

Inter-tester reliabilitet på screening av danserne ved
Nasjonalballetten i Norge
- En tverrsnittstudie



Forskningsdesign-oppgave

Antall ord: 10777

Masterprogram i helsefag – HEL320A Forskningsmetode og - etikk

Desirée Are

Universitetet i Bergen

Høsten 2014

FORORD

Denne masteren er en del av et todelt prosjekt gjort på danserne på Nasjonalballetten i Norge. Jeg ønsker å takke alle utøvere og ansatte ved Nasjonalballetten som gjorde det mulig å utføre prosjektet.

Jeg vil også rette en takk til Kjartan Vibe Fersum for god veiledning på oppgaven.

I tillegg ønsker jeg å takke Hilde Stendal Robinson og Elisabeth Julsvoll for god hjelp med metode og statistikkdelen av oppgaven.

Takk til Roger Olsen som gjennomførte prosjektet sammen med meg!

Desirée Are

12.11.14, Oslo

Innhold

SAMMENDRAG

ABSTRACT

1. INTRODUKSJON	6
1.1 Bakgrunn for valg av tema	6
1.2 Tidligere forskning på dansere	6
1.2.1 Kvalitet på tidligere forskning	6
1.2.2 Typer muskel- og skjelettplager	7
1.2.3 Risikofaktorer	7
1.2.4 Prevalens	8
1.2.5 Den Norske Nasjonalballetten	8
1.3 Teoridel	9
1.3.1 Reliabilitet	9
1.3.2 Testene inkludert i studien	10
1.3.3 Tidligere forskning av intertester reliabilitet på de ulike screeningelementene	11
2. HENSIKT OG PROBLEMSTILLING	15
2.1 Hensikt	15
2.2 Problemstilling	15
3. METODE	15
3.1 Valg av forskningsdesign	15
3.2 Utvalg	15
3.3 Inklusjons- og eksklusjonskriterier	17
3.4 Utvalgets størrelse	17
3.5 Datainnsamling	17
3.6 Testprosedyre	18
3.7 Analyse	19
3.8 Etske hensyn	22
4. RESULTATER	23
5. DISKUSJON	25
5.1 Hovedmål med studien	25
5.2 Hovedfunn	25
5.3 Funns i lys av tidligere studier	26
5.4 Metodekritikk og begrensninger i studien	32
6. KONKLUSJON	35
LITTERATURLISTE	36
VEDLEGG	41

TITTEL: Intertester reliabilitet på en screening gjort på danserne ved Nasjonalballetten i Norge

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

ÅR: 2014

SAMMENDRAG

Bakgrunn

Dansere er svært utsatte for skader. Det er gjort lite forskning på dansere i forhold til andre idrettsutøvere og de studiene som finnes er ofte av lav kvalitet. Screening av utøvere er innført for å avdekke skaderisiko og er dermed et skadeforebyggende tiltak. Helseteamet rundt utøverne består ofte av multiple personer fra varierende yrkesbakgrunner. Det er derfor viktig at testene i et screeningbatteri har god intertester reliabilitet.

Problemstilling

Hvor god er intertester reliabilitet på et sett tester ofte brukt i screening av ulike utøvere; Functional Movement Screen (FMS), Star Excursion Balance Test (SEBT) og Lunge Lean (LL)?

Metode

I dette tverrsnittstudiet ble totalt 21 personer (9 kvinner og 12 menn, gjennomsnittsalder $30,1 \pm 6,1$) randomisert fra et større prosjekt til å gjennomgå to testrunder av to testere. Resultater fra FMS og LL ble analysert med kappa (vektet og uvektet), SEBT ble analysert med intraklasse korrelasjons koeffisient (ICC).

Resultater

Det ble funnet varierende resultater i studien. Vektete kappa-verdier i FMS var mellom negativ (-0,05) til betydelig (0,61) reliabilitet. LL har uvektede kappa-verdier fra negativ (-0,30) til brukbar (0,46) reliabilitet. SEBT kommer best ut med verdier fra moderat (0,50) til tilnærmet perfekt (0,84) reliabilitet.

Konklusjon

Relative dårlige kappa-verdier antyder at FMS og LL burde brukes med forbehold. Det ble funnet god intertester reliabilitet ved bruk av SEBT. Det er likevel flere faktorer som kunne påvirke reliabilitet i denne studien, som ved eksempel tidsrom mellom testinger. For å øke intertester reliabilitet burde videre forskning inkludere bedre standardisering og opplæring av testerne.

Nøkkelord

Intertester reliabilitet, Functional Movement Screen, Star Excursion Balance Test, Lunge Lean, Nasjonalballetten

TITLE: Intertester reliability of a screening of the National Ballet of Norway**Master in Health Sciences – Manual Therapy, Department of Public Health and Primary Health Care, The University of Bergen****Year: 2014****ABSTRACT****Objective**

Dancers are highly prone to injuries. Little research has been done on dancers compared to other athletes, and the few studies done often have poor quality. Screening of athletes was introduced to assess risk factors for injury and is seen as a preventive measure. The health team supporting the athletes often consists of multiple people with various experience and backgrounds. This is the reason why it is important that the screening-battery they use show good intertester reliability.

Research question

What is the intertester reliability of a set of tests often used in screening of athletes; Functional Movement Screen (FMS), Star Excursion Balance Test (SEBT) and Lunge Lean (LL)?

Methods

This cross-sectional study randomized 21 dancers (9 female, 12 male, mean age $30,1 \pm 6,1$) from a larger project to be tested in two test-rounds by two testers. The results from FMS and LL were analysed with kappa, SEBT was analysed with Intraclass correlation coefficient (ICC).

Results

The study found varying results. Weighted kappa in FMS was from negative (-0,05) to substantial (0,61) reliability. LL results for non-weighted kappa was from negative (-0,30) to moderate (0,46) reliability. SEBT had the best results with ICC-values from moderate (0,50) to almost perfect (0,84) reliability.

Conclusion

Relatively low kappa-values imply that FMS and LL should be used with caution. The study found good intertester reliability with use of SEBT. There are however, factors that could have affected the reliability of this study, such as time between tests. To further enhance intertester reliability, future research should include improved standardization and training of test-personnel.

Key words

Intertester reliability, Functional Movement Screen, Star Excursion Balance Test, Lunge Lean, National Ballet of Norway

1. INTRODUKSJON

1.1 Bakgrunn for valg av tema

Jeg begynte å danse ballett da jeg var 7 år gammel. Siden den gang har dansen vært en stor del av livet mitt, på et tidspunkt var jeg også på vei til å bli en profesjonell danser. Selv om dette ikke lenger er aktuelt for min del er dansen fortsatt i dag en viktig del av min hverdag. Det var denne erfaringen som førte til min interesse for kroppen og hvordan den fungerer. Som fysioterapeut har den generelle interessen for dans og dansere utviklet seg til å omhandle risiko for skader og skademekanismer hos dansere. I min hverdag i fysioterapiklinikken er ikke dette en type pasienter jeg ser ofte. Det er heller ikke en gruppe idrettsutøvere det blir satt et stort fokus på sammenlignet med andre idretter.

Dansere generelt, og særlig ballettdansere, er svært utsatt for skader. En systematisk oversikt fra 2008 (Hincapie et al.) viser en stor variasjon i skadeprevalens for dansere, dette er også grunnet store forskjeller i inklusjons- og eksklusjonskriterier i gjennomførte studier. Tallene varierer mellom 3 % og 95 %, som kan tolkes til å gjelde fra ingen til alle dansere.

Før man kan gjennomføre en screening forutsetter det at de individuelle testene i screeningen er reliable og valide. Det er relativt lite studiemateriale rundt dette temaet og jeg synes derfor det er både interessant og nødvendig å være med på å videreutvikle gode screeningverktøy for en gruppe dansere.

Et søk i Pubmed 17.10.13 med søketeksten "Dancers" AND "Screening" AND ("test" OR "tests" OR "physical examination") ga 5 treff der kun en studie er en generell screening av dansere (Karim et al., 2011). Hensikten med studien var å se på intertester reliabiliteten på tester normalt brukt i screening av muskel- og skjelettplager hos dansere. Testene var basert på resultater fra en pilotstudie fra 2006, gjennomført av Dance USA Annual Post-Hire Health Screen for Professional Dancers (2007). Organet Taskforce on Dancer Health skulle fremme tiltak som vil virke forebyggende på skader hos danserne. Studien til Karim et al. (2011) ble gjennomført av fysioterapeut og veiledere, uten kontroll av gjennomføring av testingen. Noen av testene ble også gjennomført med andre testere tilstede i samme rom. Resultatene viser dårlig til moderat intertester reliabilitet og konkluderer med behov for flere studier som tar for seg reliabilitet av screeningmetoder gjort på dansere.

1.2 Tidligere forskning på dansere

1.2.1 Kvalitet på tidligere forskning

I en gjennomgang gjort av Hincapie et al. (2008) konkluderes det med at det generelt er lite forskning på dansere, og at det som har blitt gjort ofte har dårlig kvalitet, er metodisk svakt og generelt ikke møter kravene som er satt til høy standard. Det fremkommer et ønske om videre forskning på området som er

av bedre kvalitet. 103 studier ble vurdert som relevante basert på inklusjons- og eksklusjonskriteriene. Av disse ble 71 (69%) avvist på grunn av metodologiske svakheter.

Ekegren et al. (2014) viser til samme resultat; sammenliknet med andre idretter er det gjort lite forskning på både prevalens, insidens og risikofaktorer når det kommer til dansere. I sin studie evaluerte de hvordan den intense treningen som utføres hos unge atleter kan resultere i skader. De inkluderte 286 ballettstudenter fordelt på 3 danseskoler i alderen 15-19 år. Resultatene viser 378 skader, som utgjør 1,42 skader pr pers pr år. Studien viser ingen signifikant forskjell mellom jenter og gutter.

1.2.2 Typer muskel- og skjelettplager

Steinberg et al. (2011) sin studie som fokuserte på skademønster hos unge, ikke-profesjonelle dansere poengterer at de fleste studier som er utført i dette feltet er gjort på kvinnelige, profesjonelle dansere over 18 år. Den samme studien viste at 42,6 % av de inkluderte hadde skader ved første screening. Man kom frem til at de mest prevalente skadene hos unge dansere er kneskader (29,4 %), ankel- og fottendinitter, (24,5 %) og ryggplager (16,7%). Hos den yngre delen av gruppen, 8-9 år var mer utsatt for inflammatoriske reaksjoner i sener i ankel- og fotbueledd (41%), ryggplager (28%) og kneskader (11%). Dette viser at alder kan spille en rolle i utviklingen av danserelaterte skader.

1.2.3 Risikofaktorer

Det er gjort noe forskning på risikofaktorer for skadeforekomst hos dansere og det har kommet en del ulike konklusjoner på dette feltet. Anatomisk struktur, arv, trenings-rutiner, dårlig teknikk, kvalitet på underlag, alder, BMI, muskulær ubalanse, ernæring, menstruasjonssfunksjon og disiplin blir nevnt som kausale faktorer til skader hos dansere i en studie (Steinberg et al., 2011). Luke et al. (2002) ønsket med sin studie å finne verktøy for å evaluere skadeinsidens og å dermed identifisere risikofaktorer for skade. De inkluderte medisinsk historie, danseerfaring i tillegg til de nevnte faktorene ovenfor.

En artikkel i Journal of Physical Education, Recreation and Dance (Noh et al., 2003) tar for seg psykososialt stress som en risikofaktor for skade hos dansere. Det blir hevdet at de psykososiale faktorene lenge har vært oversett som risikofaktor for skadeproblematikk i alle idretter, i tillegg til hos dansere. Danserne driver seg selv hardt fysisk for å klare avanserte teknikker og overser ofte eventuelle skader for å kunne yte på høyt nivå. Patterson et al. (1998) konkludere med at det er lite forskning på området og at det er et behov for flere studier av god kvalitet og metode som kan belyse temaet.

Tretthet/utmattelse (fatigue) blir også nevnt som en risikofaktor for skade. Knicker et al. (2011) deler tretthet/utmattethet inn i en perifer og en sentral kategori. Perifer tretthet vil påvirke vevets evne til å fungere optimalt. Ved en sentral tretthetskategori vil kognitive faktorer som lav motivasjon og emosjonell ubalanse kunne årsaksforklare utmattelsen. Liederbach et al. (2013) tegner en rett strekk mellom tretthet/utmattethet, skader og redusert evne til å prestere.

Ekegren et al. (2014) antyder at det store fysiske kravet innen dans hos unge utøvere i en alder der kroppen fortsatt er i utvikling kan ha en påvirkning på skadeutvikling.

1.2.4 Prevalens

Steinberg et al. (2011) inkluderte 1336 unge, ikke-profesjonelle dansere i alderen 8 til 16 år for å kartlegge skademønstre i denne aldersgruppen. 42,6 % av de inkluderte i studien hadde skader ved første screening. Danserne på 8 år hadde en statistikk på 1 skade pr 10 dansere, men i aldersgruppen 16 år hadde man 1 skade pr 3 dansere, som er tre ganger så høy forekomst. Gamboa et al. (2008) ønsket å beskrive distribusjon av skader hos pre-profesjonelle ballettdansere. De inkluderte 204 dansere i alderen 9-20 år som ble screenet i en periode over 5 år. Resultatene fra denne studien viser at ankel- og fotskader hadde høyest innsidens (53 %), etterfulgt av hofta (21,6 %), kne (16,1 %) og rygg (9,4 %).

1.2.5 Den Norske Nasjonalballetten

I 2002 ble det foretatt en kohortstudie av den Norske Nasjonalballetten der man ville se på innsidens av muskel- og skjelettskader hos danserne (Byhring & Bo, 2002). Resultat av denne 19 uker lange studien var at 31 av 41 dansere inkludert i studien hadde en eller flere skader, der 22 % var akutte skader i fot og ankel og 74 % var bløtvevsskader. Det er ikke spesifisert nærmere i artikkelen om de akutte skadene også er bløtvevsskader eller om de 74% bløtvevsskadene kun var kroniske. Antall personer med muskelskader var 37, ligamentskader 3, seneskader 11 og andre 13. De fleste skadene ble betegnet som milde til moderate og 16 % resulterte i sykemelding. På spørsmål rapporterte 78 % av danserne at de opplevde negativt stress på jobb, 64 % oppga at de følte at de hadde lite påvirkningskraft i arbeidssituasjonen. I denne studien finner de ingen korrelasjon mellom psykososiale forhold og muskel- og skjelettplager. Denne studien inkluderte både kvinner og menn, der tidligere studier ofte har fokusert på kvinner.

Gjennom Operaen opprettet vi kontakt med Nick Allen som er clinical director for Birmingham Royal Ballet's Jerwood Centre for the prevention and treatment of dance injuries. De foretok en 3-års studie for å undersøke mulighet for å individualisere et treningsprogram til den enkelte danser basert på deres skadehistorikk for å redusere skadeinsidens (Allen et al., 2013). Studien inkluderte fra 52-58 profesjonelle dansere. De brukte Functional Movement Screen (FMS) for å kartlegge risikofaktorer for skade, for så å skreddersy individuelle treningsopplegg for den individuelle danseren. Han har videre lagt til andre, mer dansespesifikke tester og har hatt god nytte av dette. Allen påpeker at risiko for skade ofte er bygget opp av mange faktorer og dermed burde tiltak for å jobbe forebyggende også ta hensyn til alle faktorene. Allen et al. (2013) valgte denne screeningen delvis fordi ingen av testene involverer dansespesifikke bevegelser og danserne ble derfor tatt ut sin komfortsone. I teorien ville de dermed

enkler kunne oppfatte asymmetrier og kompensasjoner hos danserne. Dersom danseren mangler styrke eller stabilitet i dette området vil bevegelsesmønsteret kunne endres og man får kompensatoriske mønstre som letter kan føre til skade (Beales et al., 2009a; Beales et al., 2009b). I tillegg til FMS-testingen som er inkludert i studien, har Allen i ettertid lagt til andre tester som er mer dansespesifikke, en modifisert Romberg test, der personen står på et ben i 30 sekunder med øynene lukket, en Biodex Athletic Single Leg Balance test, og en forward flexion lunge test fra Mark Commerfords Dynamic Matrix. Det ble utviklet individualiserte treningsprogram basert på baseline av første screening. Det konkluderes med en reduksjon i antall skader for både kvinner og menn basert på det individualiserte treningsprogrammet.

Via Nasjonalballetens kontakt med Nick Allen var det ønskelig om vår screening inkluderte så godt som mulig de testene som blir utført på danserne ved Birmingham Royal Ballet's Jerwood Centre.

Variasjoner av Rombergs test blir ofte brukt for å se på nevro-muskulære og proprioceptive kvaliteter etter skade i underekstremitetene (Olmsted et al., 2002). Vi har derimot valgt å bruke Star Excursion Balance Test (SEBT). Vi hadde ikke tilgang til Biodex Athletic Single Leg Balance test-verktøy.

1.3 Teoridel

1.3.1 Reliabilitet

I denne studien ønsket vi å undersøke reliabiliteten til tre generiske tester. Generiske måleinstrument er tester som kan brukes på alle og er ikke sykdomsspesifikke (Ljunggren, 1995). Reliabilitet brukes om hvor målesikkert et testinstrument er og om resultatene fra den aktuelle testen er reproducerbar (Streiner & Norman, 2003). Det kreves av et reliabelt instrument at man må kunne vurdere resultatene, det må være en grad av stabilitet i målingen slik at resultatene vil være reproducerbare og at instrumentet må ha intern konsistens. Ved å undersøke reliabilitet vil man i større grad kunne vurdere kvaliteten til måleinstrumentet.

Reliabilitet viser også til hvilken grad et måleresultat er fri for tilfeldige feil. Med det menes at man ser på faktorer som kan ha en innvirkning på det resultatet man kommer frem til. Ulike former for tilfeldige målefeil kan være oppmerksomhet/uoppmerksomhet i målesituasjonen, unøyaktighet, mye versus lite erfaring, uklar instruksjon i utførelse av testen og testers tolkning av resultatet (Stokes, 2011).

Den sanne svaret (true score) vil være et resultat av gjennomsnittet av et uendelig antall målinger (de Vet et al., 2011).

Reliabilitet deles inn i undergruppene absolutt og relativ reliabilitet (Carter et al., 2011b). Absolutt reliabilitet måles ved å se på de faktiske testresultatenes endring i løpet av en behandlingsperiode. Ved å regne ut standard error of measurement finner man den minste endring i testresultat som må til for at man kan si at endringen ikke er tilfeldig. Relativ reliabilitet deles inn i tre undergrupper: intertester- og

intratester reliabilitet og intern konsistens.

Som beskrevet av Shrout & Fleiss (1979) vil en studie av intertester reliabilitet inkludere et randomisert utvalg av n individer som evalueres av k antall testere.

1.3.2 Testene inkludert i studien

Functional Movement Screen

Et av testapparatene brukt i denne studien er Functional Movement Screen (FMS). FMS blir brukt for å se på evne og villighet til å utføre funksjonelle bevegelsesmønstre i 7 ulike tester. Kvalitet vurderes utfra kompensatoriske mønstre (Cook et al., 2014). Beskrivelse og bilder av testene sees i vedlegg 1. Testene tar utgangspunkt i normale bevegelsesmønstre som er viktige i ulike idretter (Cook et al., 2014). De ble utviklet for å vurdere skaderisiko hos amerikanske fotballspillere (Kiesel et al., 2007). Akkurat som hos dansere ser man at risikofaktorer for skader hos fotballspillere er bygget opp av multiple faktorer (Bahr & Krosshaug, 2005). Man så at idrettsutøvere som hadde redusert dynamisk kontroll eller asymmetrier i styrke og bevegelighet også hadde en større skadeforekomst. Kiesel et al. (2007) utførte en studie på et profesjonelt lag i amerikansk fotball ($n=46$) for å se på samsvar mellom skader og score på FMS. Basert på disse resultatene er det satt en «cut off» på 14, der score ≤ 14 gir en høyre risiko for skade.

Star Excursion Balance Test

Star Excursion Balance Test (SEBT) er en balansetest som ble utviklet for å se på balanse og motorisk, postural kontroll i dynamisk setting og vil gi et annet bilde av utøveren enn ved en statisk test (Gribble et al., 2013). Studier viser at SEBT er et godt screeningverktøy for å differensiere pasienter med skader i underekstremitetene (Olmsted et al., 2002). Et systematisk review viser til SEBT som valid i forhold til funksjon i hofteabduktorer (Kivlan & Martin, 2012). Olmsted et al. (2002) fant at rekkevidden var signifikant redusert når testpersonen stod på affektert fot enn hos kontrollgruppen og konkluderte med at testen er sensitiv for personer med kronisk ankelinstabilitet. Testen er også vist signifikant for personer med rekonstruksjon av fremre korsbånd (Herrington et al., 2009) og for patello-femorale smerter (Aminaka & Gribble, 2008).

Lunge Lean

Mottram & Comerford (2008) er kritiske når det kommer til tilgjengelige screeningverktøy for å predikere risikofaktorer for idrettsutøvere. De ønsker et større fokus på funksjonelle bevegelser der man må koordinere flere ledd og muskler. Mottram & Comerford beskriver flere tester (Dynamic Matrix), hvorav lunge lean ble plukket ut av Nick Allen i sin screening av dansere. Han synes han lettere kan vurdere hvordan danserne kontrollerer hoftene i øvelser som involverer fleksjon av trunkus. Dette fører

ofte til en innoverrotasjon av femur og kan også observere valgusering av knær og rulling på fot (Allen, 2014).

Screeningen ser etter en link mellom kompensatoriske bevegelsesmønstre og smerte, men også mellom disse mønstrene i forhold til skadeforebygging. Studier viser at slike bevegelsesmønstre kan reverseres (O'Sullivan, 2005).

1.3.3 Tidligere forskning av intertester reliabilitet på de ulike screeningelementene

Det ble utført søk i PubMed for å finne relevante studier som er gjennomført på de tre testene inkludert i denne studien. De ulike søkestrengene var

1. (Functional Movement Screen) AND (intertester OR inter-tester OR interrater OR inter-rater OR interobserver OR inter-observer) AND reliability
2. (Star Excursion Balance Test) AND (intertester OR inter-tester OR interrater OR inter-rater OR interobserver OR inter-observer) AND reliability
3. (Lunge Lean OR Dynamic Matrix) AND (intertester OR inter-tester OR interrater OR inter-rater OR interobserver OR inter-observer) AND reliability

Det ble funnet 12 studier av intertester reliabilitet på FMS og av disse ble 7 ansett som relevante. Dette var studiene til Gulgin & Hoogenboom (2014), Elias (2013), Schneiders et al. (2011), Minick et al. (2010), Teyhen et al. (2012), Shultz et al. (2013) og Smith et al. (2013). Disse studiene er oppsummert i tabell 1.

Tabell 1: Studier av intertester reliabilitet på Functional Movement Screen (FMS)				
Studie	Testere	Antall testere	Populasjon	Resultat (Statistikk)
Gulgin & Hoogenboom 2014 N=20	Mellom testere av ulik erfaring	4	Skadefrie universitetsstudenter 10 kvinner; Alder 19,62 ± 0,91 10 menn; Alder 20,44 ± 2,12	50% tilnærmet perfekt 50% rimelig til brukbart (Prosent enighet)
Elias 2013 N=5	Mellom uerfarne fysioterapi-studenter	20	Skadefrie atleter	0,91 tilnærmet perfekt (ICC og t-test)
Schneiders et al. 2011 N= 58 (28% av total i hele studien)	Mellom erfarne testere	2	Normalt aktive skadefrie I totalstudien n= 209 Kvinner 108; Alder 21,2 ± 3,0 Menn 101; Alder 22,7 ± 4,2	Totalscore: 0,97 tilnærmet perfekt (ICC) Individuelle tester: 0,70-1 betydelig-tilnærmet perfekt (Uvektet kappa)

Minick et al. 2010 N= 40	To eksperter og to nybegynnere	4	Skadefrie universitetsstudenter (13 av deltakerne universitetsatleter) Kvinner 23 Menn 17 Alder (gj.snitt) 20,8	<i>Nybegynnere (17 tester):</i> Tilnærmet perfekt i 6/17 Betydelig i 8/17 Brukbar i 3/17 <i>Eksperter (17 tester):</i> Tilnærmet perfekt i 4/17 Betydelig i 9/17 Brukbar 4/17 (Vektet kappa)
Teyhen et al. 2012 N= 64	Nybegynnere: Fysioterapi-studenter	4	Skadefrie militær-ansatte Kjønn: ikke spesifisert Alder 25,2 ± 3,8	<i>Individuelle tester:</i> 0,45 til 0,82 fra brukbart til tilnærmet perfekt (Vektet kappa) <i>Totalscore:</i> (ICC (2,1))
Shultz et al. 2013 N= 39	5/6 opplært av en sertifisert FMS-administrator, 1/6 selvlært	6	Skadefrie universitetsatleter Kvinner 21; Alder 19,6 ± 1,5 Menn 18; Alder 19,7 ± 1,0	<i>Individuelle tester:</i> 0,38 svak (Krippendorff's α) <i>Totalscore:</i> Nybegynnere: 0,44 Moderat Erfarne: 0,18 svak (ICC)
Smith et al. 2013 N=19	1 sertifisert 1 erfaren, men ikke-sertifisert 1 uerfaren, men mye erfaring med bevegelses-analyse 1 uerfaren fysioterapi-student	4	Skadefrie Kvinner 9 Menn 10 Alder (median) 26	0,89-0,87 tilnærmet perfekt (ICC)

Resultatene fra de ulike studiene varierer noe når det kommer til intertester reliabilitet. Verdiene varierer fra svake (Shultz et al., 2013) til tilnærmet perfekt samsvar mellom testere (Minick et al., 2010; Schneiders et al., 2011; Teyhen et al., 2012; Elias, 2013; Smith et al., 2013; Gulgin & Hoogenboom, 2014). Selv om de fleste av studiene viser høy intertester reliabilitet er det også stor variasjon av resultater innen de ulike studiene. Gulgin & Hoogenboom (2014) så at ca. 50% av testene hadde tilnærmet perfekt prosent enighet og rimelig til brukbar enighet i 50%. Minick et al. (2010) fant kappa fra 0,54-1 (nybegynnere) og 0,43-0,95 (eksperter). I studien til Teyhen et al. (2012) sprikte resultatene mellom 0,45-0,82, fra brukbart til tilnærmet perfekt.

Variasjonen man ser i resultatene kan ha flere forklaringer. Det er en variasjon i yrke og erfaringsnivå til testerne. Det er også en forskjell i testprotokollene i studiene og dermed standardisering.

Det er variasjon i utvalget når man ser på sammensetning og størrelse. En studie har få deltakere (n=5) (Elias, 2013) og studien med flest deltakere (n=64) er den gjort av Teyhen et al. (2012). Sistnevnte studie er gjort på ansatte i militæret. Studiene til Gulgin & Hoogenboom (2014), Schneiders et al. (2011) og Smith et al. (2013) inkluderer skadefrie, normalt aktive personer. Shultz et al. (2013) inkluderer studenter knyttet til et universitets idrettslag og Elias (2013) har med profesjonelle atleter i sin studie. Minick et al. (2010) har et utvalg der noen er normalt aktive og noen er atleter ved et universitets idrettslag. De fleste studiene har en jevn fordeling av kvinner og menn og har generelt sett deltakere i en ung aldersgruppe.

Det er tydelig en stor variasjon i de analysemetodene som er brukt i de ulike studiene. Dette vil tas opp i analyse og diskusjonskapittelet.

Søkestrengen i PubMed ga kun 4 resultater for studier gjort på Star Excursion Balance Test. Disse er studiene til Gribble et al. (2013), Plisky et al. (2009), Hyong & Kim (2014) og Clark et al. (2010). I tillegg ble det funnet et studie via referanseliste til de andre studiene som ble vurdert som relevant (Hertel et al., 2000). Oppsummering av studiene vises i tabell 2.

Tabell 2: Studier av intertester reliabilitet på Star Excursion Balance Test (SEBT)				
Studie	Testere	Antall testere	Populasjon	Resultat (Statistikk)
Gribble et al. 2013 N=29	En erfaren tester og to uerfarne som fikk opplæring av den erfarene	3	Skadefrie universitetsstudenter Kvinner 19 Menn 10 Alder 31,72 ± 10,8	<i>Normalisert:</i> ANT: 0,88 PM: 0,91 PL: 0,88 Alle tilnærmet perfekt <i>Ikke-normalisert:</i> ANT: 0,92 PM: 0,92 PL: 0,92 Alle tilnærmet perfekt (ICC (1,1))
Plisky et al. 2009 N= 15	1 fysioterapi assistent 1 fysioterapeut	2	Skadefrie fotballspillere på skolelag Menn 15; Alder 19,7 ± 0,81	ANT: 0,99-1 PM: 1 PL: 0,99-1 Alle tilnærmet perfekt (ICC (2,1))
Hertel et al. 2000 N=16	Erfarne	2	Skadefrie kvinner	0,35-0,93 fra rimelig til tilnærmet perfekt (ICC)

Hyong & Kim 2014 N=67	Ikke beskrevet	2	Skadefrie, normalt aktive Kvinner 49; alder 20,5 ± 0,6 Menn 18; alder 21,4 ± 1,6	0,88-0,96 tilnærmet perfekt (ICC (2,1))
Clark et al. 2010 N=34	Ikke beskrevet	2	United States Military Academy, West Point Skadefrie Kjønn: ikke spesifisert Alder: mellom 17 og 25, ikke spesifisert gj.snitt	0,91 tilnærmet perfekt (ICC)

Disse studiene viser tilnærmet perfekt samsvar når det kommer til intertester reliabilitet, bortsett fra studien til Hertel et al. (2000). I den sistnevnte studien ser man at de ulike retningene hadde ICC-score som varierte mellom rimelig- og tilnærmet perfekt samsvar.

Her ser man igjen en variasjon når det kommer til størrelse og sammensetning av utvalg i de ulike studiene. Antall deltakere spriker fra 15 (Plisky et al., 2009) til 67 (Hyong & Kim, 2014). Hyong & Kim (2014), Gribble et al. (2013) og Hertel et al. (2000) inkluderte skadefrie, normalt aktive personer i sine studier. En studie hadde deltakere fra et universitets fotballag (Plisky et al., 2009) og Clark et al. (2010) gjorde sitt studie på ansatte i militæret. Gjennomsnittsalderen på deltakerne er igjen lav, fra 19,7 (Plisky et al., 2009) til 31,72 (Gribble et al., 2013). Man ser en ujevn fordeling av menn og kvinner i disse studiene. Hertel et al. (2000) og Plisky et al. (2009) har kun inkludert kvinner og menn, respektivt.

Ved søk i PubMed på Lunge Lean kom 0 resultater. Via litteratur som omhandler Mottram & Comerford (2008) Dynamic Matrix fant jeg én studie som inkluderte denne testen. Dette var studien til Monnier et al. (2012).

Tabell 3: Studier av intertester reliabilitet på Lunge Lean (LL)				
Studie	Testere	Antall testere	Populasjon	Resultat (Statistikk)
Monnier et al. 2012 N=33	Erfarne fysioterapeuter	2	Skadefrie ansatte i marinen Kvinner 1 Menn 32 Alder 28,7 ± 5,9	> 0,61 betydelig (kappa)

Studien ønsket å vurdere inter- og intratester reliabilitet for klinisk passende tester bevegelsesmønstre- og kontroll for rygg og hofta i marinen (Monnier et al., 2012). De valgte ut 6 tester basert på beskrivelsene til Mottram & Comerford, derav Lunge Lean.

Studien viser betydelige kappaverdier. Studien inkluderte ansatte i marinen og har ujevn fordeling av kvinner og menn. Igjen er utvalget tatt av en relativt ung aldersgruppe.

2. HENSIKT OG PROBLEMSTILLING

2.1 Hensikt

I en klinisk hverdag der helseteamet i et idrettsmiljø ønsker å gjennomføre screening av utøverne for å avdekke skaderisiko og deretter bruke informasjonen fra denne screeningen til forebyggende tiltak for utøverne er det viktig at testene er reliable.

Karim et al. (2011) understreker at det ikke er gjort noen intertester reliabilitetsstudier på screening av profesjonelle samtidsdansere før deres studie i 2009 og at det dermed er behov for videre forskning på dette feltet. Dette er viktig i de tilfellene der det er flere personer som skal gjennomføre testingen av danserne (Minick et al., 2010).

Målet med studien var å undersøke reliabilitet til et tenkt sett av screeningvektøy.

Denne studien var en del av et todelt studie der intertester reliabilitet og et forsøk på validering av de samme testene ble gjennomført.

2.2 Problemstilling

Hvor god er intertester reliabiliteten til disse testene som ofte blir brukt i screening av ulike utøvere: Functional Movement Screen, Star Excursion Balance Teste og Single Leg Small Knee Bend + Lunge Lean.

3. METODE

3.1 Valg av forskningsdesign

Denne studien kan beskrives som en reliabilitetsstudie, der to terapeuter har testet de samme testene på samme person med et gitt tidsintervall imellom. I studier der man ser på samsvar mellom resultater fra to ulike testere klassifiserer man dette som intertester reliabilitet. Det ble brukt et kvantitativt, ikke-eksperimentelt forskningsdesign (Carter et al., 2011a). Reliabilitetsstudier kommer innunder kategorien kvantitativ forskningsdesign. Med det menes at resultatene måles og man bruker tall eller mengde for å beskrive resultatene (Halvorsen, 2002).

3.2 Utvalg

Danserne blir invitert til å delta i studien av to manuellterapistudenter, samt fysioterapeut, osteopat,

akupunktør og ledelse ved Den Norske Opera. Utvalget blir beskrevet som et bekvemmelighetsutvalg (sample of convenience) fordi personene som ble testet var gitt på bakgrunn av arbeidsplassen deres (Carter et al., 2011c). Det vil si at de som deltok i studien ikke ble randomisert til å være med. Ettersom vi ikke kan vite hvor representative disse danserne er fra alle andre dansere er det ikke lett å generalisere resultater til den generelle dansepopulasjonen.

Av totalt 70 dansere deltok 43 for å danne en baseline i forbindelse med det større prosjektet. Resten av danserne var enten i permisjon eller ekskludert fra studien grunnet tidskonflikter og skader. 21 av danserne ble retestet som en del av reliabilitetsstudien. Beskrivelse av utvalget kan ses i tabell 4. Man ser at det er stor grad av variasjon i alder, antall skader og smertenivå. Smertenivå ble målt med Numeric Pain Rating Scale der testpersonen angir smerteintensitet med et helt tall fra 0-10 (Gallasch & Alexandre, 2007). Det var også god spredning i roller ved Nasjonalballetten, fra aspiranter til solister. Dette kan ha effekt på hvor mye den enkelte trener og dermed på totalbelastning. Terapeutene var blindet for disse faktorene.

Alle danserne måtte gi sitt samtykke for deltakelse i studien. Det ble utformet et informasjonsskriv med informasjon om hva studien innebar og hva som var hensikten. Danserne måtte i tillegg signere en samtykkeerklæring. Nasjonalballetten består av mange internasjonale dansere og disse dokumentene ble derfor utformet på engelsk.

Tabell 4: Deskriptiv statistikk	
Kategorier	Alle deltakere
Alder	
Gjennomsnitt	30,14
Standard avvik	6,05
Variasjonsbredde	19-39
Kjønnfordeling	
Kvinner	42,9%
Menn	57,1%
Antall skader pr person	
Gjennomsnitt	4,29
Standard avvik	3,15
Variasjonsbredde	1-11
Antall dager mellom testing	
Gjennomsnitt	10,90
Standard avvik	8,88
Variasjonsbredde	2-34
Smerte (NPRS)	
Gjennomsnitt	2,10
Standard avvik	1,73
Variasjonsbredde	0-5

3.3 Inklusjons- og eksklusjonskriterier

Inklusjonskriterier:

- Profesjonell danser ved den norske nasjonalballetten
- Alder mellom 18-40

Eksklusjonskriterier:

- Svangerskap
- Relevante kirurgiske inngrep de siste seks (6) månedene
- Akutte skader som fører til uteblivelse av trening/forestilling

Eksklusjonskriteriene er valgt for å utelukke faktorer som kan påvirke de testresultatene vi får. Inngrep og skader vil kunne redusere fysisk funksjon og dermed kan vi få ekstreme verdier (outliers) i resultatene. Mye av tidligere studier er gjort på kun kvinnelige dansere, men vi valgte å inkludere både de kvinnelige og de mannlige danserne for å få med nyansene i hele utvalget.

3.4 Utvalgets størrelse

Utvalgets størrelse er viktig for å kunne generalisere funnene til en større populasjon (Carter et al., 2011c). Ved bruk av ICC er det mulig å regne ut det laveste antall deltakere som kreves for at konfidensintervallet skal få en ønsket bredde (De Vet et al. 2011). Ved bruk av Cohens Kappa er det ikke like gode muligheter for å finne dette antallet.

I eksperimentell forskning ønsker man gjerne 30 eller flere deltakere for å kunne generalisere (Carter et al., 2011c). Andre mener at det være minimum 50 deltakere for å få god statistisk power (Terwee et al., 2007). Det var derfor ønskelig at alle 41 dansere som deltok i prosjektet ville vært tilgjengelig for retest. Totalt deltok 21 av danserne på retest. Det var satt opp flere retester, men grunnet konflikter i timeplanen var det flere som ikke møtte opp.

3.5 Datainnsamling

Forskerne bestod av to blindede terapeuter, begge masterstudenter i manuellterapi ved UiB. All testingen ble utført i to standardiserte lokaler ved Den Norske Opera. Som en del av det større prosjektet ble 43 dansere randomisert til å bli testet av en av de to testerne. Denne randomiseringen ble utført av administrative medlemmer av Nasjonalballetten. Danserne ble tildelt et tilfeldig nummer før screeningen og har kun blitt identifisert via dette nummeret.

Over 30 ble randomisert til retest. Av ulike årsaker var det til slutt 21 som ble retestet.

Ved andregangs testing ble testerne byttet, slik at danseren ble testet en gang av hver tester. Det var tilfeldig hvem av testerne som testet ved baseline og hvem som tok retest.

Ved førstegangs test fikk danseren generell informasjon om prosjektet og en forklaring av de ulike testene. Alle instruksjoner ble gitt muntlig da vi ville unngå å demonstrere testen. Ved baselinetesting skulle alle danserne fylle ut et skaderegistreringsskjema der de scoret aktuell smerte på NPRS og skrev ned alle relevante skader. Dette førte til at førstegangs test tok ca. ½ time. Andregangs testing tok noe kortere tid, ca. 20min.

Timeplan for testingen ble også organisert av administrasjonen for å unngå at screeningen tok danserne bort fra klasser og prøver. Dansernes normale arbeidsdag er fra 09.30 til 17.00 og for å unngå overtid ble testingen utført innenfor denne tidsrammen. Dette var et krav fra ledelsen ved Nasjonalballetten for å gjennomføre prosjektet. Det var ønskelig med et standardisert mellomrom mellom hver testing. Dette for å forhindre tretthet ved for nære testinger, som det ville blitt dersom en danser ble testet to ganger på samme dag. Vi ønsket også å unngå at det gikk for lang tid mellom hver testing, slik at eksterne faktorer ikke skulle innvirke på resultatet, for eksempel skader eller sykdom. Det ble dessverre vanskelig å få til dette, noe som har ført til store variasjoner i tidsintervall mellom testene.

Minick et al. (2010) definerer i sin studie en ekspert som en som har vært med på å utvikle tester og/eller som har mer enn 10 års erfaring. En nybegynner blir definert som en som har utført et standard kurs og har mindre enn 1 års erfaring. Smith et al. (2013) definerer en ekspert som en som har flere enn 100 gjennomførte tester.

De to testerne i dette prosjektet blir begge kategorisert som nybegynnere innen disse definisjonene.

All innsamling av data forgikk i tidsrommet januar-april 2014.

3.6 Testprosedyre

Bilder og beskrivelse av testene finnes i vedlegg nummer 1.

- *Functional Movement Screen*: De 7 testene ble scoret på en skala fra 0-3. 0 ble gitt dersom utøver har smerter i provokasjonstest, 3 er perfekt utføring av oppgave. Poeng fra de enkelte testene ble summert til en sluttscore. Begge sider ble testet og dersom de ble scoret ulikt ble laveste score valgt.
- *Star Excursion Balance Test*: danseren stod med hælen på testbenet i midten av en opp-teipet stjerne og strakk ut det andre benet så langt som mulig til hver av strålene og gjentok med det andre benet. Hver person fikk to øvinger på hvert ben før vi målte rekkevidden.

Det var opptegnet centimeter på teipen og rekkevidden ble markert med en finger og så skrevet ned.

- *Lunge Lean*: Testpersonen stod med to fotlengder mellom hælen på det bakerste benet og tærne på det fremste benet, med bena plassert på linje. Personen overførte vekt til det fremste benet og strakk ut det andre slik at det falt i samme linje som resten av kroppen på en diagonal linje i 45°. Denne stillingen ble holdt i minimum 5 sekunder. Vi observerte i hvilken grad personen klarte å opprettholde stillingen med nøytral rygg, bekken og hofter, samt valgusering av knær og rulling på fot.

Standardisering av testingen

Ved at terapeutene samarbeidet i utformingen av testprotokollen ledet dette til standardisering av testene. De ble gjennomgått teoretisk og det ble sett videodemonstrasjon i utførelse og analyse av FMS og SEBT. Dette var ikke tilgjengelig for LL og testerne kom til enighet om kriterier for vurdering av testen. Alle testene ble prøvd ut i praksis av begge terapeuter. Det ble holdt daglig kontakt i løpet av test-perioden, som også påvirket standardiseringen.

3.7 Analyse

Deltagerne ble beskrevet med deskriptiv statistikk. Etter innsamlingen av dataene ble fullført ble de analysert. Dataene i denne studien kommer innenfor både nominalnivå (*Lunge Lean*), ordinalnivå (FMS) og på intervallnivå (SEBT).

I analyse der man ser på grad av samsvar mellom flere testere brukes ofte kappa koeffisient når data er på nominal- og ordinalnivå og intraklasse korrelasjons koeffisient (ICC) ved intervallnivå (de Vet et al., 2011).

Intraklasse korrelasjons koeffisient

I klassisk teori blir det beskrevet at relativ reliabilitet ved observasjon er bygget på to elementer: et faktisk resultat og målefeil (Streiner & Norman, 2003; Viera & Garrett, 2005). Alle typer målinger foretatt i atferdsvitenskap er utsatt for målefeil, særlig når det involverer vurderinger gjort av mennesker (Shrout & Fleiss, 1979). Man ønsker å finne et samsvar i resultater der to testere har brukt det samme instrumentet med samme betingelser, i tillegg til å se om bruk av samme instrument på det samme individet vil gi det samme resultatet.

Det finnes flere versjoner av ICC og hver av dem vil analysere data ulikt og dermed gi ulike resultater (Shrout & Fleiss, 1979). Som en ser i tabell 2 er det uenighet i hvilken metode som er den beste når det kommer til intertester reliabilitet. De aktuelle versjonene for denne studien er ICC (2,1) og ICC (3,1).

Dersom man vil se på absolutt samsvar (absolute agreement) mellom to tester brukes ICC (2,1), der testerne blir sett som tilfeldige utvalgte testere i et stort utvalg av testere (Shrout & Fleiss, 1979). ICC (2,1) ser på testerne som tilfeldige faktorer; som et utvalg av alle mulige testere (Rousson et al., 2002) Det vil være en naturlig varians i måleresultater hvis man ser på alle mulige testere og alle mulige utøvere (Shrout & Fleiss, 1979). Her blir det ikke justert for systematiske feil fordi man ønsker å generalisere resultatene til den større populasjon. Rousson et al. (2002) anbefaler å bruke ICC (2,1) ved intertester reliabilitet, da rekkefølgen på testerne ikke har noe å si.

Når to testere skal vurdere de samme individene blir koherens (consistency agreement) målt med ICC (3,1), da testerne blir sett på som fikserte (Shrout & Fleiss, 1979). ICC (3,1) brukes ofte der man har to testere. Testerne blir sett på som fikserte og man er kun interessert i samsvaret mellom disse to. Det tas hensyn til rekkefølgen av testingen, og dermed læringseffekt og tretthet, på repeterte målinger. Fordi det i ICC (3,1) kun er to testere, justerer denne analysen for systematiske feil i dataene, f.eks. dersom den ene testeren systematisk gir høyere eller lavere score enn den andre. Når dataene er korrigert for systematiske feil gjenstår bare tilfeldige feil (de Vet et al., 2011).

Absolute samsvar brukes for å se om testere gjør lik vurdering av testene og er det som brukes mest i medisinsk forskning (Streiner & Norman, 2003; de Vet et al., 2011). Verdier for absolutt samsvar vil alltid være mindre enn ved koherens (Streiner & Norman, 2003; de Vet et al., 2011).

Det ble valgt å gjøre begge analysene med absolutt samsvar for å se på differansene mellom de to. Resultatene fra de to analysene ble identiske.

Styrkegrad			Styrkegrad	
Svak	0,0 – 0,2		Svak	0,00 – 0,20
Rimelig	0,3 – 0,4		Rimelig	0,21 – 0,40
Moderat	0,5 – 0,6		Brukbart	0,41 – 0,60
Sterk	0,7 – 0,8		Betydelig	0,61 – 0,80
Tilnærmet perfekt	>0,8		Tilnærmet perfekt	0,81 – 1,00

Oversikt intraklasse korrelasjons koeffisient

Oversikt Cohens kappa

(Landis & Koch, 1977)

I denne tabellen vises kun en rangering av verdier mellom 0-1. Man kan også få verdier mellom -1-0. Dersom man får en negativ kappa-verdi vil det si at samsvaret i resultatene er svakere enn det som ville vært forventet ved ren gjetning (Sim & Wright, 2005). Dette oppstår vanligvis dersom det kun er to testere og er uvanlig når det er flere enn to testere.

I flere av studiene gjort på FMS er det brukt ICC som analysemetode, både av de individuelle testene og av totalsummen (Elias, 2013; Smith et al., 2013). De individuelle variablene i FMS-testen er på ordinalnivå og det ble derfor brukt kappa-analyse på disse testene i min studie. Totalsum for alle testene er på intervallnivå og ble dermed analysert med ICC.

Cohens Kappa Koeffisient

Når man har data med nominale eller ordinale variabler blir Kappa koeffisient brukt for å se på nivå av samsvar mellom de ulike testerne (Streiner & Norman, 2003). En Kappa analyse vil være «chance-corrected» der man tar høyde for hvor stor sannsynlighet det er for samsvar utover tilfeldighet basert på de dataene som finnes (Landis & Koch, 1977). Kappa-verdiene blir regnet ut basert på forskjellen mellom forventet resultat og det faktiske resultatet.

Kappa ble derfor brukt som analysemetode for Lunge Lean (nominal) og Functional Movement screen (ordinal) i min studie. Denne analysemetoden er meget sensitiv for distribusjon av resultater. Dersom det er en overvekt av et av scoringsalternativene, vil stor enighet i høy grad være basert på tilfeldighet (chance) (Kundel & Polansky, 2003). På en skala med ordinale verdier vil det være et forhold mellom de ulike verdiene. En normal kappa-analyse vil ikke vekte disse verdiene i forhold til hverandre. Ved bruk av vektet kappa blir større differanser straffes mer enn mindre differanser.

Functional Movement Screen har verdier på ordinal skala. Det ble derfor brukt en vektet kappa-analyse av denne testen, slik som beskrevet i Cohens Kappa (Sim & Wright, 2005).

Kappa-resultatene viser at det er for lite variasjon/spredning i scorene. Dette fører til meget svake resultater i denne studien. Det er i tillegg derfor regnet ut prosent enighet (percentage agreement). Her blir det ikke tatt høyde for i hvilken grad samsvaret mellom de to testerne er tilfeldig, slik som i Cohens Kappa (de Vet et al., 2011). Dersom testerne kun er enige som et resultat av tilfeldigheter, så er de ikke egentlig enige (Sim & Wright, 2005). Kappa-verdier vil derfor alltid være lavere enn ved prosent enighet.

Analysen av Intraklasse korrelasjons koeffisient og generalisert kappa ble gjort ved hjelp av Statistical Package for the Social Sciences (SPSS, versjon 22 for Word). Vektet kappa ble regnet ut i Excel 2013.

Konfidensintervall sier noe om eventuelle feilmarginer i en måling. Intervallene indikerer om resultatene

er innenfor den sanne verdien av den variabelen man har evaluert. Dette vises ved et 95% konfidensintervall (Ringdal, 2009). Konfidensintervall for uvektet kappa ble funnet ved hjelp av bootstrapping.

Det ble også gjort en parret t-test for å finne systematiske forskjeller hos de to testerne.

3.8 Etiske hensyn

Studien fulgte de etiske prinsippene som er angitt i Helsinki-deklarasjonen (WMA - World Medical Association, 2014). Helsinki-deklarasjonen setter fokus på at hensyn til en persons helse og integritet er viktigere enn forskning. Det var frivillig å delta i studien og det var viktig at det ikke inngikk et press fra arbeidsgiver for å delta i studien. Danserne kunne på hvilket som helst tidspunkt trekke seg fra studien uten at dette fikk konsekvenser for videre arbeidsforhold. Deltakerne ble informert om studiens hensikt og skrev under på en samtykkeerklæring.

Ved at studien fant sted ved Nasjonalballetten vil total anonymitet ikke være mulig. All innhentet data ble de-identifisert ved at deltakerne kun var kjent for oss via et randomisert tildelt nummer. Alle involverte i prosjektet hadde taushetsplikt.

Siden det er få studier gjort på screeningmetoder på dansere ser vi studien og det større prosjektet som et ledd i å forebygge prevalens av skader hos danserne. Studien bidrar dermed i å evidensbasere et skadeforebyggende tiltak. Prosjektet var tidskrevende for deltakerne, men ulempene ved å være med i studien regnes som relativt små. Det medisinske teamet rundt danserne kan få tilgang til resultatene med den individuelle danserens godkjennelse.

Det er et mål at dataene i studien presenteres så objektivt og korrekt som mulig og det finnes ingen underliggende forventninger i forhold til resultater.

I forhold til at studien går innunder medisinsk- og helsefaglig forskning på mennesker ble det søkt og godkjent av regional etisk komité. REK sør-øst godkjente prosjektet 10.04.14. se vedlegg 4.

4. RESULTATER

Av 70 dansere ble 43 inkludert i det større prosjektet. Av disse ble over 30 randomisert til retest. En ble ekskludert pga. skade mellom test-tidspunktene, en ble ekskludert grunnet sykdom, en ønsket ikke å være med på retest og de 6 andre møtte ikke opp grunnet tidskonflikt.

Tabell 5: Resultater Functional Movement Screen					
Test	Enighet	Uenighet	Uvektet Kappa (95% Konfidens Intervall)	Vektet Kappa	Prosent enighet
FMS1	11	10	-0,07 (-0,35-0,27)	-0,03	52,4%
FMS2	19	2	-0,05 (-0,13-0)	-0,05	90,5%
FMS3	14	7	0,32 (-0,03-0,70)	0,32	66,6%
FMS4	19	2	0,61 (-0,07-1)	0,61	90,5%
FMS5	21	0	- (-)	-	100%
FMS6	11	10	0,36 (0,08-0,62)	0,51	52,4%
FMS7	16	5	0,19 (-0,17-0,64)	0,36	76,2%

FMS1: Squat, FMS2: Hurdle step, FMS3: Inline lunge, FMS4: Shoulder mobility, FMS5: Active straight leg raise (ASLR), FMS6: Trunk stability, FMS7: Rotary stability

Tabellen over presenterer resultatene fra FMS-testingen. Ved analysering av dataene ble det satt opp krysstabeller for alle testene og basert på disse ble kappa koeffisient og prosent enighet regnet ut. I tillegg ble kappa-resultatene vektet og 95% konfidensintervall regnet ut.

Man kan se en forskjell i vektet vs. uvektet kappa i noen av testene. I disse testene er det større spredning i resultatene. I de tilfellene verdien for vektet- og uvektet kappa er lik ser man liten spredning i resultatene. Generelt sett ser vi høy prosent enighet (66-100%) med unntak av Squat og Trunk Stability som ligger litt over 50%.

Både Squat og Hurdle Step har negative kappa-verdier (vektet og uvektet), men som nevnt over har Hurdle Step høy prosent enighet. De negative verdiene er en refleksjon av lite spredning i resultatene. Inline Lunge har rimelige kappa-verdier og moderat prosent enighet, og Shoulder Mobility viser høy prosentvis enighet og en kappa-verdi på 0,61 som vurderes som betydelig. I de to tilfellene der testerne scoret ulikt i denne testen var dette grunnet smerter på ett av testtidspunktene. I ASLR ser vi 100% enighet mellom de to testerne og kappa-verdien blir vurdert som konstant fordi det ikke er noe variasjon i dataene. Resultatene til Trunk Stability og Rotary Stability viser den største forskjellen mellom uvektede og vektete kappa-verdier grunnet større spredning i testresultatene.

Totalsummen for FMS resulterte i ICC 0,45 (rimelig, 95% KI 0,10-0,73) og hadde range mellom 11 og

20 (gjennomsnitt $15,67 \pm 1,91$).

Tabell 6: Resultater Star Excursion Balance Test			
Test	ICC	95% konfidens intervall	Cronbach's alfa
ANT høyre	0,61	0,24-0,82	0,75
ANT venstre	0,67	0,35-0,85	0,81
AL høyre	0,84	0,64-0,93	0,91
AL venstre	0,63	0,28-0,83	0,77
LAT høyre	0,73	0,45-0,88	0,84
LAT venstre	0,57	0,18-0,60	0,72
PL høyre	0,76	0,49-0,69	0,86
PL venstre	0,50	0,09-0,77	0,66
POST høyre	0,82	0,60-0,92	0,90
POST venstre	0,82	0,62-0,92	0,91
PM høyre	0,68	0,37-0,86	0,81
PM venstre	0,81	0,60-0,92	0,90
MED Høyre	0,70	0,33-0,87	0,86
MED venstre	0,76	0,50-0,89	0,86
AM høyre	0,72	0,42-0,87	0,83
AM venstre	0,54	0,17-0,78	0,71

ANT: Anterior, AL: Anteriolateralt, LAT: Lateral, PL: Posteriolateralt, POST: Posterior, PM: Posteriolateralt, MED: Medialt, AM: Anteriomedialt

Tabellen viser resultatene fra Star Excursion Balance Test. Her ser man at 4 av retningene viser tilnærmet perfekt korrelasjon mellom testerne (AL høyre, POST høyre, POST venstre og PM venstre), 4 av retningene viser sterk korrelasjon (LAT høyre, PL høyre, MED venstre og AM høyre) og 8 av retningene viser moderat korrelasjon (ANT høyre og venstre, AL venstre, LAT venstre, PL venstre, PM høyre, MED høyre og AM venstre).

Tabell 7: Resultater Lunge Lean				
Test	Enighet	Uenighet	Kappa (95% Konfidens Intervall)	Prosent enighet
LL1 høyre	16	5	0,43 (0-0,79)	76,2%
LL1 venstre	12	9	0,03 (-0,25-0,35)	57,1%
LL2 høyre	7	14	0 (0)	33,3%
LL2 venstre	12	9	0 (0)	57,1%
LL3 høyre	20	1	0 (0)	95,2%
LL3 venstre	20	1	0 (0)	95,2%
LL4 høyre	21	0	Konstant (-)	100%
LL4 venstre	21	0	Konstant (-)	100%
LL5 høyre	18	3	0,35 (0-0,88)	85,7%
LL5 venstre	18	3	0,46 (0-1)	85,7%

LL6 høyre	11	10	0,02 (-0,33-0,41)	52,4%
LL6 venstre	11	10	-0,30 (-0,49- -0,09)	52, 4%
LL7 høyre	19	2	-0,05 (-0,13-0)	90,5%
LL7 venstre	20	1	0 (0)	95, 2%

LL1: rotasjon bekken, LL2: valgus kne, rulle inn på fot, LL3: fot ut, hæl inn, LL4: krumme rygg, LL5: hyperekstensjon rygg, LL6: sidebøy/tilt bekken, LL7: holde ikke-vektbærende ben i rett linje med resten av kroppen

Tabellen over presenterer resultatene av LL-testingen. En ser høy prosentvis enighet (>75%) i 9 av testene, 4 av testene ligger på rundt 50% enighet og på en test (LL2 høyre) er det lav prosentvis enighet (33%). Det er brukbar og rimelig kappa-verdi på 3 av testene og nokså svake kappaverdier på de resterende testene. I LL2 høyre og venstre, LL3 høyre og venstre og LL7 venstre er det ingen forskjell på forventet (expected) og målt (count) resultat. Dermed får man en kappa-verdi på 0. I LL6 venstre og LL7 høyre er det minus-kappaverdier, der en kunne fått bedre samsvar ved tilfeldig utfylling av resultater. På LL4 både høyre og venstre er det 100% enighet og dette reflekteres i kappa-verdiene.

Det er generelt stor spredning i konfidensintervaller i Functional Movement Screen, Star Excursion Balance Test og Lunge Lean noe som behefter resultatene med stor usikkerhet.

5. DISKUSJON

5.1 Hovedmål med studien

Det er generelt få studier av dansere og av disse er få gjort på screeningmetoder. Valg av testbatteri for denne studien baseres på Nick Allens studie på Birmingham Royal Ballet's Jerwood Centre. Studien til Allen et al. (2013) tar ikke for seg reliabiliteten til de enkelte testene, men ser på effekt av et personalisert treningsopplegg basert på en baseline score hos danserne.

Målet med denne studien var å undersøke intertester reliabilitet på en screening av danserne ved Nasjonalballetten. Basert på samarbeidet mellom Nasjonalballetten og Nick Allen var det et ønske om å bruke noe av det samme testbatteriet i denne studien.

5.2 Hovedfunn

I vår studie ble det funnet varierende resultater i de tre testene. FMS har vektete kappa-verdier med intervall fra negativ (-0,05) til betydelig (0,61) reliabilitet. LL har uvektede kappa-verdier med intervall fra negativ (-0,30) til brukbar (0,46) reliabilitet. SEBT er den testen som kommer best ut med verdier fra

moderat (0,50) til tilnærmet perfekt (0,84) reliabilitet. De relativt store differansene leder til at det er usikkerhet i resultatene og dette vises også i de vide konfidensintervallene.

Diskusjonsdelen vil sammenligne resultatene fra denne studien opp mot tidligere studier. Det vil også diskuteres ulike faktorer som kan årsaks-forklare funnene i denne studien.

5.3 Funn i lys av tidligere studier

Functional Movement Screen

Resultater

Resultatene i denne studien viser stor variasjon av vektet kappa for FMS, fra negative verdier til betydelig verdi. I de 7 tidligere studiene av intertester reliabilitet på FMS varierte verdiene fra svake (Shultz et al., 2013) til tilnærmet perfekt samsvar (Minick et al., 2010; Schneiders et al., 2011; Teyhen et al., 2012; Elias, 2013; Smith et al., 2013; Gulgin & Hoogenboom, 2014). Jeg vil gjennomgå hva dette innebærer i mer detalj nedenfor.

Analysemetode

Det er stor variasjon i analysemetoder i de tidligere studiene. 4 av studiene (Elias, 2013; Shultz et al., 2013; Smith et al., 2013; Gulgin & Hoogenboom, 2014) har brukt andre analysemetoder, henholdsvis ICC, Krippendorfs alfa, ICC og Fisher's Exact, enn det som er brukt i vår studie og er derfor vanskelige å sammenligne. Det er usikkert hvordan det ville påvirket resultatene i vår studie dersom scorene hadde blitt analysert med disse metodene.

De tre andre studiene brukte kappa for sine analyser, både normal (Schneiders et al., 2011) og vektet (Minick et al., 2010; Teyhen et al., 2012). I disse studiene varierer resultatene fra rimelig (0,43) til tilnærmet perfekt (1) samsvar, der det ble funnet best samsvar med bruk av normal kappa-analyse (Schneiders et al., 2011). Totalt sett ligger resultatene i disse tre studiene høyere enn i min. Mulige årsaker til dette er utvalgets sammensetning, testernes erfaring og utforming av testprotokoll, noe som vil diskuteres i mer detalj nedenfor.

Resultatene fra denne studien viser svakest samsvar ved bruk av Squat og Hurdle Step. I de fleste av studiene har Squat gode resultater (Minick et al., 2010; Schneiders et al., 2011; Shultz et al., 2013), med unntak av studien til Gulgin & Hoogenboom (2014). I tillegg ble det funnet svake resultater for Inline Lunge og Rotary Stability, noe som er likt resultater fra tidligere studier (Minick et al., 2010; Schneiders et al., 2011; Shultz et al., 2013; Gulgin & Hoogenboom, 2014).

Populasjon

Sammensetning av populasjon kan ha en påvirkning på resultater i studien, spesielt der man tester fysisk

funksjon. I denne studien ble testene utført på profesjonelle ballettdansere med høy totalbelastning og forekomst av belastningsrelaterte skader. Tre av de tidligere studiene er gjort på normalt aktive personer uten skader (Schneiders et al., 2011; Smith et al., 2013; Gulgin & Hoogenboom, 2014) og dette utvalget vil differensiere nokså mye i forhold til dansere og resultatene er derfor vanskelig å sammenligne. Minick et al. (2010) har et utvalg der noen er normalt aktive og noen er atleter ved et universitets idrettslag. Eventuelle forskjeller mellom disse to gruppene blir ikke diskutert i studien og det er usikkert i hvilken grad det spiller inn. Teyhen et al. (2012) utførte sin screening på ansatte i militæret. Samlet belastning for denne gruppen kan sammenlignes med dansere, selv om bevegelsesmønsteret vil avvike en del, noe som kan gi utslag i resultatene. To av studiene (Elias, 2013; Shultz et al., 2013) har inkludert, henholdsvis, studenter ved et universitets idrettslag og profesjonelle atleter. Dette utvalget kan best sammenlignes med det brukt i vår studie selv om både analysemetode og resultater er ulike i alle tre studiene.

Alle de tidligere studiene inkluderte deltakere fra en ung aldersgruppe. Gjennomsnittsalderen varierer fra 19, 6 til 26 år og ligger litt lavere enn gjennomsnittsalder i vår studie (30,1, range 19-38). Det er usikkert hvordan resultatene i denne studien ble påvirket av alder, men det var en tendens til at de eldre danserne hadde bedre resultater enn de yngre.

Takeffekt

FMS ble utviklet for å avdekke risikofaktorer hos profesjonelle spillere i amerikansk fotball og senere også brukt på atleter i mange andre idretter. To av testene i denne screeningen hadde full score hos samtlige av danserne, med unntak av fire av danserne som ble scoret ned på grunn av smerter. Begge er mobilitetstester gjort på utøvere som har krav til stor fleksibilitet og viser at testen har en takeffekt når utført på denne gruppen (McHorney & Tarlov, 1995). Selv om alle danserne hadde full score på ASLR vil ikke det nødvendigvis bety at de har den fleksibiliteten som er kreves for dansen fordi testen krever kun at utøveren klarer å flektre benet over 90°.

Testernes bakgrunn

Begge testerne i denne studien er fysioterapeuter under videreutdanning, men med lite erfaring med FMS-testing, noe som kan ha påvirket resultatene. Dette er i samsvar med metodene brukt av Teyhen et al. (2012) og Elias (2013) der det kun ble brukt fysioterapistudenter med lite erfaring som testere. Noen av de andre studiene som er utført har sammenlignet testere med lite og mye erfaring (Minick et al., 2010; Smith et al., 2013; Gulgin & Hoogenboom, 2014). De fant en tendens til bedre samsvar hos de uerfarne testerne enn hos de erfarne, som kan ha en positiv effekt i vår studie. Schneiders et al. (2011) og Shultz et al. (2013) har kun med erfarne testere i sine studier og kan derfor sammenlignes i minst

grad med vår studie når det gjelder testernes bakgrunn.

Testprotokoll

Evalueringen av testingen i denne studien ble utført med testeren tilstede, mens noen av de tidligere studiene har tatt videoopptak for evaluering av testene (Minick et al., 2010; Elias, 2013; Shultz et al., 2013; Gulgin & Hoogenboom, 2014). Bruk av videoopptak er hensiktsmessig i den grad at man kan inkludere flere personer i studien siden utøveren kun trenger å bli testet en gang. Videoopptak vil derimot mangle dimensjoner man kun får dersom testeren er tilstede under testingen og disse studiene skiller seg derfor i stor grad fra vår studie. Gulgin & Hoogenboom (2014) lot sine testere se videoene om igjen mens de gjorde sin vurdering og kun i normal hastighet. Det er ikke angitt i de andre studiene om de kunne observere videoene flere ganger. Tilgang til videoopptak kan gi økt sikkerhet i vurderingsgrunnlag for studien, da er vanskelig å observere alle kompensasjoner direkte.

Star Excursion Balance Test

Resultater

Resultatene i vår studie viser noe variasjon av ICC for SEBT, fra moderate (0,50) til tilnærmet perfekte (>0,8) verdier. Det ble funnet 5 tidligere studier av intertester reliabilitet av SEBT som viste liten grad av spredning, da de fleste fant tilnærmet perfekt samsvar, med høyest grad av enighet hos Plisky et al. (2009) mellom 0,99-1. Et unntak er studien til Hertel et al. (2000) der ICC-score varierte fra rimelig (0,35) til tilnærmet perfekt (0,93) samsvar.

Analyse

Alle studiene har brukt ICC i analyse av sine resultater, men har brukt ulike varianter av denne analysemetoden. Som tidligere beskrevet er det uenighet i hvilken metode som egner seg best for denne type undersøkelse. I denne studien ble ICC (2,1) med absolutt samsvar brukt, som i studiene til Plisky et al. (2009) og Hyong & Kim (2014). Bruk av lik metode gjør at de bedre kan sammenlignes med vår studie. Begge studiene har tilnærmet perfekte resultater med ICC fra 0,88 til 1 noe som er litt høyere enn i vår studie. Årsak til dette kan være studienes sammensetning eller testprotokoll. I tillegg ble det gjort en ICC (3,1) analyse med absolutt samsvar av våre data, uten endring av resultatene.

To av studiene har ikke spesifisert hvilken ICC som er brukt (Hertel et al., 2000; Clark et al., 2010) og ett av studiene har brukt ICC (1,1) (Gribble et al., 2013). Dermed er det vanskelig å direkte sammenligne resultatene fra disse tre studiene.

Til tross for bruk av ulike ICC-varianter, viser de fleste studiene fra moderat til høy grad av enighet og det er usikkert om analysemetode har påvirket resultatene i stor grad.

Populasjon

Ballettdanserne i denne studien lar seg best sammenligne med utvalget i studiene til Plisky et al. (2009) og Clark et al. (2010) som har brukt, henholdsvis, spillere på et universitetsfotballag og ansatte i militæret. Begge har funnet tilnærmet perfekte resultater og ligger dermed noe høyre enn i vår studie. Andre studier har inkludert et utvalg som vanskeligere kan sammenlignes basert på totalbelastning og skadeforekomst (Hertel et al., 2000; Gribble et al., 2013; Hyong & Kim, 2014).

Testprotokoll

Det er stor variasjon i testprotokoll i de ulike studiene. De første som undersøkte reliabilitet av SEBT var Kinzey & Armstrong (1998). Deltakerne ble oppfordret til å maksimere utslag i rekkevidde ved blant annet å flektre kneet eller hofte på standbeinet eller å ekstendere ryggen. I tillegg til rekkevidde så de også på kvaliteten i bevegelsen for å mulig forklare eventuelle forskjeller mellom to ben. De oppfordrer også til innøving av testens bevegelser før selve utførelsen, siden bevegelsesmønsteret ikke er tilnærmet normalt. Danserne i vår studie ble oppfordret til å maksimere utslag på tilsvarende måte, men grunnet begrenset tidsrom fikk danserne bare to testrunder før målingen fant sted. Dette kan ha ført til at en læringseffekt har påvirket resultatene. Dette beskrives nærmere nedenfor.

Det er vanlig at stjernen er tegnet eller teipet opp på gulvet, noe som ble brukt i vår studie. Kinzey & Armstrong (1998) tegnet opp en «boks» der foten måtte plasseres inni og Plisky et al. (2009) brukte «the Y balance test kit» der utøveren står på en boks og skyver et objekt med foten langs stenger som går ut fra midten. Utøveren strekker det ene benet så langt ut på en stråle som mulig og gir hos Gribble et al. (2013) et lett trykk på tærne. Hvis det legges vekt på den utstrakte foten må denne testen gjøres igjen og Kinzey & Armstrong (1998) lot ikke utøveren være nær underlaget med den utstrakte foten. I vår studie fikk deltakeren være nær gulvet, men ikke legge vekt på foten. I noen studier holder deltakerne begge hender i hoftefeste (Robinson & Gribble, 2008; Gribble et al., 2013; Hyong & Kim, 2014), mens Hertel et al. (2000) lot sine utøvere bruke armene som balanseorgan. Vi lot danserne bruke armene fritt i vår studie og en kritikk til bruk av armene er at det stiller større krav til koordinasjon og forlenger tiden det tar for å innlære øvelsen. Deltakerne i Plisky et al. (2009) og Kinzey & Armstrong (1998) studier brukte treningssko, mens andre har sine deltakere barfot eller med sokker (Hertel et al., 2006; Gribble et al., 2013). Alle deltakere i vår studie var barføtte. Det kommer tydelig frem at det finnes mange variasjoner i standardisering av denne testen, som vil si at det er vanskelig å sammenligne resultater med hverandre eller med vår studie.

I vår studie ble alle 8 retningene av stjernen målt, noe som også ble gjort av Hertel et al. (2000) og Hyong & Kim (2014). Andre har brukt den modifiserte utgaven av testen der kun 3 av retningene blir

målt (Plisky et al., 2009; Clark et al., 2010; Gribble et al., 2013). Denne modifiserte testen ble beskrevet av Hertel et al. (2006) da de ønsket å undersøke om noen av retningene i SEBT korrelerte bedre hos personer med eller uten kronisk ankelinstabilitet enn andre. De fant at den posteriomediale retningen hadde høyest korrelasjon både hos de med og uten instabilitetsproblemer i ankel. I tillegg så man høy korrelasjon også i medial og anteriomedial retning. I vår studie fant vi best samsvar i anteriolateral, posterior og posteriomedial retning.

Antall øvinger kan ha en effekt på resultatet grunnet en læringseffekt som oppstår ved gjentakelser av en øvelse. Hertel et al. (2000) foreslo en innføring av 6 runder med øving og en hvileperiode før man gjennomfører testen. Dette er også standard i studien til Hyong & Kim (2014). Testen kan bli nokså tidkrevende med opptil 9 runder av testingen gjort bilateralt og Robinson & Gribble (2008) ønsket derfor å se om det var mulig å redusere antall øvingsrunder og fortsatt ta hensyn til læringseffekten. De fant at man fortsatt hadde høy reliabilitet ved å redusere antall test-runder til 4 som brukes i studien til Gribble et al. (2013). I vår studie ble det kun gjennomført to testrunder og til tross for dette fant vi fra moderate til tilnærmet perfekte ICC-verdier. Det er mulig at flere øvinger før man gjennomførte målingene ville ført til enda bedre resultater.

Flere av studiene har normalisert sine data (Plisky et al., 2009; Clark et al., 2010). Det vil si at man deler oppnådd rekkevidde (cm) på lengde (cm) på standbeinet. Man får dermed et forholdstall mellom de to. Gribble et al. (2013) brukte både normalisert og ikke-normalisert data i sin studie og fikk større samsvar på ikke-normalisert data enn på normalisert. I vår studie ble resultatene ikke normalisert som i en del andre studier, men som Gribble et al. (2013) fant har dette ikke alltid en positiv effekt på resultatene.

Under SEBT kan vanligvis utøveren komme tilbake til en to-fots stilling, uten å endre stillingen mye med vektoverføring (Gribble & Hertel, 2003). I vår studie beholdt danserne den utøvende foten i luften hele tiden og dette kan ha påvirket resultatene, da belastningen på standbeinet var større og kan ha påført tretthet.

Kinzey & Armstrong (1998) mener at bevegelsesmønsteret i SEBT ikke tilsvarer normale, funksjonelle bevegelser hos verken normalt aktive personer eller hos idrettsutøvere. Bevegelsen regnes dermed som ny. Gribble & Hertel (2003) ser det derimot som en fordel at testen er såpass vanskelig å utføre. Det fører til at idrettsutøvere blir mer utfordret og man lettere ser kompensasjoner. Danserne er, i mulig større grad enn andre idrettsutøvere, vant til å utfordre kroppen i ulike retninger og posisjoner, så man kan tenke seg at bevegelsesmønsteret ikke var helt nytt for dem. Den største forskjellen i testutførelsen vil være at de

måtte stå med beinet i en «parallell» posisjon, der de vanligvis arbeider med benet i en «turn-out» -stilling, med maksimal utoverrotasjon i hofte.

Resultatene fra SEBT ble i vår studie målt fra en enkelt måling. De fleste andre studier utfører tre ulike målinger og bruker et gjennomsnitt av disse målingene som verdi. Det vil alltid bli et mer reliabelt resultat ved bruk av gjennomsnittsmålinger fordi antall målefeil reduseres (de Vet et al., 2011). Dette kan ha hatt innvirkning på resultatene vår studie, da ICC-verdiene kunne vært annerledes dersom vi hadde brukt gjennomsnittsmåling.

Lunge Lean

Resultater

I vår studie viser resultatene en stor variasjon i kappa-verdier, fra negativt (-0,30) til brukbart (0,46) samsvar. Studien til Monnier et al. (2012) fant betydelige kappa-verdier (0,60/0,63). Det nevnes ikke i studien til Monnier et al. (2012) om resultatene er justert for prevalens- og bias (PABAK), noe som ikke er gjort i vår studie. Dette kan være en mulig forklaring for differensen i resultatene. Dette vil diskuteres nærmere nedenfor.

Analyse

Studien til Monnier et al. (2012) har brukt kappa som analysemetode slik som i min studie. Som nevnt ovenfor er det mulig at en PABAK-analyse kan være grunnen til differansen i resultatene.

Populasjon

Monnier et al. (2012) inkluderte ansatte i marinen i sin studie. Det finnes en overføringsverdi av totalbelastning hos disse sammenlignet med profesjonelle danserne i vår studie, selv om det er usikkert hvor høy skadeforekomst det er blant ansatte i marinen.

Aldersgruppen i studien til Monnier et al. (2012) er tilsvarende lik utvalget i vår studie, med gjennomsnittsalder på rundt 28 år. Det er derimot en meget ujevn fordeling av kvinner og menn i forhold til i vår studie. Det er usikkert i hvor stor grad dette har påvirket resultatene og også dette kan være en forklaring på den store forskjellen i resultatene.

Testernes bakgrunn

Testerne i studien til Monnier et al. (2012) er erfarne fysioterapeuter, noe som kan sammenlignes med testerne i denne studien. Lunge Lean er med i et testbatteri som ofte brukes i marinen og det er sannsynlig at testerne inkludert i studien til Monnier et al. (2012) har mer erfaring med denne testen, enn testerne i vår studie.

Testprotokoll

I studien til Monnier et al. (2012) ble evalueringen gjort simultant. Det vil si at begge testerne var tilstede i rommet samtidig, men var blindet for hva den andre testeren scoret. Selv om de ikke visste hva den andre scoret, observerte de den samme utøveren på samme tidspunkt og påvirkning av eksterne faktorer mellom tidspunkt for testing ble redusert. I vår studie var det kun en tester tilstede under testing, noe som kan føre til lavere samsvar grunnet eksterne faktorer som eksempelvis dagsform og nye skader. Dette vil bli nærmere diskutert i neste kapittel.

5.4 Metodekritikk og begrensninger i studien

Denne studien har flere begrensninger. En av disse er at utvalgets størrelse ikke ble større ($n=21$). Basert på anbefalinger gjort av Carter et al. (2011a) og de Vet et al. (2011) ønsket vi å inkludere over 30 personer til retest. Dette ble begrenset av dansernes hektiske timeplan.

Når man ser på de andre studiene varierer de en del når det kommer til størrelse på utvalg. Antall deltakere i studiene på FMS varierte fra 5 til 64, studiene på SEBT varierte fra 15-67 og antall inkluderte i studiene på LL var 33.

En konsekvens av et lite utvalg kan være lavere resultater i både kappa og ICC.

Dataanalyse

Kappa-analyse er sensitiv for prevalens- og bias indeks. Når det er meget høy eller meget liten prevalens av et resultat i det samme utvalget vil kappa resultere i svak reliabilitet selv når man har høy prosent enighet (Brennan & Silman, 1992; Cunningham, 2009). Dette blir sett på som et av kappa-statistikkens paradoks og det anbefales å legge til andre verdier, da ren kappa blir sett på som utilstrekkelig. Det er noe uenighet om hvilke andre verdier som burde legges til, men det er vanlig å ta med prosent enighet, det forventede samsvaret og prevalens- og bias indeks.

Med prosent enighet kan det se ut som samsvaret er større enn det virkelig er, fordi det kapitaliserer av tilfeldige resultater. Siden det ikke tas høyde for tilfeldige resultater mener Hunt (1986) at metoden ikke burde brukes i studier som undersøker intertester reliabilitet. Vi fant generelt høy prosent enighet, som kan indikere at danserne var et homogent utvalg.

I kappastatistikken ble forventet og faktisk samsvar regnet ut og prosent enighet ble regnet ut basert på antall tester med enighet. Bias viser til systematiske forskjeller i score for tester 1 og 2. Denne forskjellen i bias vil føre til en høyere kappa-verdi, også et av kappa-statistikkens paradoks (Viera & Garrett, 2005). Det var ønskelig å gjøre en prevalens- og bias justert kappa (Prevalence-adjusted-bias-adjusted-kappa, PABAK) – analyse, men grunnet mangel på programvare var dette beklageligvis ikke mulig å

gjennomføre.

Hoehler (2000) er kritisk til PABAK i kappa-analyse og mener at bias og prevalens' effekt på kappa-verdier gir oss mye informasjon. Han mener at disse ikke burde justeres for, da den nye kappa-verdien ikke fremstiller målesituasjonen slik den egentlig var.

En ting man kan diskutere er de nokså grove scoringsalternativene som inkluderes i FMS og i LL. I FMS ser man etter kompensasjoner i utførelsen av testen og det er noen ganger vanskelig å vurdere om en utøver scorer en høy 2 eller en lav 3. I LL scoret man kun ja eller nei ettersom utøveren klarte å unngå kompensatoriske mønstre og en ble i noen tilfeller usikker på å score testen som positiv eller negativ dersom det var antydning til kompensasjon.

Intraklasse korrelasjons koeffesient er også sensitiv for prevalens. Dersom man scorer mange av den samme testen med samme verdi er det vanskelig å skille mellom dem. I disse tilfellene vil selv små målefeil kunne gi en stor effekt i resultatene (de Vet et al., 2011). En ser at der det er større variasjon i målinger vil dette resultere i færre målefeil. ICC er dermed påvirkelig av en gruppes hetero- og homogenitet.

Standardisering

I begynnelsen av testperioden hadde ingen av testerne utført noen av de tre testene inkludert i studien på andre enn den andre tester. Selv om testene var gjennomgått både teoretisk og praktisk var man antakeligvis ikke like observant på kompensatoriske mønstre i begynnelsen. I løpet av testperioden ble man mer oppmerksom på små kompensasjoner man kanskje ikke hadde lagt merke til i begynnelsen. Et mulig resultat av dette var at man satte strengere krav til gjennomføringen av testen og dermed lettere scoret ned. Dette kan være et resultat av læringseffekt hos testerne. I tillegg begynte man å merke seg visse mønstre som gikk igjen, særlig i kvaliteten på utførelse. Det er mulig at en etter hvert begynte å se etter disse mønstrene. Dette kan føre til et tilfeldig samsvar i score, der testerne forventer et resultat på forhånd og der man da er i tvil vil gi den scoren som er forventet (Kundel & Polansky, 2003).

Resultatene fra paret t-test viste at tester 1 generelt sett ga noe høyre score enn tester 2. Dette kan si noe om for lite innøving eller standardisering av testene (Rousson et al., 2002). Andre ting som kan ha påvirket testerne er oppmerksomheten i test-øyeblikket. Dersom testerne ikke har fullt fokus på utøveren vil man kunne overse viktige faktorer. Siden testingen skulle gjennomføres raskt kan det også være unøyaktighet i målinger og nedtegnelser av resultater. Testerne kan også ha vært uklare eller forklart testen på en annen måte enn den andre tester.

Star Excursion Balance Test og Lunge Lean test blir oftest brukt i terapeutisk sammenheng. FMS blir også brukt i terapeutisk sammenheng, men er ofte brukt av trenere og andre som ikke har spesifikk

utdannelse eller erfaring i bevegelsesanalyse. I studien til Smith et al. (2013) så man at den som hadde minst erfaring med FMS-testing, men lang erfaring med bevegelsesanalyse var den som viste høyest intra-tester reliabilitet.

Tid mellom tester

Det finnes mange andre faktorer som også spiller inn på de resultatene vi ser i denne studien. En av de viktigste faktorene i denne sammenhengen er den tiden som gikk mellom de to testene hos den enkelte. Gjennomsnitt antall dager mellom testing var $10,9 \pm 8,8$ dager, med range fra 2 til 34 dager. En grunn til moderat reliabilitet kan være at resultatene til en person vil kunne variere grunnet faktorer som ulike bevegelsesstrategier i første- og andregangs testing og mental status (Kinzey & Armstrong, 1998). Ønsket var å ha 2-3 dager mellom hver testing, fordi det er mye som kan tilkomme når det går lenger tid enn dette. I løpet av uker som har gått mellom testingen til enkelte kan det ha oppstått sykdom, skader, endringer i belastning og dermed tretthet. Denne trettheten kan innvirke på optimal funksjon hos danseren (Knicker et al., 2011) og vil dermed redusere prestasjonen til den enkelte (Liederbach et al., 2013).

Smerter

Smerter vil også innvirke på resultatene til den enkelte, der overbelastningsskader spiller en like stor rolle som akutte skader (Bahr, 2009). Selv om danserne i studien blir betegnet som friske ser vi en stor variasjon i antall skader og smertenivå hos den enkelte. Et stykke ut i testperioden ble testerne oppmerksomme på at NPRS burde vært scoret begge gangene danseren var inne til testing. Dette ble derfor kun gjort på noen av de siste som ble retestet og det er dermed vanskelig å komme med gode konklusjoner på hvordan dette spilte inn på eventuelle variasjoner i testresultat. Det ses en tendens hos de som scoret på NPRS begge gangene at smertenivået lå noe høyere ved andre gangs testing. Dette kan ha med tidsperioden for testingen å gjøre. Da vi startet opp testingen etter nyttår var danserne utvilte etter en pause fra trening og forestillinger. Mot slutten av testperioden var det to forestillinger som ble øvd til og danset simultant. I FMS 4 for Shoulder Mobility var det kun to av deltakerne som ble scoret ulikt og i begge tilfellene ble danseren scoret ned pga. smerte.

Generalisering av resultater

I utgangspunktet vil resultatene fra denne studien kun være gyldig for de som deltok i studien, da man ikke kan si at deltakerne er representative for alle dansere. Vi kan dermed ikke generalisere resultatene fra denne studien til den generelle dansepopulasjonen.

6. KONKLUSJON

Denne studien var en del av et større prosjekt for å screene alle danserne ved Nasjonalballetten i Norge. Utfra baselinescore vil helseteamet ved Nasjonalballetten utvikle individualiserte treningsprogram som et skadeforebyggende tiltak. Studien hadde som formål å undersøke intertester reliabilitet på tre tester som vil bli brukt i denne screeningprosessen av danserne fremover.

Vi fant at SEBT hadde fra moderate til gode resultater på samsvar mellom de to testerne og vurderes derfor som et reliabelt screening verktøy som kan brukes av flere testere på samme utøver. Dette korrelerer godt med tidligere studier av denne testen.

To av testene (FMS og LL) hadde svake resultater og dermed lav grad av samsvar mellom de to testerne. Tidligere studier varierer mye i sine resultater av intertester reliabilitet og til tross for relativ stor grad av prosent enighet i denne studien, vil lave kappa-verdier antyde at testene burde brukes med forbehold. Det er flere faktorer som kan ha påvirket resultatene i denne studien som burde tas hensyn til i fremtidig forskning.

Fremtidig forskning

En gruppe masterstudenter gjorde et studie der man sammenlignet bruk av ulike typer kappa-analyser på samme datasett (O'Leary et al., 2014). De fant at dersom man bare målte samsvar mellom to testere var det grunn til å betvile resultatene, da de så at parete samsvar i ulike grupper varierte meget. Det er dermed grunn til å tro at vurdering fra flere testere ville bedret resultatene i vår studie.

Videre forskning på intertester reliabilitet burde ha et større utvalg for å genere mer «power» i resultatene for å kunne generalisere til en større populasjon. Gjennomføringen av testingen burde også gjøres av personer med god erfaring. Det anbefales å bruke tid på å skape en god testprotokoll og en solid standardisering av testene, for å være mer nøyaktig i både gjennomføring og evaluering. Det anbefales ytterligere at helseteamene som jobber med utøvere utarbeider sin egen standardisering og reliabilitetstesting for å øke samsvar mellom enkelttestere.

LITTERATURLISTE

- Allen N. Private e-mail. (Ed Are D, Olsen R); 2014.
- Allen N, Nevill AM, Brooks JH, Koutedakis Y, Wyon MA. The effect of a comprehensive injury audit program on injury incidence in ballet: a 3-year prospective study. *Clin J Sport Med* 2013; 23 (5): 373-378.
- Aminaka N, Gribble PA. Patellar taping, patellofemoral pain syndrome, lower extremity kinematics, and dynamic postural control. *J Athl Train* 2008; 43 (1): 21-28.
- Bahr R. No injuries, but plenty of pain? On the methodology for recording overuse symptoms in sports. *Br J Sports Med* 2009; 43 (13): 966-972.
- Bahr R, Krosshaug T. Understanding injury mechanisms: a key component of preventing injuries in sport. *Br J Sports Med* 2005; 39 (6): 324-329.
- Beales DJ, O'Sullivan PB, Briffa NK. Motor control patterns during an active straight leg raise in chronic pelvic girdle pain subjects. *Spine (Phila Pa 1976)* 2009a; 34 (9): 861-870.
- Beales DJ, O'Sullivan PB, Briffa NK. Motor control patterns during an active straight leg raise in pain-free subjects. *Spine (Phila Pa 1976)* 2009b; 34 (1): E1-8.
- Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *Bmj* 1992; 304 (6840): 1491-1494.
- Byhring S, Bo K. Musculoskeletal injuries in the Norwegian National Ballet: a prospective cohort study. *Scand J Med Sci Sports* 2002; 12 (6): 365-370.
- Carter RE, Lubinsky J, Domholdt E. *Group Designs*. In: Falk K (Ed) *Rehabilitation Research; Principles and Applications*. Missouri: Elsevier Saunders, 2011a; 105-116.
- Carter RE, Lubinsky J, Domholdt E. *Measurement Theory*. In: Falck K (Ed) *Rehabilitation Research; Principles and Applications*. Elsevier Saunders, 2011b; 229-244.
- Carter RE, Lubinsky J, Domholdt E. *Selection and Assignment of Participants*. In: (Ed) *Rehabilitation Research; Principles and Applications*. Missouri: Elsevier Sanders, 2011c; 92-104.
- Clark RC, Saxion CE, Cameron KL, Gerber JP. Associations between three clinical assessment tools for postural stability. *N Am J Sports Phys Ther* 2010; 5 (3): 122-130.
- Cook G, Burton L, Hoogenboom BJ, Voight M. Functional movement screening: the use of fundamental movements as an assessment of function - part 1. *Int J Sports Phys Ther* 2014; 9 (3): 396-409.
- Cunningham M. More than Just the Kappa Coefficient: A Program to Fully Characterize Inter-Rater Reliability between Two Raters. In: *SAS Global Forum*. Washington DC: SAS Global Forum; 2009; 7.
- Dance/USA Taskforce on Dancer Health An Annual Post-Hire Health Screen for Professional Dancers:

Phase-One of a Preventive Healthcare Program for Professional Dance Companies 2007;

Available from:

https://www2.danceusa.org/uploads/Dancer_Health/Resources_Overview_Statement.pdf

[Accessed 07.10.2014].

de Vet HCW, Terwee CB, Mokkink IB, Knol DI. *Reliability*. In: (Ed) Measurement in Medicine: A Practical Guide. Cambridge: Cambridge University Press, 2011; 96-149.

Ekegren CL, Quested R, Brodrick A. Injuries in pre-professional ballet dancers: Incidence, characteristics and consequences. *J Sci Med Sport* 2014; 17 (3): 271-275.

Elias JE. The Inter-rater Reliability of the Functional Movement Screen within an athletic population using Untrained Raters. *J Strength Cond Res* 2013.

Gallasch CH, Alexandre NM. The measurement of musculoskeletal pain intensity: a comparison of four methods. *Rev Gaucha Enferm* 2007; 28 (2): 260-265.

Gamboa JM, Roberts LA, Maring J, Fergus A. Injury patterns in elite preprofessional ballet dancers and the utility of screening programs to identify risk characteristics. *J Orthop Sports Phys Ther* 2008; 38 (3): 126-136.

Gribble PA, Hertel J. Considerations for normalizing measures of the Star Excursion Balance Test. *Measurement in physical education and exercise science* 2003; 7 (2): 89-100.

Gribble PA, Kelly SE, Refshauge KM, Hiller CE. Interrater reliability of the star excursion balance test. *J Athl Train* 2013; 48 (5): 621-626.

Gulgin H, Hoogenboom B. The functional movement screening (fms): an inter-rater reliability study between raters of varied experience. *Int J Sports Phys Ther* 2014; 9 (1): 14-20.

Halvorsen K. *Forskningsmetode for helse- og sosialfag: en innføring i samfunnsvitenskaplig metode (pocket)*: Cappelen Akademisk, 2002.

Herrington L, Hatcher J, Hatcher A, McNicholas M. A comparison of Star Excursion Balance Test reach distances between ACL deficient patients and asymptomatic controls. *Knee* 2009; 16 (2): 149-152.

Hertel J, Braham RA, Hale SA, Olmsted-Kramer LC. Simplifying the star excursion balance test: analyses of subjects with and without chronic ankle instability. *J Orthop Sports Phys Ther* 2006; 36 (3): 131-137.

Hertel J, Miller SJ, Denegar CR. Intratester and intertester reliability during the Star Excursion Balance Test. *Journal of Sports Rehabilitation* 2000; 9 (2).

Hincapie CA, Morton EJ, Cassidy JD. Musculoskeletal injuries and pain in dancers: a systematic review. *Arch Phys Med Rehabil* 2008; 89 (9): 1819-1829.

Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *J Clin Epidemiol* 2000; 53 (5): 499-503.

- Hunt RJ. Percent agreement, Pearson's correlation, and kappa as measures of inter-examiner reliability. *J Dent Res* 1986; 65 (2): 128-130.
- Hyong IH, Kim JH. Test of intrarater and interrater reliability for the star excursion balance test. *J Phys Ther Sci* 2014; 26 (8): 1139-1141.
- Karim A, Millet V, Massie K, Olson S, Morgenthaler A. Inter-rater reliability of a musculoskeletal screen as administered to female professional contemporary dancers. *Work* 2011; 40 (3): 281-288.
- Kiesel K, Plisky PJ, Voight ML. Can Serious Injury in Professional Football be Predicted by a Preseason Functional Movement Screen? *N Am J Sports Phys Ther* 2007; 2 (3): 147-158.
- Kinzey SJ, Armstrong CW. The reliability of the star-excursion test in assessing dynamic balance. *J Orthop Sports Phys Ther* 1998; 27 (5): 356-360.
- Kivlan BR, Martin RL. Functional performance testing of the hip in athletes: a systematic review for reliability and validity. *Int J Sports Phys Ther* 2012; 7 (4): 402-412.
- Knicker AJ, Renshaw I, Oldham AR, Cairns SP. Interactive processes link the multiple symptoms of fatigue in sport competition. *Sports Med* 2011; 41 (4): 307-328.
- Kundel HL, Polansky M. Measurement of observer agreement. *Radiology* 2003; 228 (2): 303-308.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33 (1): 159-174.
- Liederbach M, Schanfein L, Kremenic IJ. What is known about the effect of fatigue on injury occurrence among dancers? *J Dance Med Sci* 2013; 17 (3): 101-108.
- Ljunggren AE. Perspektiver på vurderingsmetoder innen faget fysioterapi. *Fysioterapeuten* 1995; 62 (1): 23-26.
- Luke AC, Kinney SA, D'Hemecourt PA, Baum J, Owen M, Micheli LJ. Determinants of injuries in young dancers. *Medical Problems of Performing Artists* 2002; 17 (3).
- McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995; 4 (4): 293-307.
- Minick KI, Kiesel KB, Burton L, Taylor A, Plisky P, Butler RJ. Interrater reliability of the functional movement screen. *J Strength Cond Res* 2010; 24 (2): 479-486.
- Monnier A, Heuer J, Norman K, Ang B. Inter- and intra-observer reliability of clinical movement-control tests for marines. *BMC Musculoskeletal Disorders* 2012; 13 (1): 263.
- Mottram S, Comerford M. A new perspective on risk assessment. *Physical Therapy in Sport* 2008; 9 (1): 40-51.
- Noh YE, Morris T, Andersen MB. Psychosocial Stress and Injury in Dance. *Journal of Physical Education, Recreation & Dance* 2003; 74 (4): 36-40.
- O'Leary S, Lund M, Ytre-Hauge TJ, Holm SR, Naess K, Dalland LN, McPhail SM. Pitfalls in the use of

- kappa when interpreting agreement between multiple raters in reliability studies. *Physiotherapy* 2014; 100 (1): 27-35.
- O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. *Man Ther* 2005; 10 (4): 242-255.
- Olmsted LC, Carcia CR, Hertel J, Shultz SJ. Efficacy of the Star Excursion Balance Tests in Detecting Reach Deficits in Subjects With Chronic Ankle Instability. *J Athl Train* 2002; 37 (4): 501-506.
- Patterson EL, Smith RE, Everett JJ, Ptacek JT. Psychosocial factors as predictors of ballet injuries: Interactive effects of life stress and social support. *Journal of Sport Behavior* 1998; 21 (1): 101-112.
- Plisky PJ, Gorman PP, Butler RJ, Kiesel KB, Underwood FB, Elkins B. The reliability of an instrumented device for measuring components of the star excursion balance test. *N Am J Sports Phys Ther* 2009; 4 (2): 92-99.
- Ringdal K. *Enhet og mangfold, Samfunnsvitenskapelig forskning og kvantitativ metode*: Kunnskapsforlaget, 2009.
- Robinson RH, Gribble PA. Support for a reduction in the number of trials needed for the star excursion balance test. *Arch Phys Med Rehabil* 2008; 89 (2): 364-370.
- Rousson V, Gasser T, Seifert B. Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Stat Med* 2002; 21 (22): 3431-3446.
- Schneiders AG, Davidsson A, Horman E, Sullivan SJ. Functional movement screen normative values in a young, active population. *Int J Sports Phys Ther* 2011; 6 (2): 75-82.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86 (2): 420-428.
- Shultz R, Anderson SC, Matheson GO, Marcello B, Besier T. Test-retest and interrater reliability of the functional movement screen. *J Athl Train* 2013; 48 (3): 331-336.
- Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; 85 (3): 257-268.
- Smith CA, Chimera NJ, Wright NJ, Warren M. Interrater and intrarater reliability of the functional movement screen. *J Strength Cond Res* 2013; 27 (4): 982-987.
- Steinberg N, Siev-Ner I, Peleg S, Dar G, Masharawi Y, Zeev A, Hershkovitz I. Injury patterns in young, non-professional dancers. *J Sports Sci* 2011; 29 (1): 47-54.
- Stokes EK. *Rehabilitations outcome measures*. Edinburg: Churchill Livingstone Elsevier, 2011.
- Streiner DL, Norman GR. *Reliability*. In: (Ed) *Health measurement scales - a practical guide to their development and use*. Oxford: Oxford University Press, 2003.
- Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin*

Epidemiol 2007; 60 (1): 34-42.

Teyhen D, Shaffer SW, Lorenson CL, Halfpap JP, Donofry DF, Walker MJ, Dugan JL, Childs JD. The Functional Movement Screen: A Reliability Study. *Journal of Orthopaedic & Sports Physical Therapy* 2012; 42 (6): 530-540.

Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; 37 (5): 360-363.

WMA - World Medical Association. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. 2014; Available from:

<http://www.wma.net/en/30publications/10policies/b3/> [Accessed 19.10.2014].

VEDLEGG

Vedlegg 1

TESTPROTOKOLL

Testperson kommer inn i rommet og får informasjon om hvordan testingen vil foregå. Det vil ikke bli vist hvordan testene skal utføres og tester skal ikke rette på utøver. I FMS scores alle testene fra 1-3, der 3 er best og 1 (0) er dårligst. Man legger sammen poengscoren.

De utøverne som skal testes to ganger skal randomiseres og testes i randomisert rekkefølge hos de to testerne.

1. Deep squat

Instruksjon: Stå med føttene parallelt i skulderbreddes avstand. Hold stokken over hodet i en V-form med strake albuer. Utøveren gjør en dyp knebøy med hælene i gulvet, mens hode og bryst vender fremover

Scoring:

3. Utøver gjennomfører oppgave uten problemer, med hælene i gulvet og med ansikt og bryst vendt fremover
2. Utøver må ha tillagt høyde under føttene (planken) for å utføre oppgaven
1. Klarer ikke å utføre oppgaven



2. Hurdle step

Utstyr: høyden på stangen måles til samme høydesom utøvers tuberositas tibia

Instruksjon: Plasser føttene sammen med tærne mot planken. Hold stangen med begge hender bak på øvre del av skuldre, under nakken. Skritt over stangen med ett ben og sett hælen ned i gulvet. Hold standbeinet strakt. Løft benet tilbake til startposisjon og gjenta med det andre benet. Denne øvelsen kan gjentas opp til tre ganger på hvert bein.

Scoring:

3. Utøver gjennomfører øvelsen uten å kompensere på noen måte, eks rotere eller addusere/abduere i hoftene
2. Utøver kompenserer med rotasjon eller lener seg til siden
1. Utøver mister balansen eller kommer nær stangen



3. In-line lunge

Utøver står med føttene på planken. Lengde fra gulv til tuberositas tibia måles i cm og måles opp på planken. Mål lengden fra storetå og sett et kryss på planken.

Instruksjon: Plasser den venstre hælen mot kanten av planken. Hold stangen mellom hendene bak på ryggen, med den venstre hånden i nakken og den høyre hånden ved halebenet. Planken skal være i kontakt med kroppen i hele øvelsen. Gå fremover og plasser høyre hæl mot det markerte punktet. Bøy knærne helt til venstre kne kommer ned til gulvet. Gå tilbake til startposisjon og gjenta med det andre benet og bytt grepet på planken. Denne øvelsen kan gjentas opp til 3 ganger på hvert ben.

Scoring:

3. Utøver gjennomfører øvelsen uten kompensasjon

2. Utøver gjennomfører øvelsen med kompensasjon, eks rotasjon, lener seg til siden eller frem og tilbake

1. Utøver mister balansen eller klarer ikke å gjennomføre



4. Shoulder mobility

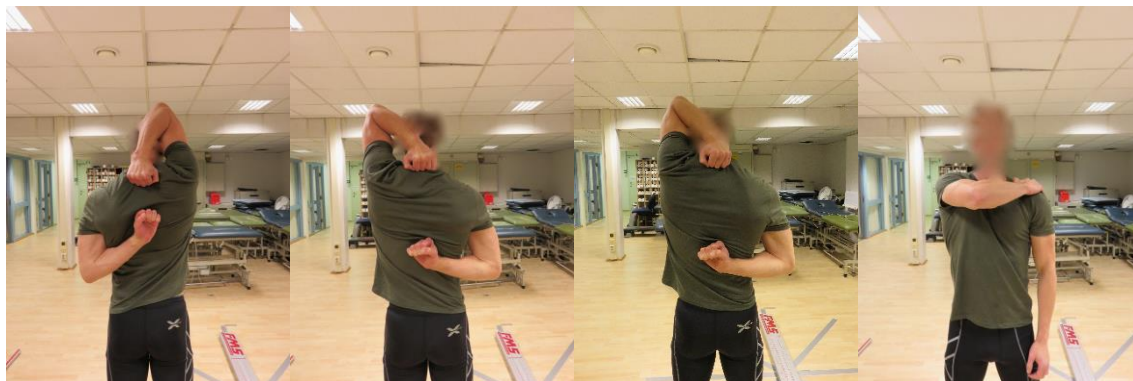
Mål lengde i cm fra handledd til tuppen av 3.finger.

Instruksjon: lag knyttnever. Strekk den ene hånden så langt ned på ryggen og den andre så langt opp på ryggen som mulig. Mål avstanden mellom hendene. Mål på begge sider.

Sikkerhetstest: Etter testen er gjennomført på en side skal utøveren ta den motsatte hånden på skulderen og løfte albuen opp. Spør etter smerte.

Scoring:

3. Avstanden mellom hendene er mindre eller den samme som lengde målt i hånden
2. Avstanden mellom hendene er innen 1 ½ gang større enn lengde målt i hånden
1. Avstanden mellom hendene er lengre enn avstanden målt i hånden
0. Smerter ved sikkerhetstest



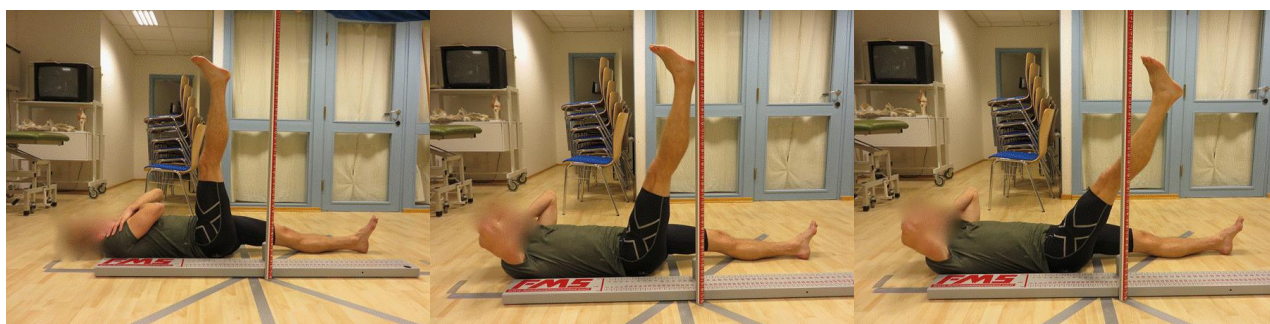
5. Active straight leg raise (ASLR)

Plasser stangen midtveis mellom SIAS og patella, planken ligger under kneet.

Instruksjon: Utøver ligger på ryggen med armene langs kroppen. Løft det ene benet strakt opp, forsøk å få låret forbi stangen og hold i ro. Senk benet langsomt ned igjen. Gjenta med det andre benet. Denne øvelsen kan gjøres opp til 3 ganger på hvert ben. Dersom utøver ikke klarer å få benet forbi stangen, plasseres stangen i nivå med malleolen på det løftede benet

Scoring:

3. Utøver klarer å gjennomføre øvelsen uten å kompensere
2. Stangen må flyttes til mellom midtpunkt på lår og patella
1. Stangen må flyttes nedenfor patella



6. Trunk stability push up

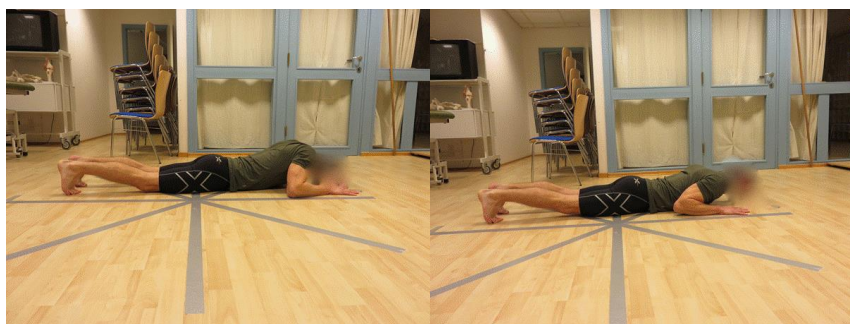
Instruksjon: Begynn i mageliggende. Hendene plasseres i skulderbreddes avstand. Menn har tomlene i høyde med pannen, kvinner har tomlene i høyde med haken. Kroppen løftes som en enhet med føttene sammen, knærne strukket og på tå Dersom dette ikke er mulig kan håndstillingen justeres, for menn ned

til høyde med haken og for kvinner ned til clavícula. Kompensering kan sees ved at utøver ikke klarer å løfte hele kroppen som en helhet og at hendene endrer posisjon.

Sikkerhetstest: Etter utøveren har gjort en repetisjon legger han/hun seg i mageliggende med albuestøtte. Utøver skyver seg opp i ryggekstensjon med bekkenet i gulvet.

Scoring:

3. Utøver gjennomfører øvelsen uten å kompensere
2. Utøver gjennomfører øvelsen, men må justere håndstilling
1. Utøver klarer ikke å gjennomføre øvelse med justert håndstilling
0. Smerter ved sikkerhetstest.



7. Rotary stability

Instruksjon: Stå i firfotstående med knær og armer plassert inntil planken. Overarmer og lår er 90 grader på overkroppen/trunkus. Dorsalfleksjon i fot. Løft samme sides arm og ben henholdsvis frem og bak, før du fører dem sammen slik at albue og kne berøres på samme side. Prøv å holde benet og armen i linje med planken. Standarm og ben skal ikke endre posisjon. Gjenta på den andre siden. Denne øvelsen kan gjøres opp til 3 ganger på hver side.

Sikkerhetstest: Utøver setter seg, med samme arm og beinstilling, tilbake på hælene. Spør etter smerter

Scoring:

3. Utøver gjennomfører øvelsen uten å kompensere
2. Utøver gjennomfører øvelsen med diagonalt mønster (motsatt arm og ben)
1. Utøver klarer ikke å gjennomføre med diagonalt mønster ut å kompensere
0. Smerter ved sikkerhetstest

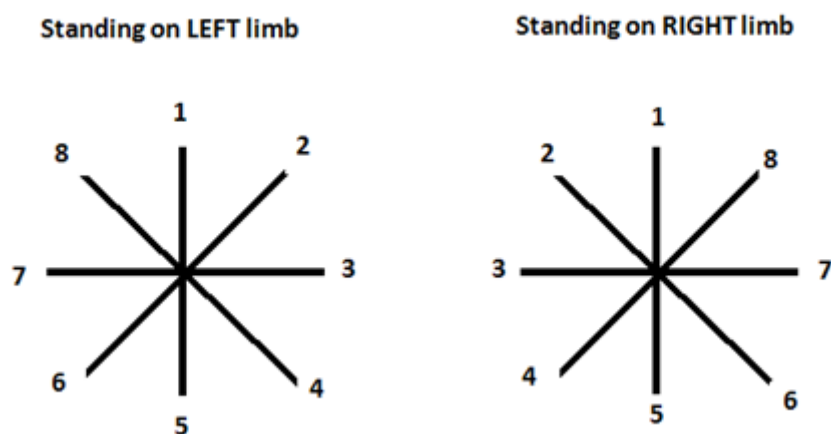


8. Star excursion balance test

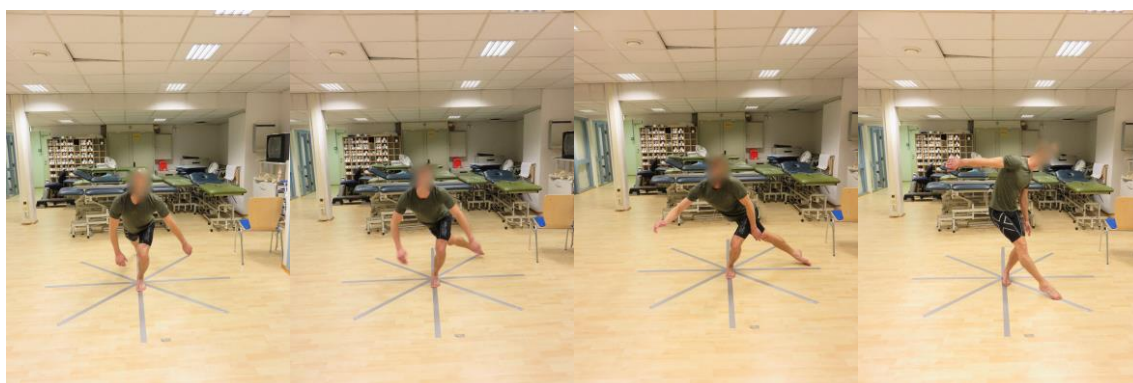
Åttearmet stjerne med 45 graders vinkel mellom armene, totalt 360 grader. Hver stråle måles til 100 cm ut fra midten.

Instruksjon: Utøver står på ett ben i midten av stjernen og strekker det andre benet så langt ut fra midten som mulig og berøre teipen uten å legge vekt på foten. Utøver begynner med strålen som går direkte fremover og fortsetter lateralt, dorsalt og medialt. Gjenta øvelsen med det andre benet. Øvelsen kan gjentas opp til 3 ganger på hvert ben. Danser får en runde med hvert ben til å prøve ut, en runde med hvert ben der han/hun blir observert, og en siste runde med hvert ben der det blir målt hvor langt han/hun klarer å strekke benet på hver stråle i stjernen

Avstand mellom høyre og venstre ben blir målt



1. Anterior 2. Anteromedial 3. Medial 4. Posteromedial 5. Posterior 6. Posterolateral 7. Lateral 8. Anterolateral



9. Lunge lean

Instruksjon: Stå med det ene benet foran det andre, la det være en hel fotlengde fra tå til hæl. Innsiden av føttene skal være på linje med hverandre og føttene peker rett frem. Bøy kneet på det fremste beinet og behold hælen i gulvet. Lår og kne skal være i linje med fotstilling. Hold ryggen rett og vertikal. Bekkenet er rettet fremover, ingen rotasjon. Behold ryggen i naturlig kurve og skift vekten over på det fremste beinet ved å flektre i hoften. Hold ca. 45 grader fleksjon i hoftene. Unngå rotasjon i bekkenet og behold kne og lår over foten. Det bakerste beinet er strakt og løftes så vidt av gulvet. Lag en rett linje fra hode til den bakre foten. Hold stillingen i 5 sekunder. Øvelsen vurderes ut fra 7 kriterier: rotasjon i bekken, endring i fotstilling på standbein, innoverrotasjon med valgusering av kne og overpronasjon, fleksjon i trunkus, hyperekstensjon i rygg, tilting/sideskift i bekken og sidebøy i trunkus og det bakre beinet er ikke i linje med overkroppen.



Scoring

Kan utøver forhindre rotasjon av bekkenet (bekkenet (blir stående rett frem)?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver forhindre valgusering av kne på standbein eller ruller ned på fotbuen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver hindre at foten dreies ut eller hælen dras inn (ruller ned på fotbuen og/eller overflekterer tær)?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver unngå å krumme ryggen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver unngå å gå i hyperekstensjon i ryggen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver hindre sidebøy av trunkus eller tilting eller sideskift av bekkenet?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver forhindre at det ikke vektbærende benet faller fra den rette linjen fra ryggen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Vedlegg 2

Scoringsskjema utvidet Functional Movement Screen**Id.....****Dato:****Deep squat test**

Score: 0 1 2 3

Hurdle step

Score: 0 1 2 3

In-line lunge

Score: 0 1 2 3

Shoulder mobility

Score: 0 1 2 3

Active straight leg raise (ASLR)

Score: 0 1 2 3

Trunk stability push up

Score: 0 1 2 3

Rotatory stability

Score: 0 1 2 3

Star excursion balance test

1	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
2	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
3	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
4	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
5	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
6	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
7	Høyre.....cm	Venstre.....cm	Forskjell:.....cm
8	Høyre.....cm	Venstre.....cm	Forskjell:.....cm

Lunge lean

Kan utøver forhindre rotasjon av bekkenet (bekkenet (blir stående rett frem)?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver forhindre valgisering av kne på standben eller ruller ned på fotbuen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver hindre at foten dreies ut eller hælen dras inn (ruller ned på fotbuen og/eller overflekterer tær)?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver unngå å krumme ryggen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver unngå å gå i hyperekstensjon i ryggen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver hindre sidebøy av trunkus eller tilting eller sideskift av bekkenet?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Kan utøver forhindre at det ikke vekt bærende benet faller fra den rette linjen fra ryggen?

Høyre: Ja..... Nei..... Venstre: Ja..... Nei.....

Vedlegg 3

Request for participation in a research project***Screening of ballet dancers of the Norwegian National Ballet*****Background and purpose**

This is a request for you to participate in a research study that intends to investigate a screening method on dancers, to discover risk factors for injuries. In this way we want to check the inter-tester reliability of the different tests and tester and to see if these tests are well suited for ballet dancers. The information retrieved in this project will result in two (2) master theses in collaboration with the study of Manual Therapy at the University of Bergen. Both testers/students have experience with ballet and/or ballet dancers, and concerning the lack of good quality studies in this area we would like to contribute with our research.

What does the study entail?

The dancers will be tested in several specific tests, administered by two testers. Some of the dancers will be tested twice to monitor the tests themselves and how reliable they are. A questionnaire of previous history of injury have to be filled out by everyone participating in the research project. The whole examination will take approximately 30 minutes. The test will be recorded on video as a baseline to see if there have been improvement on a later stage, this is not a part of the research project. The Norwegian National Ballet will use these videos taken as a baseline to compose an individualized training program. At later stages this baseline could be used as a comparison to measure the outcome of the training. The information retrieved at baseline will result in two master theses in manual therapy.

Potential advantages and disadvantages

The tests will not cause any potential danger, or discomfort. Some of the participants will be tested twice, a total of one hour of testing. This hour may cause a disruption of your regular training or rest. The tests might give information about potential risk factors for injury or potential for improvement of performance, hence there might be a beneficial factor participating in this research project.

What will happen to the samples and the information about you?

The samples and data that are registered about you will only be used in accordance with the purpose of the study as described above. All the data and samples will be processed without name, ID number or other directly recognisable type of information. A code number links you to your data and samples through a list of names.

Only authorised project personnel will have access to the list of names and be able to identify you. The test forms will be de-identified.

It will not be possible to identify you in the results of the study when these are published. Since there is a relatively small amount of dancers at the Norwegian National Ballet, publication will tell that you as a participator might be in this research project, but in any publication no results will be of such matter that anyone will have the possibility to recognize any individuals.

Voluntary participation

Participation in the study is voluntary. You can withdraw your consent to participate in the study at any time and without stating any particular reason. This will not have any consequences for your further treatment. If you wish to participate, sign the declaration of consent on the final page. If you agree to participate at this time, you may later on withdraw your consent without your treatment being affected in any way. If you later on wish to withdraw your consent or have questions concerning the study, you may contact. Roger Olsen, 905 51 676 and/or Desirée Are 907 95 290

Further information on the study can be found in Chapter A – *Further elaboration of what the study entails.*

Further information about privacy and insurance can be found in Chapter B – *Privacy, funding and insurance.*

The declaration of consent follows Chapter B.

Chapter A – Further elaboration of what the study entails

Criteria for participation:

- Professional dancer at the Norwegian National ballet

Criteria for exclusion from the project:

- Pregnancy
- Currently injured
- Surgery within the last six (6) months

In this two-part study we will conduct a screening. The main purpose of these screening tests is to find potential risk factors for injury, or factors that might hinder skill improvement. The screening consists of nine (9) tests that consist of basic movements. History of injuries will be compared with the results from the screening tests. This to see if the selected tests are able to measure that they claim to measure.

The schedule for testing will be from the end of January 2014, and till the end of March 2014. Exact day and time for testing will be worked out in a plan together with the Norwegian National ballet, in order to minimize absence from regular training. All tests will be conducted within the regular working hours, so that no one has to come in on your off time. The tests will be conducted in the office of the physiotherapist and in a neighbour office.

The tests are basic movements that are considered as movements that hold very little risk of injury, considering that all movements have risks of injury. There is a slight possibility that these 30 to 60 minutes (depending on you will be tested once or twice), might take time of regular training. The recording of previous history of injuries will take less than 5 minutes.

Compared to other groups of athletes there is very little studies conducted on dancers.

Dancers in general and especially ballet dancers are prone to injuries. In a recent review of studies on dancers and injuries 69 % of 103 published studies were found inadequate cause of poor quality. The same review showed that injuries are common among dancers, from 3 – 95 % reported injuries. Ankle/foot, hip, knee and spine injuries are common injuries among dancers. Long-term effects of injuries are unknown, but are thought to have an influence in quality of life after ended career. Prevention of injuries is therefore important in all areas. At present time there are no guidelines for preventing injuries for ballet dancers. Tests that have been used in research have been of poor quality and have shown poor reliability. We see the need for further research in mechanism of dance injuries.

Potential advantages of participating in this study could be to expose possible risk factors for injuries, and possible weaknesses that could hinder further development in ballet dancing. Findings can help individualize training programs and hence prevent injuries.

Chapter B – Privacy, funding and insurance

Privacy

Information that is registered about you is going to be processed in statistics in this research project. The de-identified information retrieved from this project will be stored safely, and the only ones having access to this information are the two persons that are conducting the

research and the tutor for these two master theses.

The screening exercises will be scored after a pre-set scale. The injury registration will be examined after the screening tests are conducted. Data will be collected and restored according to Norwegian law and the Helsinki declaration to insure that privacy is obtained. Anyone who has access to this data is bound to secrecy. The two researchers responsible for collecting the data will conduct processing of the data.

Releasing material and data to other parties

If you agree to participate in the study, you also consent to samples and de-identified data being released for publication in an article in later stages. Participating in this study you accept that the medical staff at the Norwegian National Ballet and the ballet director and assistant ballet director can watch the video recording of the test performance.

Right to access and right to delete your data and samples

If you agree to participate in the study, you are entitled to have access to what information is registered about you. You are further entitled to correct any mistakes in the information we have registered. If you withdraw from the study, you are entitled to demand that the collected samples and data are deleted, unless the data have already been incorporated in analyses or used in scientific publications.

Funding

This research project is not funded by anyone.

Insurance

The research project takes place within working hours, so work insurance will cover any potential injury.

Information about the outcome of the study

As a participator in this study you are entitled to receive information about the outcome/result of this study. And we as researchers will provide the results at the end of the project (stipulated end of the project is by the end of 2014).

Consent for participation in the study

I am willing to participate in the study.

(Signed by the project participant, date)

Proxy consent when this is warranted, either in addition to or in place of the participant's consent.

(Signed by representative, date)

I confirm that I have given information about the study.

(Signed, role in the study, date)

Vedlegg 4



Region:	Saksbehandler:	Telefon:	Vår dato:	Vår referanse:
REK vest	Arne Galbu	55078498	10.04.2014	2014/0337/REK vest
			Deres dato:	Deres referanse:
			04.03.2014	
Vår referanse må oppgis ved alle henvendelser				

[REDACTED]
Seksjon for fysioterapivitenskap
Universitetet i Bergen

2014/337 Screening av skadeforkomst hos ballettdansere ved Den Norske Opera/Den Norske Nasjonalballetten

Forskningsansvarlig: Universitetet i Bergen
Prosjektleder: [REDACTED]

Vi viser til søknad om forhåndsgodkjenning av ovennevnte forskningsprosjekt. Søknaden ble behandlet av Regional komité for medisinsk og helsefaglig forskningsetikk (REK vest) i møtet 27.03.2014. Vurderingen er gjort med hjemmel i helseforskningsloven (hfl.) § 10, jf. forskningsetikklovens § 4.

Prosjektomtale

Utgangspunktet for denne studien er at ballettdansere er utsatt for belastningsskader, men forskningslitteraturen til nå har ikke klart å vise årsakssammenhengene. I denne studien ønsker man at en screening av danserne skal bidra til økt kunnskap om skader hos ballettdansere. En legger vekt på at det er funksjonell screening som opplyser å være utføring av oppgaver som er idrettsspesifikk eller arbeidsspesifikk. Deltakerne vil være ballettdansere ved Nasjonalballetten, i alt 62 personer. Deltakelse innebærer gjennomgang av 9 funksjonelle screeningøvelser. Testene vil bli filmet. Videre skal de fylle ut skaderregistreringsskjema.

Vurdering

Søknad/protokoll er relevant og komiteen mener studien er organisert på en måte som gjør det mulig å besvare forskningsspørsmålene.

Frivillighet/Rekruttering/forespørsel om deltakelse

I punkt 2d sies det at «Alle profesjonelle dansere ved den Norske Nasjonalballetten er inkludert i studien». Da deltakelse i forskningsprosjekt skal være basert på frivillighet forutsetter vi at det menes alle blir spurt om å delta, men at det er frivillig å delta, altså at det ikke inngår som noe pliktarbeid å delta. Denne frivilligheten må være reell, også på den måten at den ikke har noen betydning for ens ansettelsesforhold om man ikke deltar, noe som det skal være opplyst om i forespørsel til deltakerne.

Et hovedprinsipp i medisinsk/helsefaglig forskning blant grupper av arbeidstakere fra spesielle bedrifter, er at der er vannede skott mellom forsker og bedrift/arbeidsgiver. Dette innebærer at ingen helseopplysninger eller individuelle opplysninger som fremkommer gjennom å delta i prosjektet skal tilfalle bedrift/arbeidsgiver, i dette tilfellet Den Norske Nasjonalballetten.

Bruk av Videopptak/datasikkerhet

Besøksadresse: Arntsen Helsehus (A111), Tverfley Nord, 2 etasje, Rom 201, Haukelandveien 28	Telefon: 55075000 E-post: rek-vest@uhb.no Web: http://helseforskning.etikk.uib.no/	All post og e-post som inngår i saksbehandlingen, bes adressert til REK vest og ikke til enkelte personer	Kindly address all mail and e-mails to the Regional Ethics Committee, REK vest, not to individual staff
--	--	---	---

I Forespørselen til deltakerne under avsnitt What does the study entail? opplyses det om testing av dansere og om spørreskjema som involverer spørsmål om evt. tidligere skader. Det sies videre at "The test will be recorded on video as a baseline to see if there have been improvement on a later stage, this is not a part of the research project. The Norwegian National Ballet will use these video taken as a baseline to compose an individualized training program. I punkt 5 h om "Tidsramme" sies det videre at Alle data eies av UiB og den Norske Nasjonalballetten for etterprøving og videre oppfølgingsmuligheter av danserne etter kunnskapen fra screeningen som blir gjennomført.

For REK anses videoopptak av testene som er tatt i forbindelse med et forskningsprosjekt etter helseforskningsloven som registrering av helseopplysninger og må behandles som dette. Hovedreglene (§ 38) er at helseopplysninger ikke skal oppbevares lengre enn det som er nødvendig for å gjennomføre prosjektet. Dette tilsier at dataene, inklusiv videoopptak, i utgangspunktet må slettes etter prosjektlutt. REK har ikke hjemmel for å vurdere eller godkjenne videre oppbevaring og bruk ut over det som gjelder for prosjektet. Dette er et forhold vi antar er regulert i helseregisterloven og vi ber om at spørsmålet tas opp med Datatilsynet eller stedlig personvernombud.

For øvrig må koblingsnøkkel og aidentifiserte opplysninger oppbevares separat.

Vilkår

- Ingen helseopplysninger eller individuelle opplysninger som fremkommer gjennom å delta i prosjektet skal tilfalle arbeidsgiver
- Alle innsamlede opplysninger, inklusiv videoopptak, må slettes etter prosjektlutt. Søknad om bruk av videoopptak/helseopplysninger ut over det som gjelder for dette prosjektet, må tas opp med datatilsynet, evt stedlig personvernombud
- Koblingsnøkkel og aidentifiserte opplysninger, må oppbevares separat.

Vedtak

REK vest godkjenner prosjektet på betingelse av at ovennevnte vilkår tas til følge.

Sluttmelding og søknad om prosjektendring

Prosjektleder skal sende sluttmelding til REK vest på eget skjema senest 31.12.2014, jf. hfl.

12. Prosjektleder skal sende søknad om prosjektendring til REK vest dersom det skal gjøres vesentlige endringer i forhold til de opplysninger som er gitt i søknaden, jf. hfl. § 11.

Klagendgang

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

Med vennlig hilsen

Ansgar Berg
komiteleder, professor

Arne Salbu
rådgiver