.2 Gjennomsnittlig maksimalkraft (N)



Figur 15: Bland-Altman plott som illustrerer forskjellen mellom test og retest målinger for gjennomsnittlig maksimal eksentrisk hamstringstyrke (NH) ($n=21$), gjennomsnittlig maksimal isometrisk hofteadduksjonsstyrke (ADD) ($n=25$) og gjennomsnittlig maksimal isometrisk hofteabduksjonsstyrke (ABD) ($n=25$). Den solide linjen representerer den gjennomsnittlige forskjellen mellom testene (mean bias), mens de stiplede linjene representerer 95% referanseintervall for målingene. De svarte prikkene illustrer resultatet til deltakerne.



Figur 16: Bland-Altman plott som illustrerer forskjellen mellom test og retest målinger for gjennomsnittlig maksimal isometrisk utadrotasjonsstyrke (UR) (n=26), gjennomsnittlig maksimal isometrisk innadrotasjonsstyrke (IR) (n=26) og gjennomsnittlig maksimal isometrisk hoftedefleksjonsstyrke (HF) (n=25). Den solide linjen representerer den gjennomsnittlige forskjellen mellom testene (mean bias), mens de stiplede linjene representerer 95% referanseintervall for målingene. De svarte prikkene illustrer resultatet til deltakerne.

5. Diskusjon

Vi undersøkte test-retest reliabiliteten for én isotonisk og fem isometriske styrketester hos skadefrie kvinnelige fotballspillere i Toppserien. Hensikten med studien var å undersøke test-retest reliabiliteten mellom to testdager. Samtidig vil studien avdekke metodiske svakheter som kan utbedres til framtidige prosjekter.

5.1 Hovedfunn resultater

Den gjennomsnittlige differansen (\bar{d}) mellom testdagene varierte fra -7,7N til 9,2N for maksimalkraft (tabell 8) og fra -7,2N til 13,2N for gjennomsnittlig maksimalkraft (tabell 9). Måling av utadrotasjonsstyrke for høyre bein viste systematisk forskjell mellom test og retest ($p=0,02$), men for de andre testene var det ingen systematiske forskjeller. Testing av eksentrisk hamstringstyrke og abduksjonsstyrke viste smalest referanseintervall for 95% LOA, mens de største referanseintervallene ble oppgitt for innadrotasjons- og hoftefleksjonsstyrke. Resultater fra alle styrketestene viste at den tilfeldige feilen var langt større enn den systematiske. Det betyr at komponenter som biologisk variasjon (testere/deltakere), mekanisk variasjon eller uoverensstemmelser med testprotokollen hadde større påvirkning på resultatene enn for eksempel læringseffekt. Dette kan ha en sammenheng med at tilfeldig målefeil er vanskeligere å kontrollere for i en test-retest protokoll enn systematisk feil (Atkinson & Nevill, 1998; Portney & Watkins, 2009; Hopkins, 2000).

Som mål på relativ reliabilitet ble ICC utregnet. ICC viste moderat til høy reliabilitet for maksimalkraft med $ICC_{3,1}$ fra 0,75 til 0,95 (95% KI: 0,51-0,98). For gjennomsnittlig maksimalkraft varierte $ICC_{3,k}$ fra 0,87 til 0,98 (95% KI: 0,81-0,99). Dette indikerer god til høy reliabilitet (Koo & Li, 2016). For både maksimalkraft og gjennomsnittlig maksimalkraft ble lavest ICC oppnådd ved hoftefleksjonstesten. Gjennomsnittlig maksimalkraft oppnådde generelt høyere ICC enn maksimalkraft, dette har også blitt observert i tidligere studier (Desmyttere et al., 2019). Dette kan ha en sammenheng med hvordan ICC ble avlest i SPSS. Generelt vil verdier basert på $ICC_{3,k}$ som blir avlest med *average measures* gi høyere verdier enn om man baserer seg på $ICC_{3,1}$ og *singel measures*. Som beskrevet i teorien kan ICC bli påvirket av flere faktorer enn kun målefeilen mellom testene. Årsaken til dette er at selv om vi brukte enighet som definisjon, blir ICC påvirket av distribueringen til styrkeverdiene (Atkinson & Nevill, 1998). ICC sier heller ingenting om presisjonen til målingene uttrykt i en gitt målenhet (Atkinson & Nevill, 1998). De høye ICC-verdiene vi oppnådde i denne studien

kan derfor ha vært et resultat av variasjonen i utvalget og beskriver ikke nødvendigvis små individuelle forskjeller i testresultater. Vi valgte å oppgi ICC for å kunne sammenligne med tidligere studier og for å kunne få et mål på *between subject* variabilitet for vårt utvalg.

SEM% for maksimalkraft varierte fra 3,2% til 8,4% for alle testene (tabell 8). De laveste verdiene ble funnet for testing av eksentrisk hamstringsstyrke og abduksjonsstyrke, mens de høyeste verdiene ble funnet for målinger av innadrotasjons- og hoftefleksjonsstyrke. SEM% for alle testene lå under 10% som er foreslått som en kritisk verdi på god reliabilitet (Hopkins, 2000). På den andre siden illustrerte Atkinson & Nevill (2000) at en SEM% på under 10% ikke nødvendigvis betyr god reliabilitet. Som beskrevet i teorien viser SEM% kun 68% av variasjonen ved repeterte målinger og er basert rundt å finne en sann skår for målingene (Atkinson & Nevill, 1998). Selv om SEM% er under 10% kan flere av deltakerne ha større individuelle variasjoner i testresultatene. Dette ble også observert i vår studie og er presentert i Bland-Altman plottene i figur 13-16.

MDC/MDC% representerer testens egenskap til å oppdage en reell endring (Carter et al., 2011). MDC95% varierte fra 9,2% til 23,2% for maksimalkraft og fra 6,7% til 17,8% for gjennomsnittlig maksimalkraft. Lavest MDC95% ble oppgitt for test av eksentrisk hamstringsstyrke og abduksjonsstyrke. De høyeste MDC95%-verdiene ble oppgitt for måling av hoftefleksjons- og innadrotasjonsstyrke. Det er derfor usikkert om de sistnevnte testene er i stand til å oppdage endring på individnivå, da det må til en over 20% endring i styrkeverdier for å kunne fastslå at en reell endring har funnet sted for målinger av maksimalkraft (tabell 8). Testing av hoftefleksjon og innadrotasjon er derfor best egnet til å vurdere styrke på gruppenivå eller for å vurdere store endringer i styrkeverdier på individnivå, som kan forekomme i forbindelse med skader (Wiesinger et al., 2020).

5.2 Reliabilitet for styrkemålingene

5.2.1 Hvorfor det er utfordrende å sammenligne med tidligere studier

Ved undersøkelse av reliabilitet er det relevant å sammenligne med eksisterende studiers reliabilitet, for deretter å vurdere gyldigheten av våre resultater. Fra et litteratursøk fant vi to relevante studier for testing av eksentrisk hamstringsstyrke med NordBord (Opar et al., 2013; Wiesinger et al., 2020), to for testing av adduksjonsstyrke med Groinbar (Ryan et al., 2019; Desmyttere et al., 2019), én for testing av abduksjons-, innadrotasjons- og utadrotasjonsstyrke med Groinbar (Desmyttere et al., 2019) og to for testing av hoftefleksjonsstyrke med HHD (Ishøi et al., 2019; Thorborg et al., 2010a). Disse studiene er presentert i tabell 1-4. En direkte

sammenligning med tidligere studier er vanskelig. Deltakerne ble testet i ulike testposisjoner, hadde ulikt tidsintervall mellom tester, hadde et annet utvalg og brukte i noen tilfeller andre måleapparater. Studiene hadde også en større grad av standardisering mellom testene enn hva vi hadde i vår studie. Utvalgsstørrelsen i de inkluderte studiene er ulik fra vår, og de statistiske analysene er i flere tilfeller blitt gjennomført på en annen måte. Alle disse faktorene understreker viktigheten av å undersøke test-retest reliabilitet til den valgte metoden relatert til vår populasjon og måleapparatene som ble brukt. Vi kan ikke ta det som en selvfølge at vårt utvalg har tilsvarende variasjon i målingene som utvalget i tidligere studier. I følge Rankin og Stokes (1998) er en direkte sammenligning av forskjellige studiers reliabilitet ikke mulig, med mindre populasjonens sammensetting og utvalgsstørrelse er identiske.

5.2.2 Eksentrisk hamstringsstyrke

For målinger av eksentrisk hamstringsstyrke skyldes målefeilene hovedsakelig tilfeldig feil. Differansen mellom test-retest var $-5,1N$ (1,5%) og $-7,7N$ (2,4%) for målinger av maksimalkraft på henholdsvis venstre og høyre bein og viste ingen systematisk forskjell mellom testdagene (tabell 8). Det betyr at komponenter som læringseffekt og at testene ble gjennomført i forskjellig lokaler hadde liten betydning for endringer i gjennomsnittresultatet for gruppen som helhet. Derimot inneholdt resultatene en større del tilfeldig målefeil. For maksimalkraft var 95% LOA fra $-40N$ til $30N$ for venstre bein og fra $-42N$ til $26N$ for høyre bein, noe som indikerer at 95% av populasjonen er forventet å ligge et sted mellom $42N$ dårligere og $30N$ bedre på den andre testen forutsatt at testprotokollen er lik (figur 13). For målingene av totalkraft for begge bein var referanseintervallene smalere (tabell 8). Dette indikerer at det kan være en fordel å regne ut totalkraften av begge bein og bruke denne i analysene. Det kan tenkes at siden deltakerne holder igjen med begge beina under en eksentrisk kontraksjon med NordBord, kan det ene beinet kompensere for det andre. Dette kan variere fra test til retest og er vanskelig å kontrollere for av testleder.

De to studiene vi sammenlignet oss med ble presentert i tabell 1. Begge gjennomførte eksentriske hamstringsstyrke-målinger på menn og hadde en utvalgsstørrelse som var ganske lik vår, med henholdsvis 30 og 25 deltakere (Opar et al., 2013; Wiesinger et al., 2020). Studiene oppga en ICC fra 0,83 til 0,94 for maksimalkraft. Dette er verdier som er ganske lik våre, som varierte fra 0,89 til 0,92. Like ICC-verdier kan tenkes å skyldes at vårt utvalg sin variasjon (SD) var ganske lik tidligere studier. Opar og medarbeidere (2013) oppga SEM% fra 5,8% til 8,5%. Vi observerte noe lavere SEM% på 4,0% for både venstre og høyre bein (tabell 8). Wiesinger og medarbeidere (2020) oppga absolutt reliabilitet fra metodefeilen

(ME) som beskrevet av (Portney & Watkins, 2009). De fant ME på 5,5% og 6,9% for henholdsvis venstre og høyre bein. For å sammenligne våre resultater opp mot disse verdiene regnet vi ut ME for våre resultater. ME ble regnet ut som SD av differansen mellom test og re-test dividert med $\sqrt{2}$. ME sin prosentvise variasjon ble utregnet som beskrevet av Portney & Watkins (2009).

$$\frac{2ME}{\text{gjennomsnitt av test 1} + \text{gjennomsnitt av test 2}} * 100$$

Vi fant en ME% på 3,7% og 4,0% for henholdsvis venstre og høyre bein (vedlegg 6).

Resultatene indikerer at det var noe mindre variasjoner i testresultatene for vårt utvalg, enn i tidligere studier. En årsak til dette kan skyldes at flere av de inkluderte deltakerne trener eksentrisk hamstringsstyrke regelmessig og derfor er kjent med øvelsen. En annen forklaring kan være knyttet til at de fleste deltakerne i vår studie var kjent med testapparatet som ble benyttet fra tidligere testing på IHS. En tredje forklaring kan muligens relateres til måleapparatene som ble brukt, da de to andre studiene brukte en egenesignet versjon av NordBord. Både Opar og medarbeidere (2013) og Wiesinger og medarbeidere (2020) hadde en tilvennings-økt i forkant av reliabilitetstestingen. Våre resultater viser i midlertid at en tilvennings-økt ikke er nødvendig for en gruppe som trener *Nordic hamstrings* øvelsen regelmessig.

MDC95% fra målingene våre indikerer at det må til en omtrent 11% endring for henholdsvis venstre og høyre bein for å finne en reell forskjell i maksimalkraft på individnivå (tabell 8). MDC95% var noe lavere i vår studie enn i de to studiene vi sammenligner oss med (Opar et al., 2013; Wiesinger et al., 2020). MDC95% er som tidligere nevnt et konservativt estimat for endring av funksjon og blir direkte påvirket av SEM-verdien. Utregningen beskriver en 95% sannsynlighet for at endringen er reell enten i negativ eller positiv retning dersom den overstiger 11% i dette tilfellet. Det videre arbeidet består da av å undersøke om denne verdien er mindre enn framgangen som blir oppnådd i kliniske studier (helst randomiserte kontrollerte studier), for å kunne konkludere med at endringen er reell. En nylig publisert studie av Vianna og medarbeidere (2021) undersøkte et treningsprogram som inkluderte *Nordic hamstrings* øvelsen for 17 profesjonelle kvinnelige fotballspillere gjennom en åtte-ukers intervensjon. De målte deltakernes eksentriske hamstringsstyrke med en prototype av NordBord (E-lastic; E-sporte Soluções Esportivas, Brasília, Brazil). Resultatene viste en framgang på 13% og 14% for deltakerne for henholdsvis venstre og høyre bein fra før til etter intervensjonen. Disse

resultatene indikerer at målinger med NordBord for kvinnelige fotballspillere på høyt nivå kan oppdage endring i funksjonen etter en treningsintervensjon og kan vurderes som reliabel også på individnivå.

5.2.3 Isometrisk adduksjons-, abduksjons-, utadrotasjons og innadrotasjonsstyrke i hoften

Adduksjons-, abduksjons-, utadrotasjons og innadrotasjonsstyrke i hoften ble målt med ForceFrame. Det kommer tydelig fram i Bland-Altman plottene at den tilfeldige målefeilen var langt større enn den systematiske målefeilen for alle testene. Dette gjelder målinger av venstre og høyre bein samt målinger for maksimalkraft og gjennomsnittlig maksimalkraft (figur 13-16). Differansen mellom testdagene var liten for alle testene og varierte fra -1,1N til -4,3N for maksimalkraft (tabell 8). Dette betyr at den systematiske forskjellen mellom testdagene var liten og faktorer som læringseffekt, ulik lokasjon og ulik lengde på testprotokollen mellom test og retest hadde minimal betydning for endring i gjennomsnittsverdien til utvalget mellom testene. På den andre siden var den tilfeldige målefeilen mellom testdagene noe stor sett i forhold til gjennomsnittsverdien. Dette gjelder særlig for måling av innadrotasjonsstyrke, hvor 95% LOA viste intervaller for både venstre og høyre bein som strakk seg over 25N i negativ og positiv retning (figur 14), fra en gjennomsnittsverdi som var fra 131N til 145N (tabell 8).

To tidligere studier har vurdert reliabiliteten til Groinbar (Vald Performance, Albion, Australia) som målemetode (Ryan et al., 2019; Desmyttere et al., 2019). Groinbar er en tidligere versjon av ForceFrame (Vald Performance, Albion, Australia), og slik vi målte hoftestyrke er det ingen forskjell mellom apparatene. Ryan og medarbeidere (2019) undersøkte test-retest reliabiliteten i en testprotokoll samme dag (ikke oppgitt tid mellom test-retest) for hofteadduksjonsstyrke (tabell 2), mens Desmyttere og medarbeidere (2019) undersøkte de samme testene som oss i et større testbatteri med 15 minutter mellom test-retest (tabell 2 og 3). Desmyttere og medarbeidere (2019) sin studie er således den studien som er mest sammenlignbar med vår. Det må likevel påpekes at begge studiene testet test-retest samme dag, mens vi undersøkte uke-til-uke reliabiliteten. Variasjonen innad i individer har en tendens til å variere mer fra uke-til-uke enn samme dag og resultatene er dermed ikke direkte sammenlignbare (O`Donoghue, 2012). Desmyttere og medarbeidere (2019) påpekte også dette og foreslo at framtidige studier burde undersøke uke-til-uke reliabiliteten til GroinBar/ForceFrame som målemetode.

Når det gjelder reliabilitet for adduksjonsstyrke oppga begge studiene ICC som mål på relativ reliabilitet. Ryan og medarbeidere (2019) fant høye ICC-verdier på 0,95 og 0,94 for henholdsvis venstre og høyre bein. Desmyttere og medarbeidere oppga en ICC-verdi på 0,85 for totalkraften av begge bein samlet. Vi fant ICC-verdier på 0,91 og 0,94 for henholdsvis venstre og høyre bein (tabell 8). I likhet med Desmyttere og medarbeidere (2019) ble våre ICC-verdier noe høyere under utregning av gjennomsnittlig maksimalkraft (tabell 9) enn under utregning av maksimalkraft (tabell 8). Som mål for absolutt reliabilitet oppga Ryan og medarbeidere (2019) CV% på 6,7% og 6,3% for henholdsvis venstre og høyre bein, Desmyttere og medarbeidere (2019) oppga SEM% på 8,2% for begge bein samlet. Våre resultater viste noe lavere SEM% på 6,2% og 4,9% for henholdsvis venstre og høyre bein (tabell 8). Dette indikerer noe høyere reliabilitet i vår studie enn tidligere studier, men man må ta forbehold om at Ryan og medarbeidere (2019) utregnet reliabilitet på en annen måte enn oss.

For måling av abduksjonsstyrke oppga Desmyttere og medarbeidere (2019) en SEM% på 6,2%, våre verdier var på 3,3% og 4,2% for henholdsvis venstre og høyre bein (tabell 8). For utadrotasjonsstyrke var deres SEM% på 5,7%, våre verdier var på 6,3% og 4,4% for henholdsvis venstre og høyre bein (tabell 8) og for innadrotasjonsstyrke var deres SEM% på 13,0% våre verdier var på 8,4% og 7,6% (tabell 8). Vi hadde jevnt over noe lavere SEM% enn Desmyttere og medarbeidere (2019). Resultatene fra både vår studie og Desmyttere og medarbeidere (2019) viste at målinger av innadrotasjonsstyrke inneholdt størst grad av målefeil. Årsaken til dette blir videre drøftet i metodediskusjonen.

Ingen tidligere studier har vurdert MDC95% for ForceFrame/Groinbar. Vi fant MDC95%-verdier på 17,2% og 13,1% for adduksjon, 9,2% og 11,7% for abduksjon, 17,4% og 12,1% for utadrotasjon og 23,0% og 21,0% for innadrotasjon for henholdsvis testing av venstre og høyre bein (tabell 8). Noe lavere verdier ble funnet for gjennomsnittlig maksimalkraft (tabell 9). Hva som anses som en akseptabel verdi, er som nevnt i teorien, knyttet til hensikten med studien. I en typisk klinisk setting på individnivå vurderer vi en MDC95% på over 20%, som i dette tilfellet inkluderer målinger for både venstre og høyre bein for innadrotasjon som for stor til å kunne finne en endring i funksjonen. Det kan også argumenteres for at verdiene for adduksjon og utadrotasjon er for stor til å kunne brukes på individnivå. Derimot kan testene anvendes på gruppenivå. Som nevnt i teorien vil reliabilitetsverdier på gruppenivå være mindre avhengig av små individuelle forskjeller enn verdier som blir anvendt på individnivå. For våre resultater vil MDC95% på gruppenivå tilsvare 4,5% og 4,1% for henholdsvis venstre og høyre

bein for testing av maksimalkraft for innadrotasjonsstyrke. Dette viser at testen er brukbar på gruppenivå, men inneholder for mye målefeil til å kunne brukes på individnivå.

Vi testet isolerte hoftestyrketester. En liten endring i kroppsposisjon til deltakerne kan derimot øke styrkeresultatet ved å for eksempel aktivere hoftefleksorene under testing av hofteadduksjonsstyrke. Hoftefleksorene er en sterk muskelgruppe sammenlignet med hofteadduktorene og en aktivering av disse kan potensielt forbedre resultatet til deltakerne (Dahl & Rinvik, 2010). Protokollen vår for vurdering av adduksjonsstyrke ga en nøyaktig beskrivelse av testene, og vi tror at testleder forsikret at denne aktiveringen ikke oppstod under gjennomføringene. På en tilsvarende måte som kroppsposisjonen til deltakerne kan påvirke resultatet, kan plasseringen av sensorene på ForceFrame muligens påvirke resultatet. Dersom kraftsensorene var plassert anterior eller posterior for lateral malleol eller mediale malleol, kan det redusere eller øke deltakerens kraftproduksjon. Selv om protokollen var kjent for tester, kan slike feil oppstå. Testers motivasjon eller tretthet kan også redusere årvåkenhet og videre lede til feil. En feil fra tester kan være en mulig faktor for både systematisk og tilfeldig målefeil (Portney & Watkins, 2009).

5.2.4 Isometrisk hoftefleksjonsstyrke

Forskjellen mellom testdagene var liten for maksimalkraft og varierte fra 1,5N (0,3%) til 9,2N (1,8%) (tabell 8). Dette indikerer at det ikke var noen systematiske forskjeller mellom testdagene. Referanseintervallene for 95% LOA viste store intervaller som strakk seg over 100N i negativ og positiv retning for både venstre og høyre bein, fra en gjennomsnittsverdi på noe over 500N. Dette antyder at det forekom store individuelle forskjeller mellom testdagene for flere av deltakerne. Visuell inspeksjon av plottene i figur 14 viser at de aller fleste observasjonene var innenfor 50N mellom testdagene, men enkelte observasjoner lå over 100N i forskjell mellom testdagene og påvirket derfor LOA. En mulig forklaring på de store individuelle forskjellene mellom testdagene kan skyldes at testen var ny for flere av deltakerne. For å redusere målefeilen kunne vi inkludert flere oppvarmingsforsøk enn de to vi hadde, og eventuelt hatt en lengre pause enn 10 sekunder mellom hver kontraksjon. Det hadde også vært en fordel om vi hadde hatt en tilvenning til testen i forkant, for å gjøre deltakerne kjent med testen. Andre faktorer som dynamometerets plassering og leddvinkel kan også ha påvirket reliabiliteten. Eksempelvis om en deltaker presset rett opp mot dynamometeret eller skrått oppover, kan det påvirke kraften som blir registrert av dynamometeret. Selv om det var en erfaren tester som gjennomførte målingene og vi brukte fiksert belte for at testerens styrke ikke skulle ha innvirkning på resultatet, kan små justeringer av kroppsposisjonen til

deltakerne ha påvirket kraftproduksjonen mellom testdagene (Ishøi et al., 2019). Vi observerte også mye større styrkeverdier i vår studie enn i tidligere studier. Dette har en sammenheng med at vi testet deltakerne med hoften på 0°, mens de to studiene vi sammenligner oss med har målt hoftefleksjonsstyrke med hoften på 90°.

De to studiene vi sammenligner oss med ble presentert i tabell 4 i teoridelen (Ishøi et al., 2019; Thorborg et al., 2010a). Ishøi og medarbeidere (2019) undersøkte test-retest samme dag, med 30 minutter mellom test og retest. Thorborg og medarbeidere (2010a) undersøkte test-retest med en uke mellom testdagene. Studiene presenterte ICC for maksimalkraft på 0,95 og 0,91, noe som er høyere verdier enn vi fant i vår studie (0,75 for venstre bein og 0,8 for høyre bein). Som nevnt tidligere blir ICC påvirket av om utvalget er homogent eller heterogent. Begge de to studiene hadde et mer heterogent utvalg enn vår studie, og de inkluderte både kvinner og menn. Dette kan være med å forklare hvorfor de fant høyere ICC-verdier enn oss. Ishøi og medarbeidere (2019) fant en SEM% på 5,9%, mens Thorborg og medarbeidere (2010a) fant en SEM% på 4,9%. Dette er lavere verdier enn vi fant, med SEM% på 6,9% for venstre bein og 7,7% for høyre bein (tabell 8). Bedre SEM% i tidligere studier kan skyldes at de hadde en større grad av standardisering mellom testene. Standardiseringen til testene blir videre diskutert i metodediskusjonen. Det er heller ikke usannsynlig at metoden vi brukte for å måle hoftefleksjon på 0° er en mindre reliabel metode enn å vurdere hoftefleksjonsstyrke på 90°.

Når det gjelder MDC95% oppga Ishøi og medarbeidere en verdi på 16,4% på individnivå. Vi fant MDC95% på 19,4% for venstre bein og 21,4% for høyre bein under målinger av maksimalkraft (tabell 8). Ishøi og medarbeidere (2019) inkluderte også en utregning av MDC95% basert på utvalgsstørrelsen i studien. De fant da en MDC95% på 4,0% på gruppenivå, noe som indikerer at dersom gruppen som helhet øker eller reduserer styrkeverdiene med 4% blir det regnet som en reell forskjell. De konkluderer med at testing av hoftefleksjonsstyrke med HHD dynamometer er reliabelt på gruppenivå. På den andre siden er testen lite reliabel på individnivå med mindre testen brukes i forbindelse med skader og retur til idrett hvor det ofte blir observert store forskjeller i styrke fra før skade til etter skade (Ishøi et al., 2019). For våre resultater vil dette si at målefeilen vi observerte var for stor til å kunne bruke testen i en klinisk setting på individnivå, men testen kan anses som reliabel på gruppenivå. Med hensyn til MDC95% og 95% LOA er det tydelig at det trengs en endring i testprotokollen for å redusere de individuelle målefeilene. I vår studie testet vi uke-til-uke

reliabiliteten og det er mange faktorer som kan ha påvirket kraftproduksjonen og bistått til tilfeldig målefeil mellom testdagene. Disse faktorene blir videre drøftet i metodediskusjonen.

5.2.5 Oppsummering

Resultatene fra testene viste moderat til høy relativ reliabilitet med ICC_{3,1} verdier fra 0,75 til 0,95 (95% KI: 0,51-0,98). Resultatene fra samtlige tester viste også SEM% verdier under 10%, som er blitt foreslått som et mål på akseptable absolutte reliabilitetsverdier innen kliniske studier (Hopkins, 2000). Bland-Altman plottene for alle målingene viste at målefeilen i denne studien hovedsakelig var knyttet til tilfeldig målefeil mellom testdagene. MDC95% for de fleste testene viste akseptable verdier, hvor en tenkt intervensjon kan tenke seg å gi en større framgang på individnivå enn MDC95%. Det må allikevel påpekes at MDC95% for innadrotasjons- og hoftefleksjonstesten var over 20% for maksimalkraft, noe som indikerer at det trengs store endringer på individnivå for å kunne fastslå en endring av klinisk betydning. De sistnevnte testene er derfor best egnet til å vurdere styrke på gruppenivå, eller brukes i forbindelse med skader og retur til idrett for kvinnelige fotballspillere i Toppserien.

6. Metodediskusjon

Det er flere faktorer ved metoden som kan påvirke reliabiliteten til testene. Herunder testere, utvalget, datainnsamlingen, målemetoder og de statistiske analysene. Disse faktorene vil bli drøftet videre i dette kapittelet, som tar for seg en diskusjon av den valgte metoden.

6.1 Testere

Når test-retest reliabiliteten skal vurderes i en studie er man avhengig av hvor god testerne er til å utføre målingene likt fra gang til gang (Atkinson & Nevill, 1998). Testerne i denne studien bestod av to stipendiater og én mastergradsstudent ved NIH. De hadde ulik grad av erfaring, uten at dette nødvendigvis påvirket kvaliteten på målingene. En erfaren tester er ikke en forutsetning for presise målinger (O'Donoghue, 2012). Det faktum at testerne hadde gjennomført tilnærmet lik testprotokoll med alle Toppserielagene før fotballsesongen 2020 er med å styrke denne studien. Testerne fikk da et innblikk i hvordan protokollen ble gjennomført og de fikk observert vanlige feil deltakerne gjorde. Alle testerne fikk opplæring i apparatene som ble benyttet av personell på IHS og de øvde også på hverandre i forkant av testingen. For å redusere risikoen for at testerne skulle gjøre feil var det samme tester som gjennomførte de samme målingene på både test og retest for alle deltakerne. Samlet sett skulle disse grepene gjøre at målingene ble utført tilnærmet likt fra test til retest. Faktorer som framtoning, humør og motivasjon i interaksjonen mellom tester og deltaker kan likevel ha bistått til målevariasjon.

6.2 Utvalg

Utvalget i denne studien bestod av en homogen gruppe kvinnelige Toppseriespillere når man ser på deskriptive variabler som høyde, vekt og alder (tabell 6). De var en undergruppe fra hovedstudien som kartlegger skader og sykdom for kvinnelige fotballspillere i Norge. Den opprinnelige planen bestod av å rekruttere minimum 50 deltakere, for å gjennomføre test-retest med spillere fra forskjellige klubber i Toppserien. Med et utvalg på 50 deltakere kunne vi gjennomført robuste statistiske analyser i tråd med anbefalingene for reliabilitetsstudier og dermed fått et sikrere estimat på reliabilitet for vårt utvalg (Kottner et al., 2011). Det ville vært enklere å rekruttere flere deltakere til studien dersom vi hadde inkludert kvinnelige fotballspillere på lavere nivå, men på den andre siden ville dette redusert representativiteten til utvalget i hovedstudien. Av lagene i hovedstudien var det seks lag som var basert i Oslo som var aktuelle for å bli inkludert i reliabilitetstesting. To av lagene kunne ikke delta grunnet

Champion League-deltakelse og et lag fikk ikke gjennomført re-testen fordi det ble innført korona-restriksjoner i forkant (figur 3). Vi fikk derfor rekruttert tre lag og burde med det hatt nok deltakere til å overstige antallet på 50. Totalt møtte det opp 56 spillere fra de inkluderte lagene på IHS for testing, men av disse tok kun 28 del i test-retest studien (figur 3). Flere deltakere møtte ikke opp til retest grunnet skade, sykdom, skole, ferie eller at tidspunkt for re-testen ikke passet av andre grunner. Vi hadde bare mulighet til å re-teste spiller fra samme klubb ved én anledning, da vi var avhengig av å frakte utstyr med bil fra IHS til de ulike klubbhusene. Deltakere som bare ble testet én gang ble ekskludert fra analysene. Med et lite utvalg blir reliabiliteten i større grad påvirket av utliggerer. Et lite utvalg kan derfor påvirke reliabiliteten til testene (Carter et al., 2011). I denne studien ville flere testforsøk, testdager eller et større utvalg muligens forbedret reliabiliteten. Årsaken til dette er at desto flere målinger er det en tendens til at målefeilen reduseres (Hopkins, 2000). Alle de statistiske analysene blir også påvirket av utvalgsstørrelsen (ICC, SEM, LOA, MDC).

Eksklusjonskriterier

Eksklusjonskriteriene ble basert på kriterier gitt i tidligere studier (Desmyttere et al., 2019; Wiesinger et al., 2020). Vi la til et ekstra kriterium, baserte på deltakernes subjektive oppfattelse av smerte etter hoftefleksjons- og adduksjonstesten. Vi ekskluderte spillere som oppga at de kjente smerte i og før gjennomføringen av testene fordi dette kunne tenkes å påvirke evnen til å ta i maksimalt. Dette førte til at av de 28 deltakere som ble inkludert i studien var det 21 som gjennomførte test-retest for eksentrisk hamstringstyrke, 26 for innad- og utadrotasjonsstyrke, 25 for adduksjons- og abduksjonsstyrke og 26 for hoftefleksjonsstyrke (figur 3). Når hensikten med studien var å teste skadefrie kvinnelige Toppseriespillere var det nødvendig å ekskludere deltakere med «plager». En alternativ tilnærming til eksklusjon kunne vært at vi kun ekskluderte deltakere som hadde påviste skader og ikke tok del i trening eller kamp. En slik tilnærming ville ført til et større utvalg for de forskjellige testene. Dette kunne vært problematisk for reliabiliteten til testene da vi vet at mange av skadene i kvinnefotball er belastingsrelaterte og ikke fører til fravær fra trening/kamp (Harøy et al., 2017a).

Variasjon i testene

Selv om deltakerne bestod av et homogent utvalg (kvinnelige fotballspillere i Toppserien), var det likevel stor spredning i prestasjon på styrketestene. En årsak til dette kan være at utvalget bestod av både erfarne landslagspillere og juniorspillere som nettopp var tatt opp i A-troppen. På hoftefleksjonstesten varierte resultatene for maksimalkraft fra 325N til 785N for deltakeren med lavest og høyest resultat. Også de andre testene viste stor forskjell mellom deltakere som

presterte best og de som presterte dårligst. Videre var det store individuelle forskjeller i testresultat for noen av deltakerne. For målinger av innadrotasjons- og hoftefleksjonsstyrke var det tre deltakere som hadde over 20% forskjell i maksimalkraft mellom de to testdagene. Det er ikke utenkelig at de spillerne som hadde lite erfaring med styrketester (juniorspillere) var de som hadde den største variasjonen mellom testene, og vi sann sett hadde fått en høyere reliabilitet om vi bare hadde inkludert erfarne Toppseriespillere.

6.3 Datainnsamling

Vi fulgte en delvis standardisert testprotokoll som var utarbeidet før studiestart av personell på IHS. I denne studien ble ikke inter-tester reliabiliteten undersøkt. Undersøkelse av inter-tester reliabilitet ville gitt oss et mål på hvor mye samsvar det var mellom ulike testere under de samme målingene (Carter et al., 2011; O'Donoghue, 2012). Denne prosessen hadde vært tidkrevende for både testerne og deltakerne og hadde krevd at vi gjennomførte flere målinger med det samme utvalget. For testing av adduksjons-, abduksjons-, innadrotasjons-, utadrotasjons- og eksentrisk hamstringsstyrke er risikoen lav for at testenes reliabilitet hadde blitt påvirket av forskjellige testere, da protokollen er enkel og apparatene er stasjonære. For måling av hoftefleksjonsstyrke med HHD derimot, vil resultatene kunne bli påvirket av tester sin styrke. Vi brukte derfor samme tester og fikserte belter for å prøve og eliminere tester sin styrke som mulig kilde til målefeil, noe som er observert i tidligere studier (Thorborg et al., 2013; Kemp et al., 2013).

Forskjeller mellom testdagene

Studien vår skiller seg ut fra andre reliabilitetsstudier ved at test-retest var i forskjellige lokaler. Vi hadde bare mulighet til å teste de inkluderte lagene på IHS ved én anledning, grunnet andre bookinger av testlokale. Re-testen ble derfor gjennomført på lagene sitt klubbhus. Vi fikk da illustrere at måleapparatene er portable og at testene kan standardiseres likt på ulike lokasjoner. Det var også tidsbesparende og praktisk for deltakerne å gjennomføre re-testen på klubbhuset i forbindelse med en treningsøkt fremfor å måtte dra til IHS. Forskjellige lokaler kan være en mulig kilde som bistod til målevariasjon. Den første testen ble gjennomført på IHS, som en del av den standardiserte helseundersøkelsen for kvinnelige fotballspillere. Det er ikke utenkelig at noen av deltakerne presterte bedre i en slik klinisk setting grunnet økt motivasjon (Atkinson & Nevill, 1998; Portney & Watkins, 2009).

Det ble gjennomført flere tester på IHS enn på re-testen. Selv om vi førte opp testrekkefølgen fra den første testdagen, var det noen av deltakerne som hadde gjennomført beinpress-test og

CMJ-test i forkant av testene hvor reproduserbarheten ble undersøkt. Dette er fordi vi fordelte alle deltakerne tilfeldig ut på de fem stasjonene (beskrevet i metoden) for å få økt effektivitet i testingen. CMJ-test og beinpress test ble ikke utført på re-testen da vi ikke hadde som målsetning å undersøke reproduserbarheten til disse testene. Det er derfor ikke utenkelig at noen deltakere presterte dårligere på IHS grunnet økt grad av muskeltretthet. Hovedsakelig vil det være den isotoniske testen av eksentrisk hamstringsstyrke som ble påvirket av flere tester i forkant, ettersom isotoniske tester i større grad blir påvirket av muskeltretthet enn isometriske (Winter, 2007). Tiltak som ble satt inn for å redusere graden av muskeltretthet var minimum 5 min pause mellom de ulike testene.

Optimalt burde vi ha gjennomført den samme testprotokollen og utført testene i samme lokale ved både test og retest, for å sikre like forutsetninger på begge testdagene. Når vi ser på våre resultater (t-tester) var det likevel kun utadrotasjon for høyre bein (tabell 8), som viste en systematisk forskjell i favør den første testdagen. Forskjellige lokaler og ulik testprotokoll virker derfor til å ha hatt liten betydning for systematisk feil i vår studie. På den andre siden kan disse faktorene ha bistått til tilfeldig målefeil mellom testdagene. Som nevnt i teorien er en av svakhetene med t-testen at den ikke tar hensyn til store individuelle variasjonen i testresultatene. I vårt tilfelle kan dette ha ført til at enkelte deltakere presterte bedre på IHS grunnet økt motivasjon, mens andre presterte bedre på re-testen grunnet en kortere testprotokoll. Forskjellige lokaler og ulik testprotokoll kan derfor ha bistått til økt grad av tilfeldig målefeil i vår studie.

Tidsintervall mellom testene

Tidsintervallet ble satt til 7 dager mellom testdagene. Det er en nøye vurdering av hvor lang tid det skal være mellom testdagene for å redusere læringseffekten og muskeltrettheten fra første til andre testdag (Atkinson & Nevill, 1998). Med én uke mellom testene er det ikke sannsynlig at deltakerne har forbedret styrken, men det kan ha forekommet fysiologiske endringer, endring i mentalitet eller forskjellig grad av muskeltretthet for deltakerne (Portney & Watkins, 2009).

Når de inkluderte lagene ble testet var to av de akkurat ferdig med inneværende sesong, mens det tredje laget ble testet etter de var i gang med sesongoppkjøringen til neste sesong. Treningsbelastningen før begge testdagene kan derfor ha vært ulik for de inkluderte deltakerne. Vi hadde ingen aktivitetsbegrensninger mellom eller før testdagene i vår studie. Dette kan være problematisk for reliabiliteten, da vi også vet at to av lagene hadde trening tett

på re-testen. En studie av kvinnelige fotballspillere i Norge av Andersson og medarbeidere (2008) viser at deltakelse i kamp kan føre til akutt tretthet og reduserte styrke de neste timene og dagene etter kamp. Studien rapporterte individuelle forskjeller for deltakerne, men for noen deltakere tok det opp til tre dager før de var restituert og styrken var tilbake på samme nivå som før kamp. Det er derfor ikke utenkelig at hard styrketrening eller fotballtrening i forkant av den første eller andre testdagen kan ha ført til redusert styrke på den påfølgende testdagen. Treningsbelastningen til vårt utvalg kan i høyeste grad ha bistått til både systematisk og tilfeldig målefeil for begge testdagene, hvilket vi ikke kunne kontrollere for.

Standardisering

Begrensninger til denne studien var at den ikke inkluderer oversikt over deltakernes eksponering for fotballtreninger mellom testene. Studien inkluderer heller ikke en oversikt over spesifikk styrketrening eller intervensjoner før og mellom testdagene. Optimalt sett burde vi kontrollert for disse faktorene bedre og standardisert hva deltakerne gjorde i uken før og mellom testene. Deltakerne burde hatt en rolig uke med trening før begge testdagene, for å sikre en lik omstendighet på begge testdagene, men dette var ikke mulig å få til med denne gruppen med toppidrettsutøvere. Det faktum at vi ikke kontrollerte for disse faktorene, kan ha påvirket reliabiliteten til testene i negativ grad. En tilvennings-økt til testene var heller ikke mulig, men de fleste deltakerne i studien hadde gjennomført den samme testprotokollen før fotballsesongen 2020. Alle testene hadde også en øvelsesspesifikk oppvarming for å gjøre deltakerne kjent med testene. Det var likevel noen av deltakerne som ikke hadde testet muskelstyrke på IHS før, så det kan tenkes at de hadde større læringseffekt mellom testdagene enn deltakerne som hadde testet på IHS tidligere.

På den andre siden er det en styrke med denne studien at den reflekterer en «real-world setting», hvilket kan styrke studiens eksterne validitet (Carter et al., 2011). Graden av standardisering til testene er et stort diskusjonstema. Design av studien vår, logistikk og gjennomføring av testingen gjenspeiler «praksis som den er» i en testsituasjon for kvinnelige fotballspillere i hovedstudien. Vi kan beskrive designet som en delvis-standardisert tilnærming (Carter et al., 2011). En streng standardisering ville muligens gitt en kunstig høy reliabilitet og ikke gjenspeile resultatet deltakerne oppnår ved en senere anledning (Carter et al., 2011). Vi kunne gjennomført en mer detaljert utforming av testprotokollen, hvor vi i større grad kontrollerte for faktorer som påvirket resultatet mellom testene. Risikoen for at noen av deltakeren hadde ulik grad av muskeltretthet, forskjellig inntak av mat og drikke, søvnmangel osv., som påvirker reliabiliteten er derfor større i denne studien enn under studier

med et strengere nivå av standardisering. Vi erkjenner også at utvalget er homogent, noe som gjør at resultatene er vanskelig å generalisere til andre grupper enn kvinnelige fotballspillere på høyt nivå.

6.4 Målemetoder

Vi undersøkte uke-til-uke reliabiliteten for seks standardiserte styrketester. Eksentrisk hamstringsstyrke ble målt med NordBord. Adduksjons-, abduksjons-, innadrotasjons-, - og utadrotasjonsstyrke med ForceFrame og hoftefleksjonsstyrke med HHD.

Målemetodene vi brukte ga alle resultater i form av kontinuerlig data målt i Newton. Dette sikrer i utgangspunktet en større spredning av skår enn data på nominal/ordinal nivå. Vi korrigerer ikke styrkevariablene for kroppsvekt, høyde eller beinlengde. Korreksjon for høyde/vekt i vår studie kunne muligens ført til mindre spredning i resultatene for vårt utvalg, dette kunne gitt oss en mer homogen gruppe og videre påvirket størrelsen av ICC.

Eksentrisk hamstringsstyrke

Testing av eksentrisk hamstringsstyrke ble gjennomført likt i vår studie som de vi sammenligner oss med tanke på utføring av testen illustrert i figur 7 (Wiesinger et al., 2020; Opar et al., 2013). I vår studie var det ingen pause mellom hver av de tre kontraksjonene slik som i Opar og medarbeidere (2013) sin studie. Wiesinger og medarbeidere (2020) hadde en pause mellom repetisjonene på 1 min. Mangel på pause kan ha ført til at deltakerne ikke fikk restituert seg mellom forsøkene. Ifølge Fleck (1983) tar det opp til tre min før adenosintrifosfat og kreatinfosfat regenereres etter et maksimalt-muskelarbeid. Om dette har ført til målevariasjonen i vår studie er usikkert, siden det ikke var pause mellom de ulike kontraksjonene på begge testdagene. For fremtidige studier kan det derimot være en fordel å ha en lengre pause mellom testforsøkene for å sikre at deltakerne er restituert til hver kontraksjon. En annen faktor som kan bidra til målevariasjon var hvilken hastighet deltakerne hadde på det eksentriske hamstringsarbeidet. Wiesinger og medarbeidere (2020) er den eneste studien som har undersøkt ulike hastigheter for eksentrisk hamstringstesting med NordBord. De fant like verdier for testing av maksimalkraft, hvor deltakerne lente seg så sakte som mulig framover på NordBord og når deltakerne lente seg framover med en hastighet på 30° i sekundet. Andre hastigheter er ikke undersøkt og det kan tenkes at det er en forskjell i kraftproduksjon etter hvor hurtig en deltaker gjennomfører bevegelsen. Vi mener uansett at testleder passet på at bevegelsen ble gjennomført som beskrevet i protokollen med tregest mulig hastighet, og derfor sørget for at hastigheten på bevegelsen ikke hadde innvirkning på

resultatet i studien. En annen faktor som er viktig å nevne er kraftproduksjonen for spillere som kontrollerer bevegelsen hele veien. Flere av deltakerne i vår studie klarte å kontrollere bevegelsen gjennom hele bevelsesbanen (hovedsakelig lette spillere), og noen klarte å holde kroppen horisontalt med bakken i flere sekunder før de slapp seg ned. Når hensikten med reliabilitetsstudien var å måle maksimalkraft-produksjon i hamstring, er vi usikker på om NordBord klarer å fange opp maksimalkraft-produksjonen på samme måte som IKD for spillere som kontrollerer hele bevegelsesbanen. Et forslag til framtidige studier kan være å bruke vektvest for deltakere som kontrollerer hele bevegelsesbanen, slik at de har mulighet til å gjøre et maksimalt voluntært arbeid. Uansett må vektvest undersøkes på både spillere som kontrollerer hele bevegelsesbanen og spillere som ikke kontrollerer hele bevegelsesbanen. Dette må gjøres for å se om det vil påvirke resultatet til begge gruppene.

Adduksjons- og abduksjonsstyrke

For måling av adduksjons- og abduksjonsstyrke med ForceFrame var vår metode noe annerledes enn i tidligere studier. Tidligere studier har målt adduksjons- og abduksjonsstyrke med kraftsensorene plassert på mediale og laterale femor condyles, og med kneleddet på 60° (Ryan et al., 2019) og kneleddet på 45° (Desmyttere et al., 2019). Vi målte styrke med strake bein, hvor kraftsensorene ble plassert på mediale malleol og laterale malleol, illustrert i figur 8 og 9. Dette gjorde at vi unngikk å måle opp vinkler i hoften og kneleddet til deltakerne, og testene kunne derfor bli gjennomført mer effektivt og var enklere å standardisere. I tillegg har en studie gjennomført på 20 fotballspillere med HHD vist at testing med strake hofter hadde bedre reliabilitet enn testing med hoftene på 45°, under måling av adduksjonsstyrke (Light & Thorborg, 2016).

Innad- og utadrotasjonsstyrke

For måling av innad- og utadrotasjonsstyrke gjennomførte vi testene annerledes enn det som er gjort i den tidligere studien (Desmyttere et al., 2019). Deltakerne lå i vår studie på magen, illustrert i figur 10 og 11. Med den valgte metoden oppdaget testleder at deltakerne ofte bommet på kraftsensorene eller enten beveget overkropp eller bein for å treffe sensorene på ForceFrame. Flere av deltakerne ga også tilbakemelding at testen føltes unaturlig og at de hadde en subjektiv følelse av å ikke klare å presse så mye de kunne. Dette kan være med på å forklare hvorfor testing av innadrotasjonsstyrke hadde stor målevariasjon. For andre studier som ønsker å undersøke innad- og utadrotasjonsstyrke med bruk av ForceFrame som målemetode, virker det mer gunstig å gjennomføre testene ryggliggende, som beskrevet av Desmyttere og medarbeidere (2019).

Hoftefleksjonsstyrke

For å måle hoftefleksjonsstyrke brukte vi et HHD. Metoden vi brukte skiller seg ut fra andre studier som tester hoftefleksjonsstyrke (Thorborg et al., 2010a; Ishøi et al., 2019). Deltakerne ble plassert på en undersøkelsesbenk med hoften på 0°, illustrert i figur 12. I tidligere studier har deltakerne blitt plassert på kanten av en undersøkelsesbenk og gjennomført hoftefleksjon sittende med hoftene i 90°. Vi gjennomførte hoftefleksjonstesten likt som tidligere studier når vi testet Toppserie-spillerne før fotballsesongen 2020. Protokollen ble endret til etter sesongen fordi metoden sittende var vanskelig å standardisere og overkroppen til deltakerne ikke var fiksert under testingen. Deltakerne vippet ofte med overkroppen for å skape kraft, hvilket var vanskelig å kontrollere for av testleder. Med bruk av belter som festet deltakerne til en undersøkelsesbenk, sikret vi bedre at andre muskelgrupper ikke bisto til å skape kraft. Det kan også argumenteres for at å måle hoftefleksjonsstyrke med hoftene på 0° framfor 90° er en mer fotballspesifikk øvelse, ettersom fotballspillere skaper kraft når hoften er rundt 0°, eksempelvis når de sparker ball eller under akselerasjonsfasen i løp (Kloskowska et al., 2016).

Type muskelarbeid

Forskjellen mellom isotonisk, isokinetisk og isometrisk testing ble beskrevet i teorien. Hvilke tester som velges er hovedsakelig knyttet til hva en ønsker å undersøke, for eksempel om man vurderer skaderisiko eller undersøker tester som er relatert til prestasjon i idretten (Winter, 2007). De fleste testene som ble brukt i denne studien var isometriske (hoftefleksjon, adduksjon, abduksjon, innadrotasjon og utadrotasjon). Isometrisk testing kan være en fordel for å redusere fysisk påkjenning for deltakerne og kan anvendes for spillere med en kjent patologi (Thorborg et al., 2010a). Disse testene kan brukes i forbindelse med retur til idrett for en spiller som har vært skadet eller gi et bilde av en spillers styrke og svakheter (Desmyttere et al., 2019). Testene gir også informasjon om sideforskjeller, og forholdet mellom agonister og antagonister. På den andre siden er disse testene mindre idrettsspesifikke, og kan ikke nødvendigvis relateres direkte til idrettsprestasjon. Det er en vesentlig forskjell mellom den spesifikke situasjonen man tester muskelstyrken i isometrisk, og de funksjonelle krav som stilles i idretten eller aktiviteten spilleren skal tilbake til (Winter, 2007). Styrkeverdier på de isometriske testene i denne reliabilitetsstudien har derfor liten betydning for prestasjon i fotball.

Bilateral og unilateral testing

I et idrettsmiljø med lite tid til klinisk testing er det viktig at testingen foregår effektivt. Fem av testene i denne studien ble målt bilateral og testprotokollen blir da kortere. Det har tidligere

blitt illustrert at maksimale voluntære kontraksjoner målt bilateralt viser lavere totalsum sammenlignet med å teste et bein om gangen (Škarabot et al., 2016). Dette ble også observert av O'Brien og medarbeidere (2018) når de sammenlignet HHD med ForceFrame under vurdering av adduksjons- og abduksjonsstyrke. Deltakerne i studien oppnådde bedre resultater med HHD. Forfatterne spekulerer i om dette skyldes at deltakerne ikke klarte å produsere maksimal voluntær kraft bilateralt, men forskjellen kan også, ifølge dem, skyldes en liten forskjell i testposisjon (O'Brien et al., 2018). Forskjellen mellom unilaterale og bilaterale målinger blir i teorien omtalt som bilateralt underskudd, og kan skyldes målemetoders konfigurering eller justeringer av deltakernes kropp (Simoneau et al., 2015). Med bakgrunn i dette er ikke resultater og reliabilitetsverdier som ble oppnådd bilateralt i denne studien direkte overførbare til unilaterale målinger.

6.5 Statistiske analyser

I reliabilitetsstudien brukte vi parametriske analyser for alle testene. Vi brukte en uavhengig t-test for å studere forskjeller i styrke mellom utvalget og øvrige spillere. ICC ble brukt for å uttrykke relativ reliabilitet. For absolutt reliabilitet ble SEM/SEM% og LOA estimert. MDC/MDC% ble estimert som mål for minste reelle endring som har en klinisk betydning.

Parametriske analyser

Vi valgte å gjøre parametriske analyser for alle variablene, selv om det var noen av variablene som ikke var normalfordelte. Med et lite utvalg er ikke alltid alle variablene normalfordelte, selv om en styrkevariabel vanligvis er normalfordelt. Sentralgrenseteoremet beskriver en teori innenfor sannsynlighet. Den sier at fordelingen av et gjennomsnitt for et stort antall uavhengige variabler vanligvis er normalfordelt, uavhengig av den underliggende fordelingen. Dette vil si at fordelingen av uavhengige variabler vil nærme seg normalfordeling når utvalgsstørrelsen øker (O'Donoghue, 2012). Vi begrunner valget av de statistiske analysene som ble gjennomført på denne teorien. Dersom vi hadde hatt et større utvalg hadde sannsynligvis alle styrkevariablene vært normalfordelte. Parametriske analyser er foretrukket innenfor statistikken fordi de gir mer omfattende analyser enn ikke-parametriske tester. Grunnen til dette er blant annet at parametriske analyser gir flere estimater, er basert på gjennomsnittet til deltakerne og kan presenteres med konfidensintervall (Altman & Bland, 2009). Det er derfor en fordel å anvende parametriske analyser.

Gjenspeiler utvalget populasjonen?

Siden formålet med studien var å undersøke reliabiliteten til testene for kvinnelige Toppseriespillere, er det en fordel å vite om utvalget i vår studie (n=28), gjenspeiler store

deler av populasjonen ($n=74$). For å undersøke om de inkluderte deltakerne i reliabilitetsstudien hadde tilsvarende styrke som hele populasjonen, ble maksimalkraft fra alle testene analysert med en uavhengig t-test. Den første testen til vårt utvalg ble da vurdert mot øvrige spillere. En vurdering av gjennomsnittet til alle styrkeverdiene for maksimalkraft viste at øvrige spillere var sterkere i abduksjon enn utvalget ($P < 0,05$) (tabell 7), mens det for de øvrige styrketestene ikke var signifikante forskjeller mellom gruppene. Med bakgrunn i de statistiske analysene vil vi si at utvalget som helhet gjenspeilet store deler av populasjonen på en god måte.

6.5.1 Relativ reliabilitet

Vi undersøkte relativ reliabilitet med bruk av $ICC_{3,1}$ for maksimalkraft for alle testene (toveis ANOVA, blandet effektmodell; enighet, enkel skår) og modell $ICC_{3,k}$ for gjennomsnittlig maksimalkraft fra alle testforsøkene (toveis ANOVA, blandet effektmodell, enighet, gjennomsnittsskår). Begrunnelse for valg av ICC-modell ble tatt på bakgrunn av teoretisk anbefaling (Weir, 2005; Koo & Li, 2016; Shrout & Fleiss, 1979). Som nevnt i teorien brukes ICC-modell 3 vanligvis for test-retest analyser og er et objektivt mål på reliabilitet uavhengig av utvalgsstørrelsen (Hopkins, 2000). Med bruk av enighet kan ICC også ta høyde for systematisk og tilfeldig målevariasjon. Dette anbefales ved undersøkelse av reliabilitet i test-retest studier (Koo & Li, 2016; Shrout & Fleiss, 1979). Ettersom formålet med studien var å undersøke reliabiliteten til testene for vårt utvalg, kan det forsvares at vi brukte modell 3. Skulle resultatene blitt generalisert til andre grupper, burde modell 2 blitt anvendt (Koo & Li, 2016). Generelt vil modell 2 gi lavere ICC-verdier enn modell 3, særlig om resultatene inneholder systematiske forskjeller. For våre resultater ville modell 2 og 3 gitt ganske like verdier siden dataene inneholdt liten grad av systematisk forskjell mellom testdagene. Enighet vil også generelt gi lavere ICC-verdier enn konsistens ved repeterte målinger (Koo & Li, 2016).

Av de seks studiene vi sammenligner oss med baserte to studier seg på ICC modell 3 (Desmyttere et al., 2019; Wiesinger et al., 2020), mens to andre studier baserte seg på modell 2 (Thorborg et al., 2010a; Ishøi et al., 2019). De siste to oppga ikke modell (Opar et al., 2013; Ryan et al., 2019). Dette illustrerer noe av utfordringene når man skal sammenligne på tvers av studier. Med bruk av ulike måter å utregne ICC på reduseres sammenligningsgrunnlaget og overførbarhetsverdien mellom studiene (Weir, 2005; Atkinson & Nevill, 1998).

6.5.2 Absolutt reliabilitet

Vi brukte SEM og LOA for å estimere den absolutte reliabiliteten til testene. I teorien er statistiskere uenige i hvilken metode som egner seg best for å undersøke den absolutte reliabiliteten ved repeterte målinger, særlig når man har et lite utvalg som i vår studie (Hopkins, 2000; Atkinson & Nevill, 1998; Atkinson & Nevill, 2000). Hopkins (2000) argumenterte for at LOA er for strengt, særlig for idrettsutøvere hvor små endringer kan ha stor betydning for prestasjon. Vi brukte derfor begge estimatene i vår studie. Fordelen med LOA er at den representerer 95% referanseintervall for test-retest variasjon, og den gir et mål for både systematisk og tilfeldig målefeil. SEM representerer 68% referanseintervall av den sanne skåren sin feil (Atkinson & Nevill, 2000). Vi utregnet SEM som $SEM = SD * (\sqrt{1-ICC})$. Ved å benytte denne kalkuleringen vil SEM være påvirket av heterogeniteten av målingene og ha en begrensning når en skal sammenligne med andre studier. SEM vil da bare være representativt for et utvalg med tilsvarende variabilitet som brukt i denne studien (Weir, 2005). Med en utregning av SEM fra feilkomponenten fra en ANOVA ($\sqrt{\text{mean error variance}}$), vil SEM være uavhengig fra ICC-verdien, og minimere heterogeniteten sin påvirkning av resultatet (Weir, 2005). Av studiene vi sammenlignet oss med regnet tre ut SEM fra $SEM = SD * (\sqrt{1-ICC})$ (Thorborg et al., 2010a; Ishøi et al., 2019; Desmyttere et al., 2019). Ingen av studiene regnet ut SEM fra feilkomponenten fra en ANOVA. Vi baserte valget av utregning av SEM på hva tidligere studier på området har gjort. Vi tok hensyn i måten å utregne SEM ved å rekruttere deltakere til reliabilitetsstudien som var representativ for hovedstudies målgruppe (skadefrie kvinnelige fotballspillere i Toppserien).

I studiene vi sammenlignet oss med var det én som oppga LOA-verdier (Desmyttere et al., 2019). LOA blir framstilt grafisk i et Bland-Altman plott for å kunne studere de individuelle deltakernes variasjon over repeterte målinger. Vi mener derfor at en slik framstilling av resultatene er mer oversiktlig enn kun å få oppgitt en SEM-verdi. I tillegg ble MDC/MDC% regnet ut for alle testene. Som tidligere nevnt er MDC et konservativt estimat for å evaluere endring i funksjon og beskriver hvor mye endring som må til i deltakernes styrkeverdier for at en med 95% sannsynlighet kan si at det har skjedd en reell endring (Weir, 2005). Da den reelle forskjellen må være større enn målefeilen blir MDC regnet ut med å legge til $\sqrt{2}$ i ligningen (Portney & Watkins, 2009). $MDC = SEM * 1,96 * \sqrt{2}$. Av studiene vi sammenlignet oss med var det fire som oppga MDC95% verdi (Thorborg et al., 2010a; Ishøi et al., 2019; Opar et al., 2013; Wiesinger et al., 2020) og en som oppga MDC95% på gruppenivå utregnet som $MDC95\% / \sqrt{n}$ (Ishøi et al., 2019).

6.6 Oppsummering metodediskusjon

Metoden vi brukte i denne studien bestod av en delvis standardisert testprotokoll, hvor vi undersøkte test-retest reliabiliteten av seks standardiserte styrketester for kvinnelige Toppseriespillere i fotball uten skader. Testerne hadde god opplæring i bruk av apparater som ble benyttet. Begrensninger med studien består av liten utvalgsstørrelse, at vi ikke kontrollerte treningsbelastningen i forkant og mellom testene, utvalgets homogenitet som gjør at resultatene kun kan generaliseres til den valgte populasjonen og at vi ikke undersøkte inter-tester reliabiliteten. En styrke med denne studien var at testene ble gjennomført med samme grad av standardisering som testing i hovedstudien. Dette øker den eksterne validiteten. Statistiske analyser er også en styrke med studien hvor vi undersøkte ICC som mål på relativ reliabilitet, SEM og LOA for absolutt reliabilitet og MDC for å undersøke hvor mye endring i testresultatet som kreves for en klinisk betydning.

6.7 Praktiske implikasjoner

I denne studien har vi undersøkt seks standardiserte styrketester som blir brukt i den standardiserte helseundersøkelsen for kvinnelige Toppseriespillere. Dette har gitt oss en mulighet til å studere testprosedyrene i hovedstudien. Ved å undersøke test-retest reliabiliteten kan vi presentere i hvilken grad målingene er presise/reliable. I det større bildet vil denne studien vil gi et mål på kvaliteten av målingene i hovedstudien.

Reliabiliteten fra de ulike testene varierte, hvor teting av eksentrisk hamstringsstyrke og abduksjonsstyrke var testene som inneholdt minst grad av variasjon. Våre resultater gir en indikasjon på at testing av hoftefleksjon og innadrotasjon var de minst reliable testene. For testing av hoftefleksjon ble det brukt et fiksert belte for å eliminere testers styrke som mulig kilde til målefeil. Det var likevel tre deltakere som hadde en målevariasjon på over 20% mellom test-retest, noe som kan sies å være uakseptabelt for klinisk testing. Det er mulig at plassering av dynamometerhodet har hatt innvirkning på resultatet til deltakerne under testing av hoftefleksjonsstyrke. Dersom andre studier skal bruke den samme metoden som i vår studie, anbefaler vi å inkludere flere testforsøk og ha lengre pause enn 10 sekunder mellom hver kontraksjon for å sikre at deltakerne er uthvilt. Det kan også være en fordel å inkludere flere oppvarmingsforsøk enn de to vi hadde for å sikre at deltakerne blir kjent med testen.

Som tidligere nevnt var det flere av deltakerne som ga uttrykk for at måten vi gjennomførte testing av innad- og utadrotasjonsstyrke var vanskelig å utføre. Vi kunne gjennomført flere tilvenningsforsøk eller endret protokollen slik at deltakerne ble testet ryggliggende, som

beskrevet av Desmyttere og medarbeidere (2019). Derimot kan ingen ukritisk modifisering av protokollen som ble brukt i hovedstudien bli iverksatt. Modifiseringer av protokollen i hovedstudien kan påvirke de resultatene som allerede er samlet inn å redusere sammenlignbarhets-grunnlaget mellom testene. En mulighet er å sammenligne testene før modifisering og etter modifisering for å undersøke om de viser tilsvarende målinger.

Gjennomsnittlig maksimalkraftverdier viste generelt høyere reliabilitet enn maksimalkraftverdier i vår studie. Desmyttere og medarbeidere (2019) fant også bedre reliabilitet for målinger av gjennomsnittlig maksimalkraft enn for maksimalkraft under målinger med ForceFrame. De konkluderte med at man derfor burde bruke gjennomsnittete av flere målinger under vurdering av muskelstyrke for fotballspillere. Jeg stiller meg usikker til om dette er den beste tilnærmingen ettersom maksimalkraft gir et bedre mål på styrke enn gjennomsnittlig maksimalkraft. Det kan tenkes at deltakerne ikke gir maksimalt på alle testforsøkene, særlig når vi inkluderte en protokoll med 5-30 sekunders pause mellom hver kontraksjon som i vår studie. Selv om reliabiliteten ble bedre for målinger av gjennomsnittlig maksimalkraft, mener jeg at reliabilitetsverdiene som blir presentert for maksimalkraft-målingene gir et bedre mål på reliabiliteten enn målinger av gjennomsnittlig maksimalkraft.

7. Konklusjon

I denne studien undersøkte vi test-retest reliabiliteten for seks standardiserte styrketester av kvinnelige Toppseriespillere. Målefeilene for samtlige tester skyldes hovedsakelig tilfeldig målefeil. Testene viste generelt moderat til høy relativ reliabilitet (ICC) og samtlige tester viste akseptable SEM%-verdier $\leq 10\%$. MDC95% for de fleste testene viste akseptable verdier, hvor en tenkt intervensjon kan tenke seg å gi en større framgang på individnivå enn MDC95%. Det må allikevel påpekes at MDC95% for innadrotasjons- og hoftefleksjonstesten var over 20% for maksimalkraft, noe som indikerer at det trengs store endringer på individnivå for å kunne fastslå en endring av klinisk betydning. De sistnevnte testene er derfor best egnet til å vurdere styrke på gruppenivå, eller brukes i forbindelse med skader og retur til idrett for kvinnelige fotballspillere i Toppserien.

Standardisering av testene gjenspeilet praksis som den er under testing på IHS, så resultatene fra denne reliabilitetsstudien kan relateres til en reell testsituasjon for kvinnelige fotballspillere i Toppserien.

På grunn av metodiske svakheter med studien bør resultatene bli tolket med forsiktighet og kun anvendes for utvalg med tilsvarende variabilitet som i denne studien.

Referanseliste

- Al Attar, W. S. A., Soomro, N., Sinclair, P. J., Pappas, E., & Sanders, R. H. (2017). Effect of Injury Prevention Programs that Include the Nordic Hamstring Exercise on Hamstring Injury Rates in Soccer Players: A Systematic Review and Meta-Analysis. *Sports Med*, 47(5), 907-916. <https://doi.org/10.1007/s40279-016-0638-2>
- Altman, D. G., & Bland, J. M. (2009). Parametric v non-parametric methods for data analysis. *Bmj*, 338, a3167. <https://doi.org/10.1136/bmj.a3167>
- Andersson, H., Raastad, T., Nilsson, J., Paulsen, G., Garthe, I., & Kadi, F. (2008). Neuromuscular fatigue and recovery in elite female soccer: effects of active recovery. *Med Sci Sports Exerc*, 40(2), 372-380. <https://doi.org/10.1249/mss.0b013e31815b8497>
- Arnason, A., Andersen, T., Holme, I., Engebretsen, L., & Bahr, R. (2008). Prevention of hamstring strains in elite soccer: An intervention study. *Scandinavian journal of medicine & science in sports*, 18, 40-48. <https://doi.org/10.1111/j.1600-0838.2006.00634.x>
- Arnason, A., Sigurdsson, S. B., Gudmundsson, A., Holme, I., Engebretsen, L., & Bahr, R. (2004). Risk factors for injuries in football. *Am J Sports Med*, 32(1 Suppl), 5s-16s. <https://doi.org/10.1177/0363546503258912>
- Atkinson, G., & Nevill, A. (2000). Typical error versus limits of agreement. *Sports Med*, 30(5), 375-381. <https://doi.org/10.2165/00007256-200030050-00005>
- 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. <https://doi.org/10.2165/00007256-199826040-00002>
- Bahr, R. (2009). No injuries, but plenty of pain? On the methodology for recording overuse symptoms in sports. *British Journal of Sports Medicine*, 43(13), 966. <https://doi.org/10.1136/bjism.2009.066936>

- Bahr, R. (2016). Why screening tests to predict injury do not work-and probably never will...: a critical review. *Br J Sports Med*, *50*(13), 776-780. <https://doi.org/10.1136/bjsports-2016-096256>
- Bahr, R., Clarsen, B., Derman, W., Dvorak, J., Emery, C. A., Finch, C. F., Hägg, M., Junge, A., Kemp, S., Khan, K. M., Marshall, S. W., Meeuwisse, W., Mountjoy, M., Orchard, J. W., Pluim, B., Quarrie, K. L., Reider, B., Schwellnus, M., Soligard, T., Stokes, K. A., Timpka, T., Verhagen, E., Bindra, A., Budgett, R., Engebretsen, L., Erdener, U., & Chamari, K. (2020). International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sport 2020 (including STROBE Extension for Sport Injury and Illness Surveillance (STROBE-SIIS)). *British Journal of Sports Medicine*, *54*(7), 372-389. <https://doi.org/10.1136/bjsports-2019-101969>
- Bahr, R., & Krosshaug, T. (2005). Understanding injury mechanisms: a key component of preventing injuries in sport. *British Journal of Sports Medicine*, *39*(6), 324-329. <https://doi.org/10.1136/bjism.2005.018341>
- Bakken, A., & Norges idrettshøgskole. (2018). *The benefits of periodic health evaluation in professional football : a focus on musculoskeletal screening*. Norwegian School of Sport Sciences
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307-310.
- Boling, M., Padua, D., Marshall, S., Guskiewicz, K., Pyne, S., & Beutler, A. (2010). Gender differences in the incidence and prevalence of patellofemoral pain syndrome. *Scand J Med Sci Sports*, *20*(5), 725-730. <https://doi.org/10.1111/j.1600-0838.2009.00996.x>
- Brehm, M.-A., Scholtes, V., Dallmeijer, A., Twisk, J., & Harlaar, J. (2011). The importance of addressing heteroscedasticity in the reliability analysis of ratio-scaled variables: An example based on walking energy-cost measurements. *Developmental medicine and child neurology*, *54*, 267-273. <https://doi.org/10.1111/j.1469-8749.2011.04164.x>
- Buchheit, M., Cholley, Y., Nagel, M., & Poulos, N. (2016). The Effect of Body Mass on Eccentric Knee-Flexor Strength Assessed With an Instrumented Nordic Hamstring Device (Nordbord) in Football Players. *Int J Sports Physiol Perform*, *11*(6), 721-726. <https://doi.org/10.1123/ijsp.2015-0513>

- Carter, R. E., Lubinsky, J., Domholdt, E., & Domholdt, E. (2011). *Rehabilitation research : principles and applications* (4th ed.). Elsevier Saunders.
- Charlton, P. C., Mentiplay, B. F., Grimaldi, A., Pua, Y.-H., & Clark, R. A. (2017). The reliability of a maximal isometric hip strength and simultaneous surface EMG screening protocol in elite, junior rugby league athletes. *Journal of Science and Medicine in Sport*, 20(2), 139-145.
<https://doi.org/https://doi.org/10.1016/j.jsams.2016.06.008>
- Croisier, J.-L., Ganteaume, S., Binet, J., Genty, M., & Ferret, J.-M. (2008). Strength Imbalances and Prevention of Hamstring Injury in Professional Soccer Players: A Prospective Study. *The American Journal of Sports Medicine*, 36(8), 1469-1475.
<https://doi.org/10.1177/0363546508316764>
- Cross, K. M., Gurka, K. K., Saliba, S., Conaway, M., & Hertel, J. (2013, Apr). Comparison of hamstring strain injury rates between male and female intercollegiate soccer athletes. *Am J Sports Med*, 41(4), 742-748. <https://doi.org/10.1177/0363546513475342>
- Crossley, K. M., Patterson, B. E., Culvenor, A. G., Bruder, A. M., Mosler, A. B., & Mentiplay, B. F. (2020, Sep). Making football safer for women: a systematic review and meta-analysis of injury prevention programmes in 11 773 female football (soccer) players. *Br J Sports Med*, 54(18), 1089-1098. <https://doi.org/10.1136/bjsports-2019-101587>
- Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports Med*, 38(4), 297-316. <https://doi.org/10.2165/00007256-200838040-00003>
- Dahl, H. A., & Rinvik, E. (2010). *Menneskets funksjonelle anatomi : med hovedvekt på bevegelsesapparatet* (3. utg. ed.). Cappelen akademisk.
- Datson, N., Hulton, A., Andersson, H., Lewis, T., Weston, M., Drust, B., & Gregson, W. (2014, 2014/09/01). Applied Physiology of Female Soccer: An Update. *Sports Medicine*, 44(9), 1225-1240. <https://doi.org/10.1007/s40279-014-0199-1>
- De Vet, H., Bouter, L., Bezemer, P., & Beurskens, A. (2001). Reproducibility and responsiveness of evaluative outcome measures. *International Journal of Technology Assessment in Health Care*, 17, 479-487.

- Desmyttere, G., Gaudet, S., & Begon, M. (2019). Test-retest reliability of a hip strength assessment system in varsity soccer players. *Phys Ther Sport*, 37, 138-143. <https://doi.org/10.1016/j.ptsp.2019.03.013>
- Dick, R. W. (2009). Is there a gender difference in concussion incidence and outcomes? *Br J Sports Med*, 43 Suppl 1, i46-50. <https://doi.org/10.1136/bjism.2009.058172>
- Drew, M. K., & Finch, C. F. (2016). The Relationship Between Training Load and Injury, Illness and Soreness: A Systematic and Literature Review. *Sports Med*, 46(6), 861-883. <https://doi.org/10.1007/s40279-015-0459-8>
- Dvorak, J., Grimm, K., Schmied, C., & Junge, A. (2009). Development and implementation of a standardized precompetition medical assessment of international elite football players--2006 FIFA World Cup Germany. *Clin J Sport Med*, 19(4), 316-321. <https://doi.org/10.1097/JSM.0b013e3181b21b6e>
- Ekstrand, J. (2013). Keeping your top players on the pitch: the key to football medicine at a professional level. *British Journal of Sports Medicine*, 47(12), 723-724. <https://doi.org/10.1136/bjsports-2013-092771>
- Ekstrand, J., Hägglund, M., Kristenson, K., Magnusson, H., & Waldén, M. (2013). Fewer ligament injuries but no preventive effect on muscle injuries and severe injuries: an 11-year follow-up of the UEFA Champions League injury study. *British Journal of Sports Medicine*, 47(12), 732-737. <https://doi.org/10.1136/bjsports-2013-092394>
- Ekstrand, J., Hägglund, M., & Waldén, M. (2011). Injury incidence and injury patterns in professional football: the UEFA injury study. *Br J Sports Med*, 45(7), 553-558. <https://doi.org/10.1136/bjism.2009.060582>
- Engebretsen, A. H., Myklebust, G., Holme, I., Engebretsen, L., & Bahr, R. (2010). Intrinsic Risk Factors for Groin Injuries among Male Soccer Players: A Prospective Cohort Study. *The American Journal of Sports Medicine*, 38(10), 2051-2057. <https://doi.org/10.1177/0363546510375544>
- Espinosa, G., Pöyhönen, T., Aramendi, J., Samaniego, J., Emparanza, J., & Kyröläinen, H. (2015). Effects of an eccentric training programme on hamstring strain injuries in women football players. *Biomedical Human Kinetics*, 7. <https://doi.org/10.1515/bhk-2015-0019>

- Faude, O., Junge, A., Kindermann, W., & Dvorak, J. (2005). Injuries in female soccer players: a prospective study in the German national league. *Am J Sports Med*, *33*(11), 1694-1700. <https://doi.org/10.1177/0363546505275011>
- Feeley, B. T., Kennelly, S., Barnes, R. P., Muller, M. S., Kelly, B. T., Rodeo, S. A., & Warren, R. F. (2008). Epidemiology of National Football League training camp injuries from 1998 to 2007. *Am J Sports Med*, *36*(8), 1597-1603. <https://doi.org/10.1177/0363546508316021>
- Feiring, D. C., Ellenbecker, T. S., & Derscheid, G. L. (1990). Test-retest reliability of the biodex isokinetic dynamometer. *J Orthop Sports Phys Ther*, *11*(7), 298-300. <https://doi.org/10.2519/jospt.1990.11.7.298>
- Fleck, S. (1983). Interval: Physiological basis. *Strength & Conditioning Journal*, *5*(5), 40-40. https://journals.lww.com/nsca-scj/Fulltext/1983/10000/Interval__Physiological_basis.8.aspx
- Fousekis, K., Tsepis, E., Poulmedis, P., Athanasopoulos, S., & Vagenas, G. (2011). Intrinsic risk factors of non-contact quadriceps and hamstring strains in soccer: a prospective study of 100 professional players. *British Journal of Sports Medicine*, *45*(9), 709-714. <https://doi.org/10.1136/bjism.2010.077560>
- Fuller, C. W., Ojelade, E. O., & Taylor, A. (2007). Preparticipation medical evaluation in professional sport in the UK: theory or practice? *British Journal of Sports Medicine*, *41*(12), 890-896. <https://doi.org/10.1136/bjism.2007.038935>
- Gaulrapp, H., Becker, A., Walther, M., & Hess, H. (2010). Injuries in women's soccer: a 1-year all players prospective field study of the women's Bundesliga (German premier league). *Clin J Sport Med*, *20*(4), 264-271. <https://doi.org/10.1097/JSM.0b013e3181e78e33>
- Harøy, J., Clarsen, B., Thorborg, K., Hölmich, P., Bahr, R., & Andersen, T. E. (2017a). Groin Problems in Male Soccer Players Are More Common Than Previously Reported. *Am J Sports Med*, *45*(6), 1304-1308. <https://doi.org/10.1177/0363546516687539>
- Harøy, J., Clarsen, B., Wiger, E. G., Øyen, M. G., Serner, A., Thorborg, K., Hölmich, P., Andersen, T. E., & Bahr, R. (2019). The Adductor Strengthening Programme prevents groin problems among male football players: a cluster-randomised controlled trial.

British Journal of Sports Medicine, 53(3), 150-157. <https://doi.org/10.1136/bjsports-2017-098937>

Harøy, J., Thorborg, K., Serner, A., Bjørkheim, A., Rolstad, L. E., Hölmich, P., Bahr, R., & Andersen, T. E. (2017b). Including the Copenhagen Adduction Exercise in the FIFA 11+ Provides Missing Eccentric Hip Adduction Strength Effect in Male Soccer Players: A Randomized Controlled Trial. *The American Journal of Sports Medicine*, 45(13), 3052-3059. <https://doi.org/10.1177/0363546517720194>

Hopkins, W. (2000). Measures of Reliability in Sports Medicine and Science. *Sports medicine (Auckland, N.Z.)*, 30, 1-15. <https://doi.org/10.2165/00007256-200030010-00001>

Häggglund, M., Waldén, M., & Ekstrand, J. (2009). Injuries among male and female elite football players. *Scand J Med Sci Sports*, 19(6), 819-827. <https://doi.org/10.1111/j.1600-0838.2008.00861.x>

Häggglund, M., Waldén, M., Magnusson, H., Kristenson, K., Bengtsson, H., & Ekstrand, J. (2013). Injuries affect team performance negatively in professional football: An 11-year follow-up of the UEFA Champions League injury study. *British Journal of Sports Medicine*, 47. <https://doi.org/10.1136/bjsports-2013-092215>

Impellizzeri, F. M., Bizzini, M., Rampinini, E., Cereda, F., & Maffiuletti, N. A. (2008). Reliability of isokinetic strength imbalance ratios measured using the Cybex NORM dynamometer. *Clinical Physiology and Functional Imaging*, 28(2), 113-119. <https://doi.org/10.1111/j.1475-097X.2007.00786.x>

The International Olympic Committee (IOC) consensus statement on periodic health evaluation of elite athletes: March 2009. (2009, Sep-Oct). *Journal of athletic training*, 44(5), 538-557. <https://doi.org/10.4085/1062-6050-44.5.538>

Ishøi, L., Hölmich, P., & Thorborg, K. (2019). MEASURES OF HIP MUSCLE STRENGTH AND RATE OF FORCE DEVELOPMENT USING A FIXATED HANDHELD DYNAMOMETER: INTRA-TESTER INTRA-DAY RELIABILITY OF A CLINICAL SET-UP. *International journal of sports physical therapy*, 14(5), 715-723. <https://pubmed.ncbi.nlm.nih.gov/31598409>

- Jacobson, I., & Tegner, Y. (2007). Injuries among Swedish female elite football players: a prospective population study. *Scand J Med Sci Sports*, 17(1), 84-91.
<https://doi.org/10.1111/j.1600-0838.2006.00524.x>
- Jensen, J., Hölmich, P., Bandholm, T., Zebis, M. K., Andersen, L. L., & Thorborg, K. (2014). Eccentric strengthening effect of hip-adductor training with elastic bands in soccer players: a randomised controlled trial. *Br J Sports Med*, 48(4), 332-338.
<https://doi.org/10.1136/bjsports-2012-091095>
- Kemp, J. L., Schache, A. G., Makdissi, M., Sims, K. J., & Crossley, K. M. (2013). Greater understanding of normal hip physical function may guide clinicians in providing targeted rehabilitation programmes. *Journal of Science and Medicine in Sport*, 16(4), 292-296. <https://doi.org/https://doi.org/10.1016/j.jsams.2012.11.887>
- Kloskowska, P., Morrissey, D., Small, C., Malliaras, P., & Barton, C. (2016). Movement Patterns and Muscular Function Before and After Onset of Sports-Related Groin Pain: A Systematic Review with Meta-analysis. *Sports medicine (Auckland, N.Z.)*, 46(12), 1847-1867. <https://doi.org/10.1007/s40279-016-0523-z>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., Roberts, C., Shoukri, M., & Streiner, D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Journal of Clinical Epidemiology*, 64(1), 96-106. <https://doi.org/https://doi.org/10.1016/j.jclinepi.2010.03.002>
- Langhout, R., Weir, A., Litjes, W., Gozeling, M., Stubbe, J. H., Kerkhoffs, G., & Tak, I. (2019). Hip and groin injury is the most common non-time-loss injury in female amateur football. *Knee Surg Sports Traumatol Arthrosc*, 27(10), 3133-3141.
<https://doi.org/10.1007/s00167-018-4996-1>
- Larruskain, J., Lekue, J. A., Diaz, N., Odriozola, A., & Gil, S. M. (2018). A comparison of injuries in elite male and female football players: A five-season prospective study. *Scandinavian journal of medicine & science in sports*, 28(1), 237-245.
<https://doi.org/10.1111/sms.12860>

- Lee, M. J., Reid, S. L., Elliott, B. C., & Lloyd, D. G. (2009). Running biomechanics and lower limb strength associated with prior hamstring injury. *Med Sci Sports Exerc*, *41*(10), 1942-1951. <https://doi.org/10.1249/MSS.0b013e3181a55200>
- Light, N., & Thorborg, K. (2016). The precision and torque production of common hip adductor squeeze tests used in elite football. *Journal of Science and Medicine in Sport*, *19*(11), 888-892. <https://doi.org/https://doi.org/10.1016/j.jsams.2015.12.009>
- Maniar, N., Shield, A. J., Williams, M. D., Timmins, R. G., & Opar, D. A. (2016). Hamstring strength and flexibility after hamstring strain injury: a systematic review and meta-analysis. *British Journal of Sports Medicine*, *50*(15), 909-920. <https://doi.org/10.1136/bjsports-2015-095311>
- McCleary, R. W., & Andersen, J. C. (1992). Test-retest reliability of reciprocal isokinetic knee extension and flexion peak torque measurements. *Journal of athletic training*, *27*(4), 362-365.
- McNulty, K. L., Elliott-Sale, K. J., Dolan, E., Swinton, P. A., Ansdell, P., Goodall, S., Thomas, K., & Hicks, K. M. (2020). The Effects of Menstrual Cycle Phase on Exercise Performance in Eumenorrhic Women: A Systematic Review and Meta-Analysis. *Sports Medicine*, *50*(10), 1813-1827. <https://doi.org/10.1007/s40279-020-01319-3>
- Meeuwisse, W. H. (1994). Assessing Causation in Sport Injury: A Multifactorial Model. *Clinical Journal of Sport Medicine*, *4*(3). https://journals.lww.com/cjsportsmed/Fulltext/1994/07000/Assessing_Causation_in_Sport_Injury__A.4.aspx
- Mjølsnes, R., Arnason, A., Østhaugen, T., Raastad, T., & Bahr, R. (2004). A 10-week randomized trial comparing eccentric vs. concentric hamstring strength training in well-trained soccer players. *Scand J Med Sci Sports*, *14*(5), 311-317. <https://doi.org/10.1046/j.1600-0838.2003.367.x>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*, *63*(7), 737-745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>

- Montalvo, A. M., Schneider, D. K., Silva, P. L., Yut, L., Webster, K. E., Riley, M. A., Kiefer, A. W., Doherty-Restrepo, J. L., & Myer, G. D. (2019). 'What's my risk of sustaining an ACL injury while playing football (soccer)?' A systematic review with meta-analysis. *Br J Sports Med*, 53(21), 1333-1340. <https://doi.org/10.1136/bjsports-2016-097261>
- Mosler, A. B., Weir, A., Eirale, C., Farooq, A., Thorborg, K., Whiteley, R. J., Hölmich, P., & Crossley, K. M. (2018). Epidemiology of time loss groin injuries in a men's professional football league: a 2-year prospective study of 17 clubs and 606 players. *Br J Sports Med*, 52(5), 292-297. <https://doi.org/10.1136/bjsports-2016-097277>
- Murphy, D. F., Connolly, D. A., & Beynonn, B. D. (2003). Risk factors for lower extremity injury: a review of the literature. *Br J Sports Med*, 37(1), 13-29. <https://doi.org/10.1136/bjism.37.1.13>
- NFF. (2018). Årsrapport 2018. Hentet 15 juni 2020 fra <https://www.fotball.no/tema/nff-nyheter/2019/her-er-nffs-arsrapport-2018/>
- Nicholas, S. J., & Tyler, T. F. (2002). Adductor muscle strains in sport. *Sports Med*, 32(5), 339-344. <https://doi.org/10.2165/00007256-200232050-00005>
- Nilstad, A., Andersen, T. E., Bahr, R., Holme, I., & Steffen, K. (2014). Risk Factors for Lower Extremity Injuries in Elite Female Soccer Players. *The American Journal of Sports Medicine*, 42(4), 940-948. <https://doi.org/10.1177/0363546513518741>
- O'Brien, M., Bourne, M., Heerey, J., Timmins, R., & Pizzari, T. (2018). A novel device to assess hip strength: Concurrent validity and normative values in male athletes. *Physical Therapy in Sport*, 35. <https://doi.org/10.1016/j.ptsp.2018.11.006>
- O'Donoghue, P. (2012). *Statistics for Sport and Exercise Studies: An Introduction*. Florence: Routledge. <https://doi.org/10.4324/9780203133507>
- Opar, D. A., Drezner, J., Shield, A., Williams, M., Webner, D., Sennett, B., Kapur, R., Cohen, M., Ulager, J., Cafengiu, A., & Cronholm, P. F. (2014). Acute hamstring strain injury in track-and-field athletes: A 3-year observational study at the Penn Relay Carnival. *Scand J Med Sci Sports*, 24(4), e254-259. <https://doi.org/10.1111/sms.12159>
- Opar, D. A., Piatkowski, T., Williams, M. D., & Shield, A. J. (2013). A novel device using the Nordic hamstring exercise to assess eccentric knee flexor strength: a reliability and

- retrospective injury study. *J Orthop Sports Phys Ther*, 43(9), 636-640.
<https://doi.org/10.2519/jospt.2013.4837>
- Opar, D. A., Williams, M. D., Timmins, R. G., Hickey, J., Duhig, S. J., & Shield, A. J. (2015). Eccentric hamstring strength and hamstring injury risk in Australian footballers. *Med Sci Sports Exerc*, 47(4), 857-865.
<https://doi.org/10.1249/mss.0000000000000465>
- Orchard, J., Marsden, J., Lord, S., & Garlick, D. (1997). Preseason hamstring muscle weakness associated with hamstring muscle injury in Australian footballers. *Am J Sports Med*, 25(1), 81-85. <https://doi.org/10.1177/036354659702500116>
- Orchard, J. W. (2015). Men at higher risk of groin injuries in elite team sports: a systematic review. *British Journal of Sports Medicine*, 49(12), 798-802.
<https://doi.org/10.1136/bjsports-2014-094272>
- Petersen, J., Thorborg, K., Nielsen, M. B., Budtz-Jørgensen, E., & Hölmich, P. (2011). Preventive effect of eccentric training on acute hamstring injuries in men's soccer: a cluster-randomized controlled trial. *Am J Sports Med*, 39(11), 2296-2303.
<https://doi.org/10.1177/0363546511419277>
- Pincivero, D. M., Lephart, S. M., & Karunakara, R. A. (1997). Reliability and precision of isokinetic strength and muscular endurance for the quadriceps and hamstrings. *Int J Sports Med*, 18(2), 113-117. <https://doi.org/10.1055/s-2007-972605>
- Pollard, C. W., Opar, D. A., Williams, M. D., Bourne, M. N., & Timmins, R. G. (2019). Razor hamstring curl and Nordic hamstring exercise architectural adaptations: Impact of exercise selection and intensity. *Scandinavian journal of medicine & science in sports*, 29(5), 706-715. <https://doi.org/10.1111/sms.13381>
- Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research : applications to practice* (3rd ed. ed.). Pearson Prentice Hall.
- Presland, J. D., Timmins, R. G., Bourne, M. N., Williams, M. D., & Opar, D. A. (2018). The effect of Nordic hamstring exercise training volume on biceps femoris long head architectural adaptation. *Scand J Med Sci Sports*, 28(7), 1775-1783.
<https://doi.org/10.1111/sms.13085>

- Raastad, T., Paulsen, G., Refsnes, P. E., Rønnestad, B. R., & Wisnes, A. R. (2010). *Styrketrening : i teori og praksis*. Gyldendal undervisning.
- Rankin, G., & Stokes, M. (1998). Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation*, *12*(3), 187-199. <https://doi.org/10.1191/026921598672178340>
- Ryan, S., Kempton, T., Pacecca, E., & Coutts, A. J. (2019). Measurement Properties of an Adductor Strength-Assessment System in Professional Australian Footballers. *Int J Sports Physiol Perform*, *14*(2), 256-259. <https://doi.org/10.1123/ijsp.2018-0264>
- Schache, A. G., Woodley, S. J., Schilders, E., Orchard, J. W., & Crossley, K. M. (2017). Anatomical and morphological characteristics may explain why groin pain is more common in male than female athletes. *British Journal of Sports Medicine*, *51*(7), 554-555. <https://doi.org/10.1136/bjsports-2016-096945>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*, *86*(2), 420-428. <https://doi.org/10.1037//0033-2909.86.2.420>
- Simoneau, E., Leteneur, S., Toumi, A., Dessurne, A., Gabrielli, F., Barbier, F., & Jakobi, J. (2015). Bilateral Strength Deficit Is Not Neural in Origin; Rather Due to Dynamometer Mechanical Configuration. *PloS one*, *10*, e0145077. <https://doi.org/10.1371/journal.pone.0145077>
- Škarabot, J., Cronin, N., Strojnik, V., & Avela, J. (2016). Bilateral deficit in maximal force production. *Eur J Appl Physiol*, *116*(11-12), 2057-2084. <https://doi.org/10.1007/s00421-016-3458-z>
- Soligard, T., Myklebust, G., Steffen, K., Holme, I., Silvers, H., Bizzini, M., Junge, A., Dvorak, J., Bahr, R., & Andersen, T. E. (2008). Comprehensive warm-up programme to prevent injuries in young female footballers: cluster randomised controlled trial. *Bmj*, *337*, a2469. <https://doi.org/10.1136/bmj.a2469>
- Stark, T., Walker, B., Phillips, J. K., Fejer, R., & Beck, R. (2011). Hand-held Dynamometry Correlation With the Gold Standard Isokinetic Dynamometry: A Systematic Review. *PM&R*, *3*(5), 472-479. <https://doi.org/https://doi.org/10.1016/j.pmrj.2010.10.025>
- Tegnander, A., Olsen, O. E., Moholdt, T. T., Engebretsen, L., & Bahr, R. (2008). Injuries in Norwegian female elite soccer: a prospective one-season cohort study. *Knee Surg*

Sports Traumatol Arthrosc, 16(2), 194-198. <https://doi.org/10.1007/s00167-007-0403-z>

Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2015). *Research methods in physical activity* (7th ed. ed.). Human Kinetics.

Thorborg, K., Bandholm, T., Schick, M., Jensen, J., & Hölmich, P. (2013). Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength. *Scandinavian journal of medicine & science in sports*, 23(4), 487-493. <https://doi.org/10.1111/j.1600-0838.2011.01405.x>

Thorborg, K., Hölmich, P., Christensen, R., Petersen, J., & Roos, E. M. (2011). The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *British Journal of Sports Medicine*, 45(6), 478-491. <https://doi.org/10.1136/bjism.2010.080937>

Thorborg, K., Petersen, J., Magnusson, S. P., & Hölmich, P. (2010a). Clinical assessment of hip strength using a hand-held dynamometer is reliable. *Scandinavian journal of medicine & science in sports*, 20(3), 493-501. <https://doi.org/10.1111/j.1600-0838.2009.00958.x>

Thorborg, K., Serner, A., Petersen, J., Madsen, T. M., Magnusson, P., & Hölmich, P. (2010b). Hip Adduction and Abduction Strength Profiles in Elite Soccer Players: Implications for Clinical Evaluation of Hip Adductor Muscle Recovery After Injury. *The American Journal of Sports Medicine*, 39(1), 121-126. <https://doi.org/10.1177/0363546510378081>

Timmins, R. G., Bourne, M. N., Shield, A. J., Williams, M. D., Lorenzen, C., & Opar, D. A. (2016). Short biceps femoris fascicles and eccentric knee flexor weakness increase the risk of hamstring injury in elite football (soccer): a prospective cohort study. *British Journal of Sports Medicine*, 50(24), 1524. <https://doi.org/10.1136/bjsports-2015-095362>

Tyler, T. F., Silvers, H. J., Gerhardt, M. B., & Nicholas, S. J. (2010). Groin injuries in sports medicine. *Sports health*, 2(3), 231-236. <https://doi.org/10.1177/1941738110366820>

van der Horst, N., Smits, D. W., Petersen, J., Goedhart, E. A., & Backx, F. J. (2015). The preventive effect of the nordic hamstring exercise on hamstring injuries in amateur

- soccer players: a randomized controlled trial. *Am J Sports Med*, 43(6), 1316-1323.
<https://doi.org/10.1177/0363546515574057>
- van Dyk, N., Bahr, R., Burnett, A. F., Whiteley, R., Bakken, A., Mosler, A., Farooq, A., & Witvrouw, E. (2017). A comprehensive strength testing protocol offers no clinical value in predicting risk of hamstring injury: a prospective cohort study of 413 professional football players. *Br J Sports Med*, 51(23), 1695-1702.
<https://doi.org/10.1136/bjsports-2017-097754>
- van Dyk, N., Bahr, R., Whiteley, R., Tol, J. L., Kumar, B. D., Hamilton, B., Farooq, A., & Witvrouw, E. (2016). Hamstring and Quadriceps Isokinetic Strength Deficits Are Weak Risk Factors for Hamstring Strain Injuries: A 4-Year Cohort Study. *The American Journal of Sports Medicine*, 44(7), 1789-1795.
<https://doi.org/10.1177/0363546516632526>
- van Dyk, N., Behan, F. P., & Whiteley, R. (2019). Including the Nordic hamstring exercise in injury prevention programmes halves the rate of hamstring injuries: a systematic review and meta-analysis of 8459 athletes. *British Journal of Sports Medicine*, 53(21), 1362-1370. <https://doi.org/10.1136/bjsports-2018-100045>
- van Dyk, N., Witvrouw, E., & Bahr, R. (2018). Interseason variability in isokinetic strength and poor correlation with Nordic hamstring eccentric strength in football players. *Scandinavian journal of medicine & science in sports*, 28(8), 1878-1887.
<https://doi.org/https://doi.org/10.1111/sms.13201>
- van Mechelen, W., Hlobil, H., & Kemper, H. C. G. (1992). Incidence, Severity, Aetiology and Prevention of Sports Injuries. *Sports Medicine*, 14(2), 82-99.
<https://doi.org/10.2165/00007256-199214020-00002>
- UEFA. (2015). WOMEN'S FOOTBALL ACROSS THE NATIONAL ASSOCIATIONS 2014-15. Hentet 15 juni 2020 fra
<https://www.uefa.com/MultimediaFiles/Download/uefaorg/WFprogramme/02/20/39/67/2>
- Vianna, K. B., Rodrigues, L. G., Oliveira, N. T., Ribeiro-Alvares, J. B., & Baroni, B. M. (2021). A Preseason Training Program With the Nordic Hamstring Exercise Increases Eccentric Knee Flexor Strength and Fascicle Length in Professional Female Soccer Players. *International journal of sports physical therapy*, 16(2), 459-467.
<https://doi.org/10.26603/001c.19452>

- Waldén, M., Hägglund, M., & Ekstrand, J. (2015). The epidemiology of groin injury in senior football: a systematic review of prospective studies. *British Journal of Sports Medicine*, 49(12), 792-797. <https://doi.org/10.1136/bjsports-2015-094705>
- Weir, A., Brukner, P., Delahunt, E., Ekstrand, J., Griffin, D., Khan, K. M., Lovell, G., Meyers, W. C., Muschaweck, U., Orchard, J., Paajanen, H., Philippon, M., Reboul, G., Robinson, P., Schache, A. G., Schilders, E., Serner, A., Silvers, H., Thorborg, K., Tyler, T., Verrall, G., de Vos, R.-J., Vuckovic, Z., & Hölmich, P. (2015). Doha agreement meeting on terminology and definitions in groin pain in athletes. *British Journal of Sports Medicine*, 49(12), 768-774. <https://doi.org/10.1136/bjsports-2015-094869>
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, 19(1), 231-240. <https://doi.org/10.1519/15184.1>
- Werner, J., Hägglund, M., Ekstrand, J., & Waldén, M. (2019). Hip and groin time-loss injuries decreased slightly but injury burden remained constant in men's professional football: the 15-year prospective UEFA Elite Club Injury Study. *British Journal of Sports Medicine*, 53(9), 539-546. <https://doi.org/10.1136/bjsports-2017-097796>
- Whiteley, R., Jacobsen, P., Prior, S., Skazalski, C., Otten, R., & Johnson, A. (2012). Correlation of isokinetic and novel hand-held dynamometry measures of knee flexion and extension strength testing. *J Sci Med Sport*, 15(5), 444-450. <https://doi.org/10.1016/j.jsams.2012.01.003>
- Wiesinger, H.-P., Gressenbauer, C., Kösters, A., Scharinger, M., & Müller, E. (2020). Device and method matter: A critical evaluation of eccentric hamstring muscle strength assessments. *Scandinavian journal of medicine & science in sports*, 30(2), 217-226. <https://doi.org/10.1111/sms.13569>
- Winter, E. M., British Association of, S., & Exercise, S. (2007). *Sport and exercise physiology testing guidelines: the British Association of Sport and Exercise Sciences guide: Volume 2, : Exercise and Clinical Testing* (Vol. Volume 2). Routledge.
- Wollin, M., Thorborg, K., Welvaert, M., & Pizzari, T. (2018). In-season monitoring of hip and groin strength, health and function in elite youth soccer: Implementing an early detection and management strategy over two consecutive seasons. *J Sci Med Sport*, 21(10), 988-993. <https://doi.org/10.1016/j.jsams.2018.03.004>

Tabelloversikt

Tabell 1: Reliabilitet for eksentrisk hamstringstyrke.....	s.33
Tabell 2: Reliabilitet for isometrisk adduksjons- og abduksjonsstyrke med ForceFrame/Groinbar.....	s.35
Tabell 3: Reliabilitet for isometrisk utad- og innadrotasjonsstyrke med ForceFrame/Groinbar.....	s.36
Tabell 4: Reliabilitet for isometrisk hoftefleksjonsstyrke med HHD.....	s.37
Tabell 5: Oversikt over testbatteriet.....	s.41
Tabell 6: Deskriptive verdier for utvalget (n=28) og øvrige spillere (n=74) ved første test. Presentert med gjennomsnitt og standardavvik.....	s.50
Tabell 7: Forskjell fra styrketestene mellom utvalget og øvrige spillere. Presentert med gjennomsnitt og standardavvik (SD).....	s.51
Tabell 8: Test-retest reliabilitet for maksimalkraft (N).....	s.52
Tabell 9: Test-retest reliabilitet for gjennomsnittlig maksimalkraft.....	s.54
Tabell 10: Test-retest reliabilitet ME og CV _{me}	s.108

Figuroversikt

Figur 1: Lokasjon på de definerte områdene for lyskesmerter.....	s.13
Figur 2: Modell som beskriver årsakssammenhenger for skader.....	s.15
Figur 3: Flytdiagram for rekruttering.....	s.39
Figur 4: Tidslinje over datainnsamlingen.....	s.40
Figur 5: Illustrerer de brukte måleenhetene i denne studien.....	s.41
Figur 6: Utdrag fra Vald Hub etter 3 repetisjoner med NH.....	s.42
Figur 7: Illustrerer eksentrisk hamstringstest med NordBord (Vald Performance, Albion, Australia).....	s.43
Figur 8: Illustrere isometrisk hofteadduksjonstest med ForceFrame (Vald Performance, Albion, Australia).....	s.44
Figur 9: Illustrere isometrisk hofteabduksjonstest med ForceFrame (Vald Performance, Albion, Australia).....	s.44
Figur 10: Illustrere isometrisk utadrotasjonstest med ForceFrame (Vald Performance, Albion, Australia).....	s.45
Figur 11: Illustrerer innadrotasjonstest med ForceFrame (Vald Performance, Albion, Australia).....	s.45
Figur 12: Illustrere isometrisk hoftefleksjonstest med dynamometer (Hoggan Health Industries Inc. West Draper, UT, USA).....	s.46
Figur 13: Bland-Altman plott for maksimalkraft (N) for eksentrisk hamstrings-, adduksjons-, og abduksjonsstyrke.....	s.56
Figur 14: Bland-Altman plott for maksimalkraft (N) for utadrotasjons-, innadrotasjons- og hoftefleksjonsstyrke.....	s.57
Figur 15: Bland-Altman plott for gjennomsnittlig maksimalkraft (N) for eksentrisk hamstrings-, adduksjons-, og abduksjonsstyrke.....	s.58
Figur 16: Bland-Altman plott for gjennomsnittlig maksimalkraft (N) for utadrotasjons-, innadrotasjons-, og hoftefleksjonsstyrke.....	s.59

Vedlegg

Vedlegg 1: Godkjenning av lokal etisk komité

Vedlegg 2: Godkjenning av Norsk senter for forskningsdata

Vedlegg 3: Informert samtykke

Vedlegg 4: Godkjenning av bruk av figur 2

Vedlegg 5: Godkjenning bruk av figur 1

Vedlegg 6: Metodefeil og prosentvis variasjon fra testene

Vedlegg 1: Godkjenning av lokal etisk komité



Thor Einar Andersen
Seksjon for idrettsmedisin

OSLO 06. januar 2020

Endringsmelding 129-051219 - 86-131218 – Sammenhengen mellom treningsbelastning, skader og fysisk prestasjonsevne i norsk elite kvinnefotball

Vi viser til endringsmelding med vedlegg mottatt 20.12.19.

I henhold til retningslinjer for behandling av søknad til etisk komite for idrettsvitenskapelig forskning på mennesker, har leder av komiteen på fullmakt konkludert med følgende:

Vedtak

På bakgrunn av forelagte dokumentasjon finner komiteen at endringene er forsvarlig og at det kan gjennomføres innenfor rammene av anerkjente etiske forskningsetiske normer nedfelt i NIHs retningslinjer. Til vedtaket har komiteen lagt følgende forutsetning til grunn:

- *Vilkår fra NSD følges*

Komiteen gjør oppmerksom på at vedtaket er avgrenset i tråd med fremlagte dokumentasjon.

Dersom det gjøres vesentlige endringer i prosjektet som kan ha betydning for deltakernes helse og sikkerhet, skal dette legges fram for komiteen før eventuelle endringer kan iverksettes.

Med vennlig hilsen

Professor Sigmund Loland
Leder, Etisk komite, Norges idrettshøgskole

Sammenheng mellom treningsbelastning, skader og fysisk prestasjonsevne i norsk elite kvinnefotball

Referanse

662612

Status

Vurdert

Åpne Meldeskjema

Vurdering

N

NSD Personvern

31.01.2020 15:33

Det innsendte meldeskjemaet med referansekode 662612 er nå vurdert av NSD.

Følgende vurdering er gitt:

NSD har vurdert endringen registrert 16.12.2019.

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 31.01.2020. Behandlingen kan fortsette.

OPPFØLGING AV PROSJEKTET

NSD vil følge opp underveis (hvert annet år) og ved planlagt avslutning for å avklare om behandlingen av personopplysningene pågår i tråd med den behandlingen som er dokumentert.

Lykke til med prosjektet!

Kontaktperson hos NSD: Belinda Gloppen Helle

Tlf. Personverntjenester: 55 58 21 17 (tast 1)

Vedlegg 3: Informert samtykke

Vil du delta i forskningsprosjektet

«ReadyToPlay:

Protecting the health of Norwegian elite football players»

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å *beskrive forekomsten av skader og sykdom i Toppserien, og undersøke risikofaktorer for skader*. I dette skrevet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Kvinnefotball er i rask utvikling, og nivået og kravene som stilles på trening og i kamp er høyere enn noen gang. Dette kan påvirke risikoen for skader og sykdom, noe som er viktig å kartlegge siden det vil påvirke prestasjon og utvikling for både lag og spiller. Informasjon om faktorer som gjør at spillere har økt risiko for skader er viktig for å kunne forebygge skader, men dette er lite kartlagt i kvinnefotball. Hensikten med denne studien er derfor å *kartlegge alle helseproblemer i Toppserien og undersøke risikofaktorer for skader*. Dette vil være med å danne grunnlaget for hvordan vi kan forebygge skader og bedre prestasjon i fremtiden.

Prosjektet er del av flere doktorgradsprosjekter og involverer etablerte forskere og medisinerere innen fotball. Anonymiserte resultater fra studien vil bli presentert på nasjonale og internasjonale konferanser, brukt i undervisningsformål, inkludert i trenerutdanningen.

Hvem er ansvarlig for forskningsprosjektet?

Norges idrettshøgskole og Senter for idrettsskedeforskning er ansvarlig for prosjektet. *Norges fotballforbund og Toppfotball Kvinner* er også med som samarbeidspartnere for prosjektet.

Hvorfor får du spørsmål om å delta?

Vi kontakter deg med denne forespørselen fordi ditt lag har sagt seg villig til å delta i prosjektet.

Vi ønsker å kartlegge samtlige lag og spillere i Toppserien, derfor får du som spiller på et toppserielag forespørselen om å delta.

Hva innebærer det for deg å delta?

Metoden som brukes i prosjektet er en prospektiv kohortstudie, dette innebærer at vi ønsker å følge en spesifikk gruppe over tid, i dette tilfellet alle spillerne i Toppserien. Du vil trene som normalt med ditt lag hele sesongen, men vi vil samle data om sykdom og skader du blir utsatt for, samt varigheten av både trening og kamp du deltar i.

Hvis du velger å delta i prosjektet;

- Vil du en gang i uken få påminnelse om å rapportere sykdom/skader og treningsmengde via mobilappen «AthleteMonitoring». Her må du svare på et kort spørreskjema, «OSTRC Questionnaire on Health Problems», og registrere treningsmengden for uken som har gått. Dette tar fra 30 sekunder til 4 minutter å svare på, avhengig av om du har hatt skade/sykdom eller ikke.
- Ditt lags fysioterapeut vil varsles umiddelbart om du rapporterer noe nytt, for å raskt kunne undersøke deg og sette i gang tiltak. Fysioterapeuten vil registre hvilken skade/sykdom som har oppstått og hvor mange dager du er borte fra trening/kamp.

- Toppfotball Kvinner gjennomfører i samarbeid med lagene i Toppserien testing av fysisk prestasjonsevne ved Idrettens Helsecenter. Her testes muskelstyrken i beina i tillegg til prestasjonstester i spenst, agility og hurtighet. Du vil også svare på et spørreskjema hvor andre potensielle risikofaktorer for skader blir undersøkt. Vi vil lagre data fra disse testene og bruke resultatene til å se etter sammenhenger med skader.
- Data fra dette prosjektet vil også kunne bli knyttet opp mot prosjekter som undersøker trenings- og kampbelastning i Toppserien, samt TV-opptak av skader i kamp.

Prosjektet vil starte etter at laget ditt har gjennomført testing på Idrettens Helsecenter (i februar/mars) og vare hele sesongen.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke ditt samtykke tilbake, uten å måtte oppgi noen grunn. Alle opplysninger om deg vil da bli anonymisert. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Det vil ikke få noen konsekvenser for deg eller dit lag dersom du ønsker å trekke deg i fra studien.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrevet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- Alle som får innsyn i dine data vil ha taushetsplikt. Kun forskere som deltar i prosjektgruppen vil ha tilgang til dine data. I tillegg vil klubbens fysioterapeut og lege ha innsyn i dine data.
- Når dine data benyttes til forskningsformål, vil de anonymiseres ved at navn og personnummer fjernes. Dataene vil bli behandlet konfidensielt.
- Applikasjonen som brukes heter «Athlete monitoring» og er utviklet av et kanadisk selskap ved samme navn. Applikasjonen er godkjent etter de nye personvernreglene, GDPR.

Alle resultater som omtales i publikasjonene etter prosjektet vil være anonymiserte og det vil ikke være mulig å gjenkjenne deg i resultatene som publiseres.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal etter planen avsluttes 31.12.2029. Alle opplysninger som kan knytte deg til materialet vil bli anonymisert og opplysninger vi har lagret om deg vil slettes.

Alle data om skader og fysisk prestasjonsevne som hentes ut for forskningsformål vil bli lagret, i anonymisert form, i en database for å kunne kartlegge hvordan omfang og utvikling endrer seg i Toppserien over tid. Materialet vil være viktig kunnskap for å forstå hvordan vi skal arbeide med forebygging av skader og sykdom, samt tilrettelegging av belastning med tanke på forebygging og utvikling av fysisk prestasjonsevne. Dataene vil kunne danne et viktig grunnlag for utarbeidelse av blant annet arbeidskrav i Toppserien.

Styret ved Norges idrettshøgskole har bestemt at forskningsdata skal lagres i fem år etter prosjektslutt for etterprøvnbarhet og kontroll. Dette innebærer at alle data, utenom personopplysninger, vil bli lagret i sin helhet i fem år hos Norges idrettshøgskole. Dette er meldt til Norsk senter for forskningsdata (NSD).

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra *Norges idrettshøgskole* har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- *Norges idrettshøgskole* ved Solveig Thorarinsdottir, solveig.thorarinsdottir@nih.no, tlf. 405 22 930 eller Roar Amundsen, roar.amundsen@nih.no, tlf. 482 97 832.
- Vårt personvernombud: Tove Riise, *Norges Idrettshøgskole*, personvernombud@nih.no
- NSD – Norsk senter for forskningsdata AS, personverntjenester@nsd.no eller tlf. 555 82 117.

Med vennlig hilsen

Professor dr. med.
Roald Bahr
(Veileder)

Solveig Thorarinsdottir
(Stipendiat)

Roar Amundsen
(Stipendiat)

Samtykkeerklæring

Dersom du ønsker å delta i forskningsprosjektet vil du kunne gi ditt samtykke elektronisk ved å godkjenne informasjonen når du logger inn i appen som brukes for å registrere skader, sykdom og treningsmengde. Informasjonen er også gjengitt i dette skrevet. Du og ditt lag vil få tilgang til appen uavhengig av om du gir ditt samtykke til at dataene dine brukes i forskningsprosjektet.

Vedlegg 4: Godkjenning av bruk av figur 2

BMJ PUBLISHING GROUP LTD. LICENSE
TERMS AND CONDITIONS

Apr 09, 2021

This Agreement between Knut Helle ("You") and BMJ Publishing Group Ltd. ("BMJ Publishing Group Ltd.") consists of your license details and the terms and conditions provided by BMJ Publishing Group Ltd. and Copyright Clearance Center.

License Number	5044851081268
License date	Apr 09, 2021
Licensed Content Publisher	BMJ Publishing Group Ltd.
Licensed Content Publication	British Journal of Sports Medicine
Licensed Content Title	Understanding injury mechanisms: a key component of preventing injuries in sport
Licensed Content Author	R Bahr,T Krosshaug
Licensed Content Date	Jun 1, 2005
Licensed Content Volume	39
Licensed Content Issue	6
Type of Use	Dissertation/Thesis
Requestor type	Individual
Format	Electronic
Portion	Figure/table/extract
Number of figure/table/extracts	3
Description of figure/table/extracts	figure 3
Will you be translating?	No
Circulation/distribution	1
Title	Reliability of isometric muscle strength measurements and eccentric hamstring strength for female soccer players
Institution name	Norwegian School of Sport Sciences
Expected presentation date	Jun 2021
Portions	figure 3
Requestor Location	Knut Helle brattelia 6B

Vedlegg 5: Godkjenning bruk av figur 1

BMJ PUBLISHING GROUP LTD. LICENSE
TERMS AND CONDITIONS

Apr 09, 2021

This Agreement between Knut Helle ("You") and BMJ Publishing Group Ltd. ("BMJ Publishing Group Ltd.") consists of your license details and the terms and conditions provided by BMJ Publishing Group Ltd. and Copyright Clearance Center.

License Number	5044860729749
License date	Apr 09, 2021
Licensed Content Publisher	BMJ Publishing Group Ltd.
Licensed Content Publication	British Journal of Sports Medicine
Licensed Content Title	Doha agreement meeting on terminology and definitions in groin pain in athletes
Licensed Content Author	Adam Weir, Peter Brukner, Eamonn Delahunt, Jan Ekstrand, Damian Griffin, Karim M Khan, Greg Lovell, William C Meyers, Ulrike Muschaweck, John Orchard, Hannu Paajanen, Marc Philippon, Gilles Reboul, Philip Robinson, Anthony G Schache, Ernest Schilders, Andreas Serner, Holly Silvers, Kristian Thorborg, Timothy Tyler, Geoffrey Verrall, Robert-Jan de Vos, Zarko Vuckovic, Per Hölmich
Licensed Content Date	Jun 1, 2015
Licensed Content Volume	49
Licensed Content Issue	12
Type of Use	Dissertation/Thesis
Requestor type	Individual
Format	Electronic
Portion	Figure/table/extract
Number of figure/table/extracts	1
Description of figure/table/extracts	1
Will you be translating?	No
Circulation/distribution	1
Title	Reliability of isometric muscle strength measurements and eccentric hamstring strength for female soccer players
Institution name	Norwegian School of Sport Sciences

Vedlegg 6: Metodefeil og prosentvis variasjon fra testene

Tabell 10: Test-retest reliabilitet ME og CV_{ME}

Måling	Maksimalkraft		Gjennomsnittlig maksimalkraft	
	ME (N)	CV_{ME} (%)	ME (N)	CV_{ME} (%)
Nordic Hamstrings				
- Venstre	12,7	4	11,2	3,6
- Høyre	12,3	3,7	12,1	3,7
- (Venstre + Høyre)/2	8,8	2,7	8,8	2,8
Adduksjon				
- Venstre	9,5	6,2	8,9	6,1
- Høyre	7,1	4,6	7,4	5
Abduksjon				
- Venstre	4,6	3,3	4,6	3,4
- Høyre	5,9	4,6	5,2	3,7
Utadrotasjon				
- Venstre	7,7	6	7,8	6,4
- Høyre	5,2	4,1	5,9	4,9
Innadrotasjon				
- Venstre	11,3	8,5	10,6	8,5
- Høyre	10,7	7,5	9,9	7,4
Hoftefleksjon				
- Venstre	37,1	7	33,5	6
- Høyre	42,9	7,8	40,4	7,6

*N, Newton; ME, metodefeil = $SD\bar{d}/\sqrt{2}$; CV_{ME} , prosentvis metodefeil = $((2ME)/(\bar{X}1 + \bar{X}2)) * 100$*