Joar Harøy

Inter-rater reliability of isometric hip abductor strength, 1RM leg press, hamstrings flexibility and the Star Excursion Balance Test in elite female football- and handball players.

- A methodological study

Master thesis in Sports Physiotherapy
Department of Sports Medicine
Norwegian School of Sport Sciences, 2013

Abstract

Introduction: Strength-, flexibility- and balance measures are tools used to examine risk factors for injuries in a clinical setting. In an on-going prospective cohort study at the Oslo Sports Trauma Research Center aimed at investigating risk factors for ACL injuries in elite female football- and handball players several of these tests are used. The different tests have shown high inter-rater reliability, but the studies are rarely conducted on elite female football- or handball players. The aim of this study was to investigate the inter-rater reliability of tests assessing the isometric hip abductor strength, 1RM in a leg press machine, the hamstrings flexibility and the Star Excursion Balance Test evaluating risk factors for ACL injury in elite female football- and handball players.

Method: We included 42 elite football and handball players (21.5±3.8 years). The players participated in test-retest measurements, within three days to seven weeks. Different testers examined the players on each test session. The inter-rater reliability was calculated with ICC, SEM and SEM %.

Results: The isometric hip abductor strength test showed an ICC value of 0.67-0.69, the 1RM leg press test showed an ICC value of 0.78-0.83, the hamstrings flexibility test showed an ICC value of 0.71-0.72 and the Star Excursion Balance Test an ICC value of 0.81-0.90. SEM % varied from 5.7-8.8 % for the strength measures, 4.7-4.8 % for the hamstrings flexibility measures and 3.0-3.5 % for the Star Excursion Balance Test.

Conclusion: The isometric hip abductor strength test, the 1RM leg press test, the hamstring flexibility test and the Star Excursion Balance Test proved to be reliable on our cohort of female football- and handball players. The ICC score varied from moderate to excellent, with SEBT and leg press showing excellent inter-rater reliability. The isometric hip abductor strength test and the hamstrings flexibility test showed moderate to good inter-rater reliability.

Acronyms

1RM	One repetition maximum
ACL	Anterior cruciate ligament
AL	Anterior lateral
CI	Confidence interval
CM	Centimetre
CV	Coefficient of variation
ICC	Interclass correlation coefficient
KG	Kilogram
MD	Mean difference
ML	Medio lateral
M	Meter
PL	Posterior lateral
RS	Relative strength
ROM	Range of motion
SD	Standard deviation
SEM	Standard error of measurement
SEBT	Star Excursion Balance Test

Table of content

Preface

I am grateful to have been given the opportunity to spend a year working on my master thesis at Norwegian School of Sports Sciences studying my favourite sport; football.

I will especially thank my supervisors Thor Einar Andersen and Agnethe Nilstad, which made this thesis possible. Thank you for always been available and had plenty of time for guidance during the process. Your knowledge, help and advices have been essential.

Many thanks for all the testers who helped with the data collection and stayed positive during long days of testing. Whiteout you this project would have been impossible. Also many thanks to the clubs and players involved in the testing for being positive and took their time to participate in the two test sessions.

Thanks to Helene Kværne Vileid, which also conducted her master thesis investigating inter-rater reliability on some of the other tests in the cohort. Thanks for a brilliant collaboration and the necessary input on language.

I would like to thank Marie for being such a supportive, kind and patient person.

Last but not least my family, which deserve a big thank you for all your support.

Joar Harøy
Oslo, May 2013

1.0 Introduction

An anterior cruciate ligament (ACL) injury represents a major area within sports medicine as they cause the most time lost from competition in football (Myklebust and Bahr 2005; Lohmander, Englund et al. 2007; Meunier, Odensten et al. 2007; Engebretsen and Bahr 2009). It causes morbidity, long disability time and a devastating influence on activity levels and quality of life, high economic costs and potential long-term effects such as early development of osteoarthritis (Elliot, Goldberg et al. 2010). An ACL injury is frequently represented in pivoting- and team sports with ball, which includes powerful deceleration, rapid changes of directions, cutting manoeuvres and landing from jumps (Griffin, Albohm et al. 2006; Renstrom, Ljungqvist et al. 2008; Brophy, Silvers et al. 2010). The highest injury risk is among top-level female athletes who compete in team handball and football (Engebretsen and Bahr 2009). Reviews have shown an ACL injury rate of 0.28 ACL injuries per 1000 athlete-exposures in female football (Renstrom, Ljungqvist et al. 2008) and 0.31 ACL injuries per 1000 playing hours in female elite handball (Myklebust, Maehlum et al. 1998; Prodromos, Han et al. 2007).

Female athletes who participate in pivoting and jumping sports have a four to six times higher risk of suffering ACL injuries compared to male athletes in the same sports (Ferretti, Papandrea et al. 1992; Arendt and Dick 1995; Myklebust, Maehlum et al. 1997). A more recent review concluded that female football players seem to have a two to three times higher ACL injury risk compared to male players (Walden, Hagglund et al. 2011). Studies, both in Norwegian elite handball (Myklebust, Maehlum et al. 1997; Myklebust, Maehlum et al. 1998) and in German (Faude, Junge et al. 2006) and Norwegian elite football (Tegnander, Olsen et al. 2008) showed that 5-10% of the players sustain an ACL injury each season. In a typical league with 12-16 teams that equals to one entire team each season.

Understanding the risk factors and the mechanisms of injury is important in order to identify those at higher risk (Bahr and Krosshaug 2005). Studies have shown that ACL injuries often occur when the players is in no physical contact with other players, thus termed non-contact ACL injuries (Boden, Dean et al. 2000; Fauno and Wulff Jakobsen 2006). Additionally the players often perform a cutting manoeuvre, deceleration or a landing from a jump at the time of injury (Arendt and Dick 1995; Krosshaug, Slauterbeck et al. 2007). Researchers have proposed and debated several risk factors for ACL injury (Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009a). However, it lacks in many cases results supporting the theories (Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009a). Several programs for preventing ACL injuries among female athletes have been designed (Silvers and Mandelbaum 2007; Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009b; Brophy, Silvers et al. 2010; Shultz, Schmitz et al. 2010). Most of these studies presents multifactorial training strategies with a focus on promoting favourable positioning of the lower limbs, in cutting manoeuvres and in landings, to eliminate specific moves (Silvers and Mandelbaum 2007; Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009b; Brophy, Silvers et al. 2010; Shultz, Schmitz et al. 2010). Although some of the studies aiming to prevent ACL injuries in female athletes have shown promising results, more knowledge on injury mechanism and preventing strategies is needed (Renstrom, Ljungqvist et al. 2008; Brophy, Silvers et al. 2010; Shultz, Schmitz et al. 2010). By targeting specifically developed methods for preventing the injury, at players in risk of injury, we hopefully can increase compliance rate (Renstrom, Ljungqvist et al. 2008).

It is important to include several different tests to detect different aspects and consequences of an ACL injury (Bent, Wright et al. 2009; Wright 2009). The different studies aiming to identity risk factors for ACL injuries among female athletes use numerous different tests to examine the attributes of the athlete (Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009a). In the on-going prospective cohort study at the Oslo Sports Trauma Research Center aimed at investigating risk factors for ACL injuries in elite female football- and handball players the different tests are; anthropometric measures, 3D motion analysis, measures of hip anteversion, genu recurvatum, knee and generalized joint laxity, hamstrings flexibility, static and dynamic balance, Star Excursion Balance Test, isometric hip abduction strength, leg extensor strength, isokinetic hamstring/quadriceps strength and subjective assessment of knee and hip control in single leg squats and drop jumps.

To trust the results from the different tests we are dependent that the tests we use possess reliability, validity and have the ability to detect changes over time (Portney and Watkins 2009; Thomas, Silverman et al. 2011). The reliability of several of the tests used in the on-going cohort has already been examined. Some of these studies have also showed high reliability on isokinetic knee extension and flexion, isometric hip abduction strength, the hamstrings flexibility and the SEBT (Shultz, Nguyen et al. 2006; Sole, Hamren et al. 2007; Plisky, Gorman et al. 2009; Munro and Herrington 2010; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012). However most of the studies assessing reliability of the different tests are conducted with small sample sizes (n=< 20) and on sample group with different characteristics compared to the football- and handball players included in the on-going cohort examining risk factors for ACL injuries.

1.1 Objective

This master thesis aims to investigate the inter-rater reliability of tests assessing the isometric hip abductor strength, 1RM in a leg press machine, the hamstrings flexibility and the Star Excursion Balance Test evaluating risk factors for ACL injury in elite female football- and handball players.

1.2 Hypotheses

The following hypotheses are based on the already existing findings, as well as a reliability pilot study done in 2007 when starting the data collection.

- The isometric hip abductor strength test will show good to excellent inter-rater reliability
- 1RM leg press will show excellent inter-rater reliability
- The Hamstrings flexibility test will show good inter-rater reliability
- The Star Excursion Balance Test will show excellent inter-rater reliability

2.0 Theory

2.1 Anatomy of the knee joint

The knee joint (shown in figure 1) is the connection of the two longest levers, femur and tibia (Dahl and Rinvik 2010). It is built for motion and bearing the weight of the body (Dahl and Rinvik 2010). Therefore it possesses both subtle movement and stability, which the construction of the knee joint reflects. The knee joint is a synovial joint consisting of three compartments. The medial and the lateral compartments, where the femur meets the tibia, and the patellofemoral joint where patella articulates with the femur. The medial and lateral compartments are the areas of weight transmission and are separated by the medial and lateral meniscus. Between the two femur condyles is a narrow pass, the intercondylar fossa (notch), where the cruciate ligaments passes through (Petersen and Zantop 2007). The tibia condyles are shaped as two shallow indentations, separated by an intercondylar elevation (tibial slope), which have a slight posterior inclination (Petersen and Zantop 2007). The knee joint is stabilized through a strong joint capsule, medial- and lateral collateral ligaments and anterior- and posterior cruciate ligaments (Dahl and Rinvik 2010). The muscles surrounding the knee joint controls the movement and additionally provide stability during lower extremity movements (Dahl and Rinvik 2010).

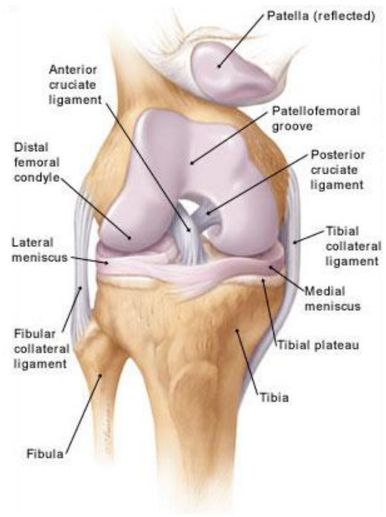


Figure 1: *The anatomy of the knee. (Medchrome.com)*

2.1.1 Knee joint function

The knee consists of six different degrees of freedom; flexion and extension, varus and valgus, internal and external rotation, anterior and posterior translation, medial and lateral translation and traction and compression (Goodfellow and O'Connor 1978). The knee joint function consists primarily of flexion, extension, internal and external rotation (Dahl and Rinvik 2010) and is most stable during maximal extension through tightening of the ligaments (Duthon, Barea et al. 2006).

During flexion the knee joint is unstable for side movement, and rotation can occur (Duthon, Barea et al. 2006).

2.1.2 Anterior cruciate ligament

The ACL is a structure of compact connective tissue and have a slightly spiral shape because of its vertical path and attachment to the femur (Odensten and Gillquist 1985; Giuliani, Kilcoyne et al. 2009). The ACL is composed of an anteromedial- and posterolateral bundle and originates from medial and anterior aspect of the tibial plateau and runs superiorly, laterally, and posteriorly through the intercondylar fossa toward its insertion on the lateral femoral condyle (Giuliani, Kilcoyne et al. 2009). The anteromedial- and posterolateral bundles contains different properties and work as synergist to optimise the ACLs function through the entire range of motion of the knee (Xerogeanes, Takeda et al. 1995). The ACL prevents forward movement of the tibia in relation to the femur, and controls rotational movement of the tibia under the femur (Duthon, Barea et al. 2006; Giuliani, Kilcoyne et al. 2009). The ACL is essential for control in pivoting movements and with a dysfunctional ACL the tibia may rotate under the femur in an anterolateral direction (Brukner and Khan 2012).

2.2 Injury mechanisms for non-contact ACL injuries

Research so far suggests a series of elements determining the mechanisms of ACL injuries. First of all, most ACL injuries (70-84 %) happen when the player is in no physical contact with other players, thus termed non-contact ACL injuries (Boden, Dean et al. 2000; Fauno and Wulff Jakobsen 2006). The mechanisms involved in most cases of non-contact ACL injuries is when the player performs a sudden change of direction or a cutting manoeuvre when decelerating, or during landing from a jump with the knee close to full extension with the foot planted (Boden, Dean et al. 2000; Olsen, Myklebust et al. 2004; Krosshaug, Slauterbeck et al. 2007; Koga, Nakamae et al. 2010). Additional mechanisms involved are knee valgus, internal rotation of the hip and external rotation of the tibia (Boden, Dean et al. 2000; Olsen, Myklebust et al. 2004; Krosshaug, Slauterbeck et al. 2007). The greatest strain on the ACL is produced by a sum of forces, however, as an isolated force the anterior translation force is the force mostly associated with ACL injuries, especially at around 20-30 degrees of flexion (Boden, Dean et al. 2000).

2.3 Risk factors for ACL injuries among female football- and handball players

In the following section several possible risk factors for ACL injuries will be presented. In this study the main focus will be on intrinsic risk factors. Extrinsic risk factors will only be presented briefly. Most research has studied isolated risk factors, therefore the consequence of clustering risk factors is unclear and ACL injuries are most probably a result of several factors appearing together as a multi-factorial cause (Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009a). A summary of the intrinsic risk factors is presented in table 1.

2.3.1 Extrinsic risk factors for ACL injuries

Little is known about the extrinsic risk factors such as footwear, playing surface, protective equipment or meteorological conditions (Renstrom, Ljungqvist et al. 2008) and the existing evidence is mainly based on American Football or Australian Football (Scranton, Whitesel et al. 1997; Orchard, Seward et al. 1999; Orchard, Chivers et al. 2005). Studies regarding extrinsic risk factors also may be limited because of several potential confounding factors (Alentorn-Geli, Myer et al. 2009a).

A higher rate of ACL tears during dry conditions on natural grass has been proposed as a potential risk factor (Heidt, Dormer et al. 1996). It could be explained with increased friction from the shoe-surface interface (Heidt, Dormer et al. 1996). Artificial turf has been presented as a another possible risk factor, but several studies have shown no difference in incidence of ACL injuries comparing artificial turf versus natural grass (Ekstrand, Timpka et al. 2006; Fuller, Dick et al. 2007; Fuller, Dick et al. 2007; Steffen, Andersen et al. 2007). In handball however, female players had a 2.4 times higher risk of ACL injury on artificial floors than on wooden floors (Olsen, Myklebust et al. 2003). Further Myklebust et al. found that female elite handball players are at higher risk of suffering an ACL injury during competition than during training (Myklebust, Engebretsen et al. 2003).

2.3.2 Intrinsic risk factors for ACL injuries

Several potential intrinsic risk factors have been proposed for ACL injuries, but there is lacking evidence linking these factors to the actual ACL injuries (Griffin, Albohm et al. 2006; Renstrom, Ljungqvist et al. 2008; Alentorn-Geli, Myer et al. 2009a). In this study the intrinsic risk factors are dived into non-modifiable- and modifiable.

Non-modifiable risk factors for ACL injuries

Non-modifiable risk factors are anatomical, hormonal, developmental or other risk factors, which are not directly modifiable to decrease risk of ACL injury (Alentorn-Geli, Myer et al. 2009a; Alentorn-Geli, Myer et al. 2009b).

Q-Angle

An increased Q-angle may give the quadriceps muscle a more lateral pull and place the knee at higher risk to static and dynamic valgus stress (Heiderscheit, Hamill et al. 2000; Mizuno, Kumagai et al. 2001). Studies have shown that females have an increased Q-angle compared to males (Conley, Rosenberg et al. 2007) and a study on female basketball players shown an increased Q-angle on players with knee injuries compared to non-injured players (Shambaugh, Klein et al. 1991). However, a Swedish study on female football players from second and third division was unable to identify any association with increased Q-angle and leg injuries (Soderman, Alfredson et al. 2001). The role of increased Q-angle as a risk factor for anterior cruciate ligament is still unclear (Alentorn-Geli, Myer et al. 2009a).

Joint laxity

The role of general joint laxity and anterior posterior knee joint laxity as a risk factor is unclear, but it seems that an increased laxity could alter dynamic lower extremity motions and has been related to increased functional valgus collapse (Ford, Myer et al. 2003; Hewett, Myer et al. 2004). There are only few studies that have examined this risk factor specific on female football or handball players. However, American Military cadets showed a greater risk of injury for cadets with increased general joint laxity and/or higher anterior-posterior knee joint laxity (Uhorchak, Scoville et al. 2003). Both the general joint laxity and the anterior-posterior knee joint laxity was significantly higher for females compared to males (Uhorchak, Scoville et al. 2003). The Swedish study among female second and third division football players identified a significantly higher risk for leg injuries for players with increased general joint laxity and knee hyper extension (Soderman, Alfredson et al. 2001). This study however, focused on leg injuries in general, not ACL injuries only. A more recent study among female football and basketball players identified a higher risk of non-contact ACL injuries among players with increased general joint laxity and anterior-posterior knee joint flexibility (Myer, Ford et al. 2008).

Tibial slope

An increased posterior tibial slope places the tibia more anterior to the femur during quadriceps contraction, which may create an increased strain on the ACL (Liu and Maitland 2003; Fening, Kovacic et al. 2008). Several studies have indicated that ACL-injured persons have a greater posterior tibial slope compared with controls (Brandon, Haynes et al. 2006; Todd, Lalliss et al. 2010; Terauchi, Hatayama et al. 2011). However there were no difference between males and females with ACL deficiency (Brandon, Haynes et al. 2006; Todd, Lalliss et al. 2010; Terauchi, Hatayama et al. 2011). Another recent study reported a significant steeper posterior tibial slope in females with ACL injury compared to ACL injured males (Hohmann, Bryant et al. 2011). None of these studies are conducted on football or handball players, and therefore it is uncertain whether these results were valid for these groups.

Intercondylar notch and ACL size

The increased risk for ACL tear in subjects with small notch width is not fully understood (Dienst, Schneider et al. 2007). It has been suggested that an impingement of the ACL at the anterior and posterior roof of the notch may occur during tibial external rotation and abduction (Dienst, Schneider et al. 2007). Studies among college- and high school athletes have shown that a smaller intercondylar notch width is a risk factor for ACL injuries (Souryal and Freeman 1993; LaPrade and Burnett 1994). However it was unable to detect any differences among males and females (LaPrade and Burnett 1994). On the other hand Schickendantz et al. and Teitz et al. did not relate narrow intercondylar notch width as a risk factor for ACL injury (Schickendantz and Weiker 1993; Teitz, Lind et al. 1997; Dienst, Schneider et al. 2007). Chandrashekar et al. have shown in cadaver studies that ACL in women were smaller in length, cross-sectional area, volume and mass when compared to men (Chandrashekar, Slauterbeck et al. 2005).

Menstrual cycle

There are indications that the menstrual cycle is a potential factor for non-contact ACL injuries (Alentorn-Geli, Myer et al. 2009a; Sutton and Bullock 2013). However, conflicting results are reported. Some studies report that most ACL injuries occur in the follicular phase (day 0-9 of the cycle) (Arendt, Bershadsky et al. 2002; Slauterbeck, Fuzie et al. 2002; Myklebust, Engebretsen et al. 2003; Ruedl, Ploner et al. 2009), some around ovulation (day 10-14) (Wojtys, Huston et al. 2002; Adachi, Nawata et al. 2008) and one in the luteal phase (day 15-28) (Moller Nielsen and Hammar 1991). Despite conflicting data authors suggest that the female athlete may be more predisposal to ACL injuries in the pre-ovulatory phase of the menstrual cycle (Griffin, Albohm et al. 2006; Hewett, Zazulak et al. 2007).

Genetic predispositions

Some studies have investigated various genetic factors as potential risk factor for ACL injuries. These studies often have reduced control on confounding factors or in example recall bias (Sutton and Bullock 2013). Posthumus et al. showed an underrepresentation of COL5A1, a gene that encodes the collagen cartilage, among females with an ACL injury (Posthumus, September et al. 2009).

Previous injury

Several studies have showed that previous ACL injury and reconstruction of the ACL ligament is a potential risk factor for ACL injury (Orchard, Seward et al. 2001; Walden, Hagglund et al. 2006; Shelbourne, Gray et al. 2009). Previous injury increases the risk of ACL injury in both re-injury of the graft and injuries in the contralateral knee (Orchard, Seward et al. 2001; Walden, Hagglund et al. 2006; Shelbourne, Gray et al. 2009). Orchard et al. showed a significant increased risk of a subsequent non-contact ACL injury in both reconstructed and contralateral knee (Orchard, Seward et al. 2001). Females also have a significant increased risk of re-injury compared to males, and the injury incidence is associated with younger age and higher activity level (Shelbourne, Gray et al. 2009). Kramer et al. also presented that athletes with a history of ACL injury were more likely to have a prior ankle sprain compared with those who had no history or ankle sprain (Kramer, Denegar et al. 2007).

Modifiable risk factors

Some biomechanical and neuromuscular factors, such as muscle strength, power and activation patterns are considered as modifiable, as they can be altered through specific training (Alentorn-Geli, Myer et al. 2009a; Myer, Brent et al. 2011).

Hamstrings-to-quadriceps ratio and recruitment

Female athletes often demonstrate quadriceps dominance, which is an imbalance between knee extensors and flexor strength, recruitment and coordination (Alentorn-Geli, Myer et al. 2009a; Sutton and Bullock 2013). Compared to males female athletes who suffered an ACL injury had decreased hamstrings strength, but similar quadriceps strength (Myer, Ford et al. 2009). Studies have also shown that females have higher quadriceps- to hamstrings recruitment (Hewett, Stroupe et al. 1996; Wojtys, Huston et al. 1996; Malinzak, Colby et al. 2001). This reduced hamstrings strength and recruitment could increase the risk of an ACL injury especially in landings from a jump. During landings from jumps, an anterior drawer force is imparted to the tibia as the quadriceps muscle is contracted in an attempt to prevent knee from collapsing (McNair, Marshall et al. 1990; Cowling and Steele 2001). This contraction exacerbates anterior tibial translation, which the ACL attempts to restrain (Cowling and Steele 2001). As an antagonist to the quadriceps muscle, the hamstrings play an important role in stabilizing the knee joint during landing (Hewett, Myer et al. 2005). A contraction of the hamstrings in landings from jumps can impart a posterior tibial drawer force, acting as a synergist to the ACL (Cowling and Steele 2001; Withrow, Huston et al. 2006; Sell, Ferris et al. 2007)

Hip abduction strength

Increased adduction movement in the hip during landing may place an increased valgus stress in the knee joint (Ford, Myer et al. 2003; Hewett, Myer et al. 2005), which is already mentioned as a risk factor for ACL injury (Hewett, Myer et al. 2005). Studies have shown that female football, basketball and volleyball players have an increased hip adduction in landings compared to male athletes in the same sport (Chappell, Creighton et al. 2007). In healthy subjects, a decreased maximal hip abductor strength (normalized to body weight) has been found among females compared to males (Jacobs, Uhl et al. 2007). They also showed a greater correlation between hip abductor strength and knee landing kinematics among females compared to males (Jacobs, Uhl et al. 2007). Decreased hip abductor strength itself is not shown to be a risk factor for ACL injury, but low muscle strength may be an underlying factor for increased knee valgus (Hewett, Myer et al. 2005).

Muscular fatigue

Since the muscles contribute to joint stability, muscular fatigue has been proposed as a potential risk factor for ACL injury, possibly because fatigued muscles are less able to absorb energy (Loudon, Jenkins et al. 1996). In fatigued conditions both female and males showed decreased lower limb control during landings (Chappell, Herman et al. 2005; McLean, Fellin et al. 2007). The decreased lower limb control could increase knee internal rotation during impact-force absorption (Nyland, Caborn et al. 1999; Chappell, Herman et al. 2005; McLean, Fellin et al. 2007).

Muscular fatigue may increase the risk of a non-contact ACL injury by altering the neuromuscular control, but may not in itself be an isolated risk factor (Alentorn-Geli, Myer et al. 2009a).

Biomechanical risk factors

Hewett et al. (2005) found that high dynamic knee valgus could predict a future ACL injury. Movement analyses of pre-season jump-landing tasks showed that players who sustained an ACL-injury during the intervention period had significant greater knee abduction, both at initial contact and at maximum displacement (Hewett, Myer et al. 2005).

In the sagittal plane, studies have shown that the more energy that is absorbed the more the hip, knee, trunk and ankle are flexed during landing tasks (Hewett 2000; Blackburn and Padua 2008). Compared to males, female football players demonstrated decreased hip and knee flexion during landing (Yu, Lin et al. 2006; Chappell, Creighton et al. 2007). It is believed that a decreased hip- and knee flexion angle will increase the risk of non-contact injury because a greater impact force will be transferred to the knee (Hewett, Stroupe et al. 1996).

I the transverse plane there have been found greater hip internal rotation and lower gluteal electromyography activity in female football, basketball and volleyball players compared to males (Lephart, Ferris et al. 2002; Zazulak, Ponce et al. 2005). The increased hip rotation in cutting manoeuvres could alter knee rotation and indirectly increase the risk of ACL injury (Besier, Lloyd et al. 2001).

Table 1: *Summary of intrinsic risk factors related to increased risk for non-contact ACL injury*

	Non-modifiable risk factor	**Author**
Anatomical	Quadriceps angle	Heiderscheit, Hamill et al. 2000; Mizuno, Kumagai et al. 2001
	General- and knee joint laxity	Ford, Myer et al. 2003; Hewett, Myer et al. 2004; Myer, Ford et al. 2008
	Tibial slope	Brandon, Haynes et al. 2006; Hohmann, Bryant et al. 2011; Todd, Lalliss et al. 2010; Terauchi, Hatayama et al. 2011
	Intercondylar tibial notch and ACL size	Dienst, Schneider et al. 2007; LaPrade and Burnett 1994; Souryal and Freeman 1993
Hormonal and developmental	Menstrual cycle	Adachi, Nawata et al. 2008; Arendt, Bershadsky et al. 2002; Griffin, Albohm et al. 2006; Hewett, Zazulak et al. 2007; Moller Nielsen and Hammar 1991; Myklebust, Engebretsen et al. 2003; Ruedl, Ploner et al. 2009; Slauterbeck, Fuzie et al. 2002; Wojtys, Huston et al. 2002
	Genetic predisposing	Posthumus, September et al. 2009; Sutton and Bullock 2013
Other	Previous ACL injury	Kramer, Denegar et al. 2007; Orchard, Seward et al. 2001; Shelbourne, Gray et al. 2009; Walden, Hagglund et al. 2006
	Modifiable risk factor	
Neuromuscular	Hamstrings/quadriceps ratio	Myer, Ford et al. 2009
	Hamstrings recruitment	Hewett, Stroupe et al. 1996; Malinzak, Colby et al. 2001; Wojtys, Huston et al. 1996
	Hip abductor strength	Ford, Myer et al. 2003; Hewett, Myer et al. 2005
	Muscular fatigue	Loudon, Jenkins et al. 1996
Biomechanical	Knee valgus	Hewett, Myer et al. 2005
	Hip, knee and ankle flexion	Blackburn and Padua 2008; Hewett 2000

2.4 Principles of measurements

A method of measurement is a way of understanding, evaluating and differentiating characteristics of people and objects (Portney and Watkins 2009). A measurement is used as a basis for making decisions or drawing conclusions (Portney and Watkins 2009). A measurement can evaluate dysfunctions, assess and predict future rehabilitation among patients (Beyer and Magnusson 2003). In healthy individuals a method of measurement can be a tool to reveal risk factors, and help predict future injuries (Dallinga, Benjaminse et al. 2012). Results and conclusions from measurements can be used as an objective tool in communication between colleagues, across professions and as quality assurance and in science (Beyer and Magnusson 2003).

To trust the results from a test, in clinical situations and in sciences, we are dependent that the test we use possess reliability, validity and have the ability to detect changes over time (Portney and Watkins 2009; Thomas, Silverman et al. 2011). However, it varies the extent to which each of these properties the tests possesses based on the application of the results and the sample group (Portney and Watkins 2009).

2.4.1 Reliability

Reliability describes to which degree a method of measurement can be free from errors, and produce consistency or repeatability of the measurement (Streiner and Norman 2008; Portney and Watkins 2009; Thomas, Silverman et al. 2011). Measurements are rarely perfect, and some type of error often occur in any measurement (Portney and Watkins 2009). The difference between the measured value and the true value is measurement error (Streiner and Norman 2008; Portney and Watkins 2009; Thomas, Silverman et al. 2011). The reliability coefficient is the ratio of the true score variance to observed score variances (Portney and Watkins 2009; Thomas, Silverman et al. 2011). Reliability can be referred to the accepted amount of measurement error for a method of measurement (Streiner and Norman 2008; Portney and Watkins 2009; Thomas, Silverman et al. 2011).

Reliability can be described in three ways; test-retest reliability, intra-rater reliability and inter-rater reliability (Portney and Watkins 2009). *Test-retest reliability* assessment is used to establish if an instrument is capable of measuring a variable with consistency (Streiner and Norman 2008; Portney and Watkins 2009). In a test-retest study a sample group is performing identical testing in to separate occasions, with all testing conditions as constant as possible (Portney and Watkins 2009).

Intra-rater reliability refers to the stability of data recorded by one individual across two or more trials (Portney and Watkins 2009). By establishing the intra-rater reliability we can describe in which degree the same person is able to measure the same result in two or more occasions (Portney and Watkins 2009). In situations where the rater´s skill is essential to the accuracy of the measurement intra-rater reliability is the same as test-retest reliability (Portney and Watkins 2009).

Inter-rater reliability is the variation between two or more raters who measure the same group of subjects (Portney and Watkins 2009). We want as good agreement as possible between the measurement results of two or more testers (Streiner and Norman 2008; Portney and Watkins 2009). Inter-rater reliability is desirably measured when two- or more raters are able to simultaneously measure a single trial. However in many measurements, simultaneously scoring/measuring is not possible due to the need of interaction from rater (Portney and Watkins 2009). By establishing inter-rater reliability the outcome of the study is more generalizable and we can assume that other raters, with the equal characteristics, are able to measure the subject´s true score (Portney and Watkins 2009).

2.4.2 Relative- and absolute reliability

It is proposed that a comprehensive set of statistical methods is required to address the reliability of measurements (Lexell and Downham 2005; Portney and Watkins 2009). It is suggested to present both relative- and absolute reliability. Relative reliability assesses the consistency a group of subjects on two separate occasions as a reliability coefficient, while absolute reliability describes the variation between two occasions in the measured unit (Lexell and Downham 2005; Portney and Watkins 2009).

Relative reliability

The intraclass correlation coefficient (ICC) is the most common used reliability coefficient for rater reliability (Lexell and Downham 2005; Portney and Watkins 2009). The ICC reflects both the degree of correspondence and the agreement among variables (Portney and Watkins 2009). It can be used assessing reliability between two or more ratings, also on small sample sizes, and the ICC does not require the equal number of raters on each subject (Lexell and Downham 2005; Portney and Watkins 2009). The ICC is an index with a range from 0.00-1.00, where 0.00 represent absence correspondence or agreement and 1.00 represent perfect correspondence and the agreement (Portney and Watkins 2009).

There are six different types of equations for calculating the ICC, divided in three models (Shrout and Fleiss 1979; Portney and Watkins 2009). Model two, $ICC_{(2.1/k)}$ is the most commonly used equation for assessing inter-rater reliability (Shrout and Fleiss 1979; Portney and Watkins 2009). In this design each subject is assessed by the same set of raters, who is randomly chosen from a group of raters that are expected to represent the group of raters the results can be generalized to (Portney and Watkins 2009). The subjects are also randomly chosen from a larger group of population who represent the population who would receive measurements (Portney and Watkins 2009).

Absolute reliability

Absolute reliability quantifies the measurement error in the same unit as measured or as coefficient of variation (CV) expressed in percent (Lexell and Downham 2005; Portney and Watkins 2009). Absolute reliability is often presented as standard error of the measurement (SEM) and SEM % (Lexell and Downham 2005; Portney and Watkins 2009). The measurement error can be visually presented in a Bland Altman plot, which analyses the differences between pairwise scores from test and re-test for each subject (Lexell and Downham 2005;

Portney and Watkins 2009). The difference is plotted against the respective individual mean. The Bland Altman plot gives a visual presentation of systematic- and random error by assessing the direction and amount of spread around the zero-line (Lexell and Downham 2005; Portney and Watkins 2009). The SD between the measurements is used to calculate upper and lower limit of agreement, were we would expect that approximately 95 % of the different scores would fall within (Portney and Watkins 2009).

What is acceptable reliability

The definition of what is acceptable reliability must be considered in conjunction to different factors (Lexell and Downham 2005; Portney and Watkins 2009). In general we may tolerate a lower reliability for measurement that are used for description, while measurements used for decision making or diagnosis need to be as high as possible (Portney and Watkins 2009). The level of acceptance should be based on the purpose of the measurement (Streiner and Norman 2008; Portney and Watkins 2009; Thomas, Silverman et al. 2011).

As a general guideline, for the relative reliability, it is often suggested that values >0.75 are indicated as excellent reliability, values <0.50 are indicated as poor reliability and values between 0.50 and 0.75 are indicated as moderate to good reliability (Lexell and Downham 2005; Portney and Watkins 2009). For the absolute reliability the measurement error must be compared to the mean of the measurement to decide if it is acceptable or not (Lexell and Downham 2005; Portney and Watkins 2009).

2.4.3 Sources of measurement error

In every measurement some amount of error consists (Streiner and Norman 2008; Portney and Watkins 2009; Thomas, Silverman et al. 2011). The reliability reflects the amount of both systematic- and random error in this measurement (Streiner and Norman 2008; Portney and Watkins 2009). There are four sources of measurement errors: the participants, the testing, the scoring and the instrumentation (Thomas, Silverman et al. 2011). Every measurement consists of a true score and an error component (Portney and Watkins 2009; Thomas, Silverman et al. 2011). The true score represent the individual´s real score and does not contain any error (Portney and Watkins 2009; Thomas, Silverman et al. 2011). The error component is the part of an observed score that is attributed to true score (Portney and Watkins 2009; Thomas, Silverman et al. 2011).

Observed score = true score ± error component

Systematic error

Systematic errors of measurement are predictable errors of measurements, who occur in one direction, and constantly overestimating or underestimating the true score (Streiner and Norman 2008; Portney and Watkins 2009). Systematic error is constant and primarily a problem for validity, but not for reliability (Portney and Watkins 2009). A typical systematic error is problem or lack of recalibration of measurement instrument/method, rater that constantly over- or underestimates the true score or learning effect from first to second test session (Portney and Watkins 2009).

Random error

Random errors of measurement are due to chance and can affect a subject´s score in an unpredictable way from trial to trial (Streiner and Norman 2008; Portney and Watkins 2009). The errors occur from unpredictable factors such as inaccuracy, fatigue, mistakes, physical- or mental status, day-to-dag variations or lack of procedures (Portney and Watkins 2009). The amount of random error there is in a measurement affects the reliability of the measurement. The fewer random errors in a measurement, the closer it would be to the true score (Portney and Watkins 2009). By conducting enough measurements, eventually the random errors will cancel each other out, and the average value of all the measurements will be a good estimate of the true score (Portney and Watkins 2009).

2.4.4 Validity

Validity is the degree to which a test or instrument measures what it intends to measure (Portney and Watkins 2009; Thomas, Silverman et al. 2011). It can further be categorized as internal- or external validity (Streiner and Norman 2008; Portney and Watkins 2009; Thomas, Silverman et al. 2011). Internal validity refers to the design of the study and to which degree the results are representative for the sample selection and the phenomenon being examined (Vincent and Weir 2012). External validity refers to the ability to generalize the results to the population which the sample selection was taken from (Vincent and Weir 2012). Still, external and internal validity influences each other. A tightly controlled experiment increases the internal validity, but decreases the external validity and makes it difficult to generalize the results to an actual situation (Vincent and Weir 2012).

2.5 Methods of measurements for muscle strength, flexibility and balance

It is important to use tests that are set to measure what we are intended to measure (Beyer and Magnusson 2003; Portney and Watkins 2009). There are several methods to measure different variables, some more preferred than others (Portney and Watkins 2009). In the following section the preferred measuring methods for strength-, flexibility- and balance measures will be presented.

Additionally I will present studies assessing inter-rater reliability of the isometric hip abductor strength test, 1RM leg press, the hamstrings flexibility test and the SEBT. A complete list of presented studies is found in table 2. The literature search in PubMed, SPORTDiscus, PEDro and Cochrane Library was conducted on the 17th of September 2012. Additional studies were collected from the studies reference lists. Only studies with ICC scores are presented and studies examining subject with any diagnosis were excluded. Another search was conducted on the 22nd of March 2013 to ensure any new studies were included.

2.5.1 Strength measures

There are several methods of measurements for assessing maximum strength. In a clinical setting it is usual to assess one repetition maximum (1RM). 1RM testing requires an isoinertial contraction – that is, a constant weight is lifted at a voluntary speed. Compared to the preferred standard, a dynamometer, 1RM is also considered as a reliable and valid measuring method for strength in both young and elderly (Verdijk, van Loon et al. 2009). An essential advantage with use of 1RM testing is the transferability to the equipment used in a training context (Abernethy, Wilson et al. 1995). However precise protocols are needed to ensure the reproducibility of 1RM strength tests (Raastad, Paulsen et al. 2010). Familiarization with the exercise, positioning and stabilization of the participant, and instruction and encouragement is important to reduce risk of measurement errors (Verdijk, van Loon et al. 2009).

Dynamometers have been referred to as the *gold standard* for assessing maximal muscle strength (Ly and Handelsman 2002; Verdijk, van Loon et al. 2009). A test of maximum strength with a dynamometer requires either an isometric or isokinetic contraction (Verdijk, van Loon et al. 2009). When appropriate standardization is applied – for example, familiarization with the exercise, positioning and stabilization of the participant, and instruction and encouragement of the participant – dynamometry has been shown to provide highly reliable test results (Abernethy, Wilson et al. 1995; Ly and Handelsman 2002). Compared with the isokinetic dynamometer, a hand held dynamometer can be regarded as reliable and valid instrument for strength assessment in a clinical setting and is substantially cheaper compared to a isokinetic dynamometer (Stark, Walker et al. 2011).

Studies assessing inter-rater reliability of isometric hip abductor strength

Some studies have previous examined inter-rater reliability of isometric hip abductor strength with a hand-held dynamometer. These studies, shown in table 2, presented an ICC value of 0.73-0.85.

Between the studies there are some differences in starting positions. Krause et al. performed the abduction strength test with the subject in side position (Krause, Schlagel et al. 2007). Both Thorborg et al. and Kelln et al. assessed hip abductor strength in a supine starting position (Kelln, McKeon et al. 2008; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012). And in their most recent study they used an external belt fixation of the hand-held dynamometer to exclude bias from tester differences (Thorborg, Bandholm et al. 2012).

Studies assessing inter-rater reliability of 1RM leg press

Some studies have reported excellent intra-rater reliability. These studies are however, mainly conducted on groups with certain diagnosis or on elderly people. Therefor there are no previous studies examining inter-rater reliability of 1RM leg press test on athletes.

2.5.2 Hamstrings flexibility measures

There are several methods of measurements for examining the hamstring flexibility. Some studies have for example measured distance in cm. like the toe-touch test or sit-and-reach test (Dallinga, Benjaminse et al. 2012). More preferred though is range of motion (ROM) measures, in degrees, to assess flexibility (Dallinga, Benjaminse et al. 2012). These measures can be performed both active and passive (Dallinga, Benjaminse et al. 2012). Askling et al. for example recommends active and passive measurement of movement, in degrees, with a standardized straight leg rise to a pain free endpoint (Askling, Nilsson et al. 2010). Flexibility measures where subjects actively stretch to an endpoint is however challenging because of the subjects personal subjective tolerance to stretch (Law, Harvey et al. 2009). To eliminate any subjective factors some studies have measured the endpoint angle with application of a standardized force (Fredriksen, Dagfinrud et al. 1997) or torque (Law, Harvey et al. 2009). In this way the measuring method will be more accurate compared to methods dependant of the subjects feedback (Ben and Harvey 2010).

Studies assessing inter-rater reliability of hamstrings flexibility

Studies examining the inter-rater reliability of combined flexibility measures like the sit-and-reach test or toe-touch test have showed excellent inter-rater reliability (Franchignoni, Tesio et al. 1998; Monnier, Heuer et al. 2012). Shultz et al. measured several anatomical clinical measures, and presented excellent inter-rater reliability on hamstrings flexibility measured as active extension of the knee with the hip flexed in 120°. In this study there are no information about the subjects activity level (Shultz, Nguyen et al. 2006). Atamaz et al. assessed several different flexibility measures in a large group of participants with different activity background (Atamaz, Ozcaldiran et al. 2011). In the hamstrings flexibility test, similar to Shultz et al., they presented moderate inter-rater reliability with an $ICC_{(2.k)}$ value of 0.68-0.73, lowest of all the tests performed in their study (Atamaz, Ozcaldiran et al. 2011). Tests in the same study measuring reaching distance, like finger to floor and sit and reach, showed $ICC_{(2.k)}$ values from 0.71-0.94 (Atamaz, Ozcaldiran et al. 2011).

2.5.3 Balance measures

Postural-control assessments are often categorized in dynamic or static categories (Gribble, Hertel et al. 2012). Static postural balances are tasks were the subject is trying to maintain a position while minimizing body movement, often on instruments as force platform or validated clinical scales as Berg Balance scale (Gribble, Hertel et al. 2012). Dynamic postural control is some form of movement of a segment or body around a stable base as SEBT (Gribble, Hertel et al. 2012).

The SEBT is widely used as a dynamic test in both clinical and research purpose. It is considered being a highly representative non-instrumented dynamic balance test for physical active individuals and can provide an objective measure to differentiate deficits and improvements in dynamic postural control (Gribble, Hertel et al. 2012). The SEBT is often used as an predictor for lower extremity injuries (Gribble, Hertel et al. 2012), however only one study has used the test to evaluate risk for ACL-injury (Herrington, Hatcher et al. 2009). Despite the limited amount of research, authors presented promising results for predicting ACL injuries with SEBT (Herrington, Hatcher et al. 2009).

Studies assessing inter-rater reliability of Star Excursion Balance Test

Some studies have previous examined the inter-rater reliability of the SEBT. As shown in table 2 these studies presented good to excellent inter-rater reliability (Hertel, Miller et al. 2000; Plisky, Gorman et al. 2009; Munro and Herrington 2010). All the studies concluded that familiarization or practice is essential to increase the reliability of the test (Kinzey and Armstrong 1998; Hertel, Miller et al. 2000; Plisky, Gorman et al. 2009; Munro and Herrington 2010).

Table 2: *Studies assessing the inter-rater reliability for hip abductor strength, hamstrings flexibility and SEBT.*

Test	Author	Year	Title	Subject	Results (inter-rater)
Hip abductor strength	Krause et al.	2007	Influence of Lever Arm and Stabilization on Measures of Hip Abduction and Adduction Torque Obtained by Hand-Held Dynamometry	n=21 (12 males, 9 females) Age 22-31	ICC (2.1)=0.73
	Kelln et al.	2008	Hand-Held Dynamometry: Reliability of Lower Extremity Muscle Testing in Healthy, Physically Active, Young Adults	n=20 (9 males, 11 females) Age19-45	ICC (2.k)=0.87
	Thorborg et al.	2011	Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength	n=50 (21 males, 29 females) Age 20-30	ICC (2.1)=0.84
	Thorborg et al.	2012	Hip- and knee-strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable	n=21 (15 males, 6 females) Age 21-28	ICC (2.1)=0.85
Hamstrings flexibility	Shultz et al.	2006	Intratester and Intertester Reliability of Clinical Measures of Lower Extremity Anatomic Characteristics: Implications for Multicenter Studies	n=16 (7 males, 9 females) Age 22-29	ICC (2.1)=0.89-0.97
	Atamaz et al.	2011	Interobserver and intraobserver reliability in lower-limb flexibility measurements	n=66 (53 males, 13 females) Age 18-28	ICC (2.k)=0.68-0.73
Star Excursion Balance Test	Hertel et al.	2000	Intratester and Intertester Reliability During the Star Excursion Balance Tests	n=16 (all males) Age 18-26	ICC (2.k)=0.78-0.93(AL), 0.35-0.81(ML), 0.58-0.88(PL)
	Plisky et al.	2009	The reliability of an instrumented device for measuring components of the Star Excursion Balance Test	n=15 (all males) Age 19-20	ICC (2.1)=0.99-1.00(PL)
	Munro et al.	2010	Between-session reliability of the Star Excursion Balance Test	n=22 (11 males, 11 females) Age 18-26	ICC (2.1)=0.87(AL), 0.91(ML), 0.92(PL)

ICC=interclass correlation coefficient, AL=anterolateral direction, ML=mediolateral direction, PL=posterolateral direction.

3.0 Methods

3.1. Study design

This project was a part of an on-going prospective cohort study at the Oslo Sports Trauma Research Center aiming at investigating risk factors for ACL injuries in elite female football and handball players. The elite handball players have participated in annual screening tests since study start in June 2007, while the football players have been included since February 2009.

This project was a methodological study involving test-retest measurements. We assessed the inter-rater reliability of the isometric hip abductor strength test, 1RM leg extensor strength (leg press), the hamstrings flexibility and the SEBT.

In this study 22 elite female football players and 20 elite female handball players took part in two days of testing. Two different raters tested the players at the two test sessions. Neither the raters nor the players were blinded for the results at any of the two days of testing. The test results from each day were not analysed or used to compare results between test sessions during the test period.

Of the 42 participants, 26 players completed the re-test session within three to ten days. The remaining 16 players participated in the re-test session six to seven weeks after the first test session.

3.2 Participants

We recruited female handball players who were already enrolled in the annual testing in the cohort study. All clubs in the elite league (Postenligaen) were contacted and invited to participate in the screening tests. A total of 22 players, who conducted the annual testing, agreed to participate in the test-retest session.

The female football players were recruited from the elite league (Toppserien) and from one 1st division club. We contacted the coaches of the clubs and sent an information letter about the project (appendix 1). A total of 31 players agreed to participate in the test- re-test session.

To ensure a sufficient number of participants, a new series of test-retest sessions were arranged four weeks later. The players invited to the second session of testing were offered 1480NOK in compensation for lost time at work.

3.2.1 Inclusion criteria

- Elite female football players
- Elite female handball players
- Players had to be able to participate in match play the day of testing

3.2.2 Exclusion criteria

Players were excluded if:

- The players had pain during testing making the player unable to complete the test with maximal effort
- Players could not provide complete results from both test sessions

If there were any uncertainties whether a player could take part in testing with maximal effort, the tester responsible for the relevant test decided if the player could accomplish the test or not.

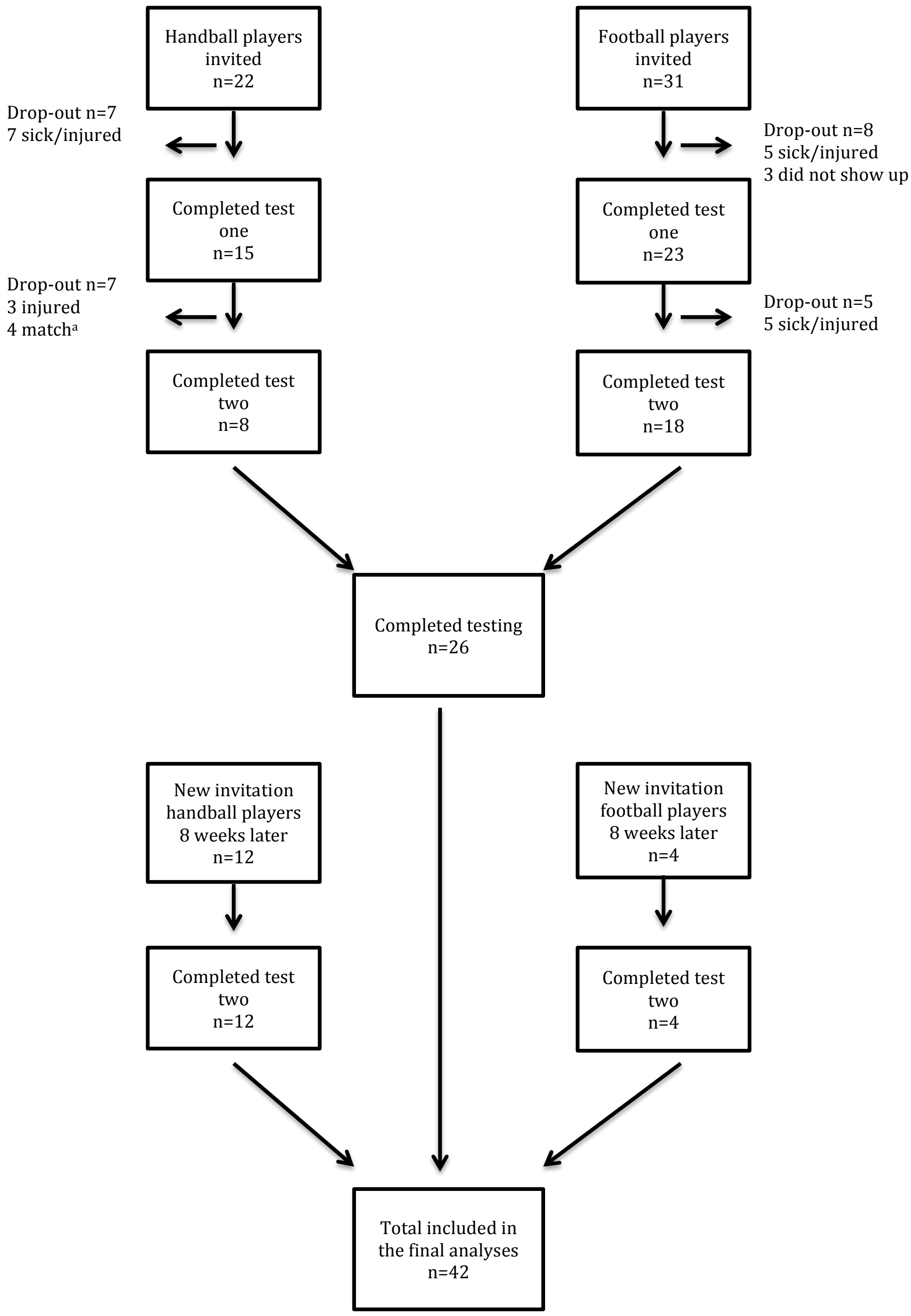


Figure 2: *The flow of the participants in the project.*
[a]=players were unable to attend testing due to participation in an international match.

3.3 Test team

The test team mainly consisted of experienced researchers, physiotherapists or medical students who had been a part of the test team during the previous data collections. Two days of preparation and pilot testing were carried out prior to the testing to ensure that all the testers were confident with the procedures before the data collection started.

The test team was organized on different days of testing to ensure that the players included in the study were assessed with different testers on test first- and second test session.

3.4 Screening tests

The test-retest session was conducted at The Norwegian School of Sports Sciences in Oslo from the 10th to the 20th of August 2012 and 29th of September to the 2nd of October 2012. On the first day of testing, each player signed a written consent form (appendix 4 and 5). The players also gave a blood sample, which were included as a part of the risk factor screening protocol.

The handball players participated for seven hours of testing on their first test session. Players were linked in pairs and rotated on seven different test stations (appendix 1). The players completed all tests in a pre-defined randomized order. The second test session for the handball players, as well as both test days for the football players, was completed in three hours. Six players rotated on six test stations in a pre-defined randomized order with 30 min on each station (appendix 2).

3.4.1 Isometric hip abductor strength test

The hip abductor strength was measured in kg with a hand-held dynamometer (Hydraulic Push-Pull Dynamometer, Baseline® Evaluation Instruments, White Plains, NY, USA). The protocol was similar to tests performed in previous reports (Thorborg, Petersen et al. 2010; Thorborg, Bandholm et al. 2011)

Figure 3: *The placement of fixation belts and testing position for isometric hip abductor strength.*

The test was conducted with the players lying in supine position with the legs extended on an examination table. To stabilize the pelvic fixation belts were positioned over the pelvic and over the thigh that was not tested as showed in figure 3. The dynamometer was positioned two cm proximal to the lateral malleolus with the leg in neutral position and the foot in slight dorsal flexion. The players' arms were held across the chest during the test and the player was told to push the leg against the tester, without lifting thigh and heel from the bench. The contraction was held for two seconds. The tester was instructed to keep the dynamometer as stable as possible. First a submaximal test trial was carried out to get familiar with the procedure and to ensure that the players performed a correct trial. After a short break of minimum 30 seconds to one minute the players conducted two maximal tests on each leg. Both scores were collected for the analyses.

3.4.2 Leg extensor strength (Leg press) test

A custom made seated leg press machine was used to test their one repetition maximum (1RM) measured in kg. As shown in figure 4, the angle of the platform was 115° and the angel of the seat was 100°. Both measured relative to the tracks of the sled. Feet were placed with 14.5 cm in between. The tester ensured that the end of the shoes were at the end of the lower edge of the platform. The players knee angle was measured from the greater trochanter to the femur condyle aligned to the midpoint of the ankle. The players released the safety handles and lowered the sled until they reached 100° knee flexion. As shown on figure 5, a stick was placed at 100°, touching the weights, to ensure that the athlete preformed a valid trial in the following test.

Figure 4: *Angles of the leg press machine.*

The players started with a five minutes warm-up on a stationary bicycle. They then performed a warm-up on the leg press with the same procedure as described for the 1RM test. Starting weight was set at 50 kg and the players performed eight repetitions to familiarize on how to achieve the correct knee angle and perform with correct technique. The players then followed the same procedure for four repetitions with 80-90 kg. The final stage of warm-up was self-chosen, one or two repetitions with 120 kg. The players were given breaks of 30 seconds to one minute and were encouraged to stand up and walk between sets.

Figure 5: *Placement of the stick for 100° position on leg press.*

The 1RM test started with 150 kg. Secondly, the load was increased if they succeeded, but with no more than 30 kg for each step. The players pulled handles to ensure not sliding upwards on the seat and were told to not look at the stick for correct angle. The testers gave verbal encouragement and told the players when the sledge had touched the stick. A trial was valid when the players had reached the required knee angle of 100°, controlled with the sledge touching the stick. If the sledge was not in contact with the stick the trial was not valid. The players conducted testing until they had three lifts that were not valid. Their maximum trial was collected for the analyses.

3.4.3 Hamstrings flexibility test

The hamstrings flexibility was measured as static range of motion (ROM) with a goniometer while lying in supine position on an examination table. Protocol for the testing followed a modified procedure of Fredriksen et al. (1997) and Shultz et al. (2006). Fixation belts were positioned over the pelvic and the non-tested leg. The hip of the tested leg was fixed at 120° flexion using a fixation belt while the players supported against further hip flexion by pressing both hands distally on the femur as shown on figure 6. The tested ankle and foot were relaxed and the hip held in neutral rotation, abduction and adduction. The tester palpated and marked three anatomical landmarks on the tested leg in a flexed position: lateral malleolus, lateral femur epicondyle and the greater trochanter. A tension meter was placed just proximal to the lateral malleolus at a 90° angle to the calf as shown on figure 6. The players were instructed to actively extend the knee like performing an overhead kick. The knee was extended with an eight kg load, as measured with a tension meter. Flexibility was measured as static ROM, in degrees, around lateral femur condyle with the axis of the goniometer in line with the lateral malleolus and the greater trochanter. One measure from each leg was collected.

Figure 6: *Placement of fixation belts and testing position for hamstrings flexibility.*

3.4.4 Star Excursion Balance Test

The players conducted a modified SEBT that assessed anterolateral, mediolateral and posterolateral direction, which are the directions found to correlate the most with chronic ankle instability (Hertel, Braham et al. 2006). The protocol used in this study is based on earlier studies (Hertel, Miller et al. 2000; Munro and Herrington 2010).

Figure 7: *Starting position and tester position of the Star Excursion Balance Test.*

Three tape measures were attached to the floor, with 45° between as shown on figure 7, and represented anterolateral, mediolateral and posterolateral directions. The starting position was the point where the three lines met and the players always conducted the testing without shoes. The player's aim were to find a balance on the standing leg and reach out with the other leg as far as possible while the heel on the standing foot had to be in contact with the floor. The players were instructed to hold their hands on the crista of the pelvic. The leg was moved out as far as possible and kept in position for a short while, without touching the floor, and moved back to starting position. The non-dominant leg was always the standing leg first. The players reached anterolateral first, then mediolateral and finally in posterolateral direction. The tester was seated on the floor in line with the mediolateral line, as shown on figure 7, and measured maximal length manually in cm of the tape measures attached to the floor. This procedure was conducted in all three directions on both feet, without break.

For a trial to be valid the players had to keep their hands on the crista of the pelvic; the reaching leg could not touch the ground providing support; and the heel of the stance leg had to be kept in position and not be lifted from the ground. All players were given one test trial in each direction on one foot to get familiar with the testing procedures. Three complete rounds were recorded on each foot.

3.4.5 Additional screening tests

In addition to tests included in this project the players conducted anthropometric measures, 3D motion analysis, measures of hip anteversion, genu recurvatum, knee laxity and generalized joint laxity, static and dynamic balance on balance 3000, isokinetic hamstring/quadriceps strength and subjective assessment of knee and hip control in single leg squats and drop jumps.

3.5 Statistical analyses

All statistical analysis was conducted using SPSS (Version 21, Mac OSX 10.8.2). Anthropometric variables age, height and weight (self reported) and years playing at the elite level are presented with means and standard deviations (SD). From the different measurements several variables were calculated and presented with mean and SD. We tested the dependant variables for normal distribution using Kolmogorov-Smirnov test. Independent samples t-tests were used to assess differences between handball and football players. For strength measures both absolute and relative strength (kg/body weight) were presented.

To assess relative reliability we used ICC. A two way random model was used to assess inter rater reliability (ICC 2.1/ICC 2.K) (Shrout and Fleiss 1979). For all the ICC´s agreement were used to take in consideration the systematic differences between testers (McGraw and Wong 1996). For all single measures or max scores $ICC_{(2.1)}$ was applied and the single measure value of ICC in SPSS output was used. Calculation of ICC for the mean values of hip abduction strength and the SEBT was conducted using $ICC_{(2.k)}$. From the output in SPSS a 95% confidence interval (CI) was specified. To interpret the ICC score we used the general guidelines that are often suggested. Values >0.75 are indicated as good to excellent reliability, values <0.50 are indicated as poor reliability and values between 0.50 and 0.75 are indicated as moderate to good (Lexell and Downham 2005; Portney and Watkins 2009). The mean difference (MD) was calculated to show difference in mean between the test sessions.

Absolute reliability was presented with standard error of the measurement (SEM). It was calculated $SEM = \sqrt{residual\ mean\ square}$ with the residual mean square collected from the ANOVA (Stratford and Goldsmith 1997). SEM % is also presented as $SEM\% = \frac{SEM}{Grand\ mean}$, expressing the measurement variability as a CV (Lexell and Downham 2005). Bland-Altman plots were presented to give a visual presentation of the absolute reliability. In these plots the difference between measurements from the two test days were plotted against the mean of the two test occasions for each subject and any systematic bias or outliers could be seen (Bland and Altman 1986).

Statistical significant differences were considered when p-values were equal or smaller then 0.05.

3.6 Ethics

The Regional Committee for Medical Research Ethics and the Norwegian Social Science Data Services have both approved the cohort study. As our project was a part of the cohort study, there was no need for further approvals. We provided the players with a sheet of information about the study (appendix 1), a written consent form (appendix 4), and for players under 18 a consent form to be signed of parents (appendix 5). All players were informed about the possibility to leave the project at any time during the course of the study.

3.7 Organization

As this study was a part of the prospective cohort study “Risk factors for ACL injuries in elite female handball and football players”, it was financially supported through the Oslo Sports Trauma Research Center.

4.0 Results

4.1 Anthropometric data for the participants

Anthropometric data for the players that attended testing are presented in table 3. There was no difference in any of the anthropometric data between the two groups (p=0.06 – 0.93).

Table 3: *Descriptive statistics for participants in the study. Mean ± SD*

	Handball (n=20)	Football (n=22)	Both groups (n=42*)	Difference (p)
Age (years)	21.9 ± 4.4	21.1 ± 3.2	21.5 ± 3.8	0.50
Height (cm)	170.0 ± 6.3	166.9 ± 3.3	168.5 ± 5.2	0.06
Weight (kg)	68.0 ± 9.0	63.6 ± 4.9	66.0 ± 7.6	0.07
Elite (years)	3.4 ± 4.2	3.3 ± 3.2	3.5 ± 3.7	0.93

*Weight (n=38) and number of years playing at the elite level (n=40)

4.2 Strength-, hamstrings flexibility- and SEBT measures for both sports

All the tests on both days were normally distributed, except for hamstrings flexibility right side. As shown in table 4 we can see that there was a tendency that the handball players scored higher in all the strength measures. But there was no difference between handball- and football players, with exception of *leg press test one*. In further analyses, handball- and football players are presented as one group, while for leg press data were presented for both groups.

Table 4: *Test results for handball and football on both days (Mean ± SD) and compares the sports for differences on both test days (p-value).*

	Sport	N	Test day one	p-value	N	Test day two	p-value
Hip strength R mean (kg)	Handball	20	16.5 ± 1.8	0.72	20	17.3 ± 2.6	0.76
	Football	22	16.7 ± 1.9		22	17.1 ± 2.2	
Hip strength R max (kg)	Handball	20	17.3 ± 1.9	0.80	20	17.8 ± 2.8	0.82
	Football	22	17.4 ± 1.9		22	17.6 ± 2.3	
Hip strength L mean (kg)	Handball	20	16.5 ± 2.3	0.67	20	16.6 ± 2.8	0.91
	Football	22	16.2 ± 2.0		22	17.7 ± 2.2	
Hip strength L max (kg)	Handball	20	17.4 ± 2.4	0.63	20	18.2 ± 2.9	0.98
	Football	22	17.0 ± 2.2		22	18.2 ± 2.4	
Hamstrings flexibility R (°)	Handball	20	136.3 ± 12.6	0.86	20	135.5 ± 13.2	0.54
	Football	22	135.6 ± 10.3		22	133.0 ± 12.2	
Hamstrings flexibility L (°)	Handball	20	136.8 ± 11.9	0.93	20	134.4 ± 13.0	0.90
	Football	22	136.4 ± 11.7		22	133.5 ± 13.3	
SEBT anterolateral R mean (cm)	Handball	20	78.0 ± 6.3	0.76	20	81.0 ± 6.5	0.37
	Football	22	77.4 ± 6.3		22	79.2 ± 6.3	
SEBT anterolateral R max (cm)	Handball	20	80.0 ± 6.2	0.58	20	83.1 ± 6.9	0.20
	Football	22	78.9 ± 6.3		22	80.4 ± 6.3	
SEBT mediolateral R mean (cm)	Handball	20	81.3 ± 5.6	0.39	20	83.2 ± 5.6	0.70
	Football	22	79.8 ± 5.6		22	82.5 ± 5.8	
SEBT mediolateral R max (cm)	Handball	20	83.5 ± 5.8	0.37	20	84.7 ± 5.9	0.68
	Football	22	81.9 ± 5.5		22	84.0 ± 5.7	
SEBT posterolateral R mean (cm)	Handball	20	87.8 ± 5.3	0.18	20	88.7 ± 6.3	0.80
	Football	22	85.6 ± 5.0		22	88.2 ± 5.1	
SEBT posterolateral R max (cm)	Handball	20	90.5 ± 6.1	0.09	20	91.5 ± 6.1	0.51
	Football	22	87.5 ± 5.1		22	90.3 ± 5.4	
SEBT anterolateral L mean (cm)	Handball	20	78.6 ± 7.1	0.62	20	81.4 ± 6.3	0.12
	Football	22	77.7 ± 5.4		22	78.4 ± 6.1	
SEBT anterolateral L max (cm)	Handball	20	80.8 ± 7.5	0.47	20	83.5 ± 6.6	0.07
	Football	22	79.3 ± 5.9		22	79.8 ± 6.2	
SEBT mediolateral L mean (cm)	Handball	20	82.6 ± 5.9	0.28	20	84.0 ± 6.3	0.31
	Football	22	80.8 ± 4.7		22	82.2 ± 5.3	
SEBT mediolateral L max (cm)	Handball	20	84.3 ± 6.1	0.46	20	86.5 ± 7.3	0.20
	Football	22	83.0 ± 5.2		22	84.0 ± 5.1	
SEBT posterolateral L mean (cm)	Handball	20	88.6 ± 5.2	0.06	20	90.7 ± 5.9	0.11
	Football	22	85.7 ± 4.4		22	88.0 ± 4.7	
SEBT posterolateral L max (cm)	Handball	20	90.9 ± 5.3	0.07	20	92.5 ± 6.6	0.11
	Football	22	88.1 ± 4.6		22	89.5 ± 4.6	
Leg press (kg)	Handball	19	186.2 ± 36.4	0.01*	16	187.7 ± 28.4	0.13
	Football	22	159.3 ± 28.5		22	171.8 ± 32.3	

*p<0.05, R=right, L=left, SEBT=Star Excursion Balance Test, abd.=abduction

4.2.1 Changes in strength- and hamstrings flexibility measures

As shown in table 5, all measures, with the exception of hamstrings flexibility, increased from test one to test two. The hamstrings flexibility for both legs and the leg press for handball players showed no difference from the first to the second test session.

Table 5: *Strength measures for football and handball combined for both test sessions (Mean ± SD).*

Test (n=42[a])	Test day one	RS day one	Test day two	RS day two	p-value
Hip abd. strength R mean (kg)	16.6 ± 1.8	0.25 ± 0.03	17.2 ± 2.4	0.26 ± 0.03	0.02*
Hip abd. strength R max (kg)	17.3 ± 1.8	0.26 ± 0.03	17.7 ± 2.5	0.27 ± 0.04	0.24
Hip abd. strength L mean (kg)	16.4 ± 2.1	0.25 ± 0.03	17.6 ± 2.5	0.27 ± 0.04	0.01*
Hip abd. strength L max (kg)	17.2 ± 2.3	0.26 ± 0.03	18.2 ± 2.6	0.28 ± 0.04	0.01*
Hamstrings flexibility R (°)	135.9 ± 11.3	-	134.2 ± 12.6	-	0.23
Hamstrings flexibility L (°)	136.6 ± 11.6	-	133.8 ± 13.0	-	0.05
Leg press (kg)	171.6 ± 36.1	2.6 ± 0.5	178.4 ± 31.8	2.7 ± 0.4	0.04*
Leg press handball (kg)	186.2 ± 36.4	2.8 ± 0.6	187.7 ± 28.4	2.8 ± 0.4	0.80
Leg press football (kg)	159.3 ± 28.5	2.5 ± 0.4	171.8 ± 32.3	2.7 ± 0.5	0.01*

[a] for leg press (n=41) test day one and (n=38) on test day two, *p<0.05, RS=relative strength (kg/body weight), R=right, L=left, abd.=abduction

4.2.2 Changes in the SEBT measures

As shown in table 6 there was an increased score from the first to the second test session for all directions. The maximum score for the SEBT in anterolateral direction with left leg was the only score that was not different from test session one to test session two. The anterolateral direction, with both legs, was the test with greatest variation.

Table 6: *The SEBT measures for football and handball combined for both test sessions (Mean ± SD).*

Test (n=42)	Test day one	Test day two	p-value
SEBT anterolateral R mean (cm)	77.7 ± 6.2	80.0 ± 6.3	0.01*
SEBT anterolateral R max (cm)	79.4 ± 6.2	81.7 ± 6.6	0.01*
SEBT mediolateral R mean (cm)	80.5 ± 5.6	82.8 ± 5.6	0.01*
SEBT mediolateral R max (cm)	82.6 ± 5.6	84.3 ± 5.7	0.01*
SEBT posterolateral R mean (cm)	86.7 ± 5.2	88.5 ± 5.6	0.01*
SEBT posterolateral R max (cm)	88.9 ± 5.8	90.8 ± 5.7	0.01*
SEBT anterolateral L mean (cm)	78.1 ± 6.2	79.8 ± 6.3	0.01*
SEBT anterolateral L max (cm)	80.0 ± 6.7	81.5 ± 6.7	0.05
SEBT mediolateral L mean (cm)	81.7 ± 5.3	83.1 ± 5.8	0.03*
SEBT mediolateral L max (cm)	83.6 ± 5.6	85.1 ± 6.3	0.04*
SEBT posterolateral L mean (cm)	87.1 ± 5.0	89.3 ± 5.4	0.01*
SEBT posterolateral L max (cm)	89.4 ± 5.1	90.9 ± 5.8	0.02*

*p<0.05, SEBT=Star Excursion Balance Test, R=right, L=left.

4.3 Inter-rater reliability for strength- and hamstrings flexibility measures

As shown in table 7, ICC values for the strength measures ranged from 0.57-0.84. Mean difference between tests (MD) is negative for most of the test. This means a systematic error showing higher score on test day two. The exception was the hamstrings flexibility, both legs, and leg press handball.

Table 7: *Relative inter-rater reliability as ICC with 95% CI and mean difference (MD) between test day one and two for the strength measures. Absolute reliability is presented with standard error of the measurement (SEM) and as a coefficient of variation SEM %.*

Test n=42	ICC	95 % CI ICC	MD	SEM	SEM %
Hip abd. strength R mean (kg)	0.67[a]	(0.62 - 0.90)	-0.6	1.2	6.9
Hip abd. strength R max (kg)	0.68[c]	(0.47 - 0.81)	-0.3	1.3	7.2
Hip abd. strength L mean (kg)	0.69[a]	(0.33 - 0.85)	-1.3	1.5	8.7
Hip abd. strength L max (kg)	0.57[c]	(0.29 - 0.75)	-1.0	1.5	8.6
Hamstrings flexibility R (°)	0.71[c]	(0.52 - 0.83)	1.7	6.5	4.8
Hamstrings flexibility L (°)	0.72[c]	(0.53 - 0.84)	2.8	6.4	4.7
Leg press (kg)	0.83[c]	(0.68 - 0.91)	-6.8	13.6	7.8
Leg press handball (kg)	0.78[c]	(0.46 - 0.92)	-1.5	16.7	8.8
Leg press football (kg)	0.84[c]	(0.33 - 0.95)	-12,5	9.5	5.7

*Leg press n=38, leg press handball n=16, leg press football n=22, [a] ICC (2.k), [c] ICC (2.1), ICC=interclass correlation coefficient, CI=confidence interval, R=right, L=left, abd.=abduction.

4.4 Inter-rater reliability for the SEBT measures

As shown in table 8, ICC values for the SEBT measures ranged from 0.66-0.90. For all the balance measures the players scored higher on test day two, resulting in a negative mean difference between tests (MD), indicating a systematic error.

Table 8: *Relative inter-rater reliability as ICC with 95% CI and mean difference (MD) between test day one and two for Star Excursion Balance Test measures. Absolute reliability is presented with standard error of the measurement (SEM) and as a coefficient of variation SEM %.*

Test n=42	ICC	95% CI ICC	MD (cm)	SEM (cm)	SEM%
SEBT anterolateral R mean (cm)	0.90[b]	(0.67 - 0.96)	-2.3	2.3	3.0
SEBT anterolateral R max (cm)	0.77[c]	(0.51 - 0.89)	-2.3	2.8	3.4
SEBT mediolateral R mean (cm)	0.89[b]	(0.59 - 0.95)	-2.3	2.2	2.7
SEBT mediolateral R max (cm)	0.81[c]	(0.61 - 0.91)	-1.7	2.2	2.7
SEBT posterolateral R mean (cm)	0.84[b]	(0.67 - 0.92)	-1.8	2.6	3.0
SEBT posterolateral R max (cm)	0.68[c]	(0.44 - 0.82)	-2.0	3.1	3.4
SEBT anterolateral L mean (cm)	0.88[b]	(0.75 - 0.94)	-1.7	2.7	3.5
SEBT anterolateral L max (cm)	0.73[c]	(0.54 - 0.84)	-1.5	3.4	4.2
SEBT mediolateral L mean (cm)	0.84[b]	(0.70 - 0.92)	-1.4	2.8	3.4
SEBT mediolateral L max (cm)	0.69[c]	(0.49 - 0.82)	-1.5	3.2	3.8
SEBT posterolateral L mean (cm)	0.81[b]	(0.54 - 0.91)	-2.2	2.7	3.0
SEBT posterolateral L max (cm)	0.66[c]	(0.44 - 0.81)	-1.5	3.0	3.4

[b] ICC (2.k) [c] ICC (2.1), SEBT=Star Excursion Balance Test, ICC=interclass correlation coefficient, CI=confidence interval, R=right, L=left.

4.5 Absolute reliability for strength-, hamstrings flexibility- and the SEBT measures

As presented in table 7, SEM varied 1.2-1.5 cm for the isometric hip abduction strength, 4.7-4.8° for the hamstrings flexibility and 5.7-8.8 kg for the leg press. The SEM% varied from 4.7-8.8 % for the strength and hamstrings flexibility measures. For the Star Excursion Balance Test the SEM varied from 2.2-3.2 cm and the SEM % varied from 2.7–4.2 % as shown in table 8.

Figure 8-13 gives a visual image of systematic- and random error of the individual variations for the strength- and flexibility measures and the SEBT. The plots include upper- and lower limits (two SD) and values for the plots are presented in table 7 and 8 with mean difference between tests (MD) representing systematic error. Most of the scores presented were within ± 2SD. (Portney and Watkins 2009).

4.5.1 Visual presentation of the absolute reliability for Strength- and hamstrings flexibility measures

Figure 8 and 9 shows that hip abductor strength and leg press, both football- and handball players had a mean score below zero line, indicating a systematic error were the players scored higher on test two. Figure 10 show that passive hamstrings flexibility test had greater score on test one.

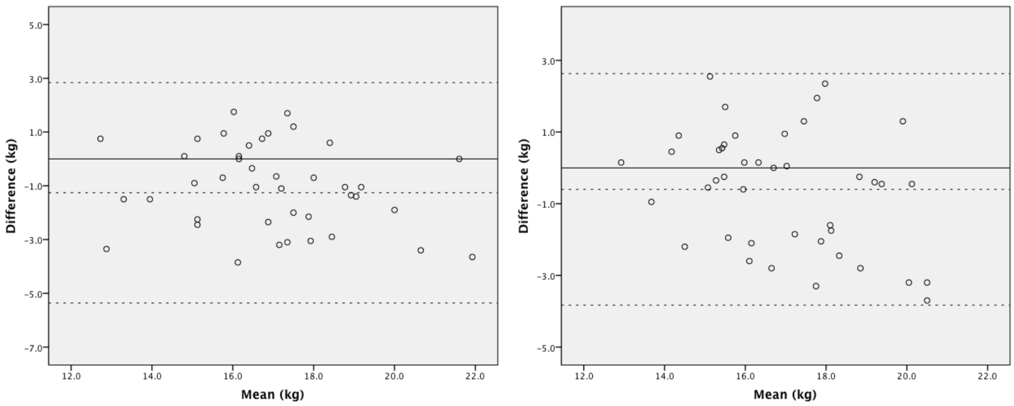


Figure 8: *Bland Altman plot for mean hip abductor strength left leg (left) and right leg (right). The difference between measurements from the two test sessions is plotted against the mean of the two test occasions for each subject. Dotted lines shows upper- and lower limit of agreement, while the central dotted line shows the mean score.*

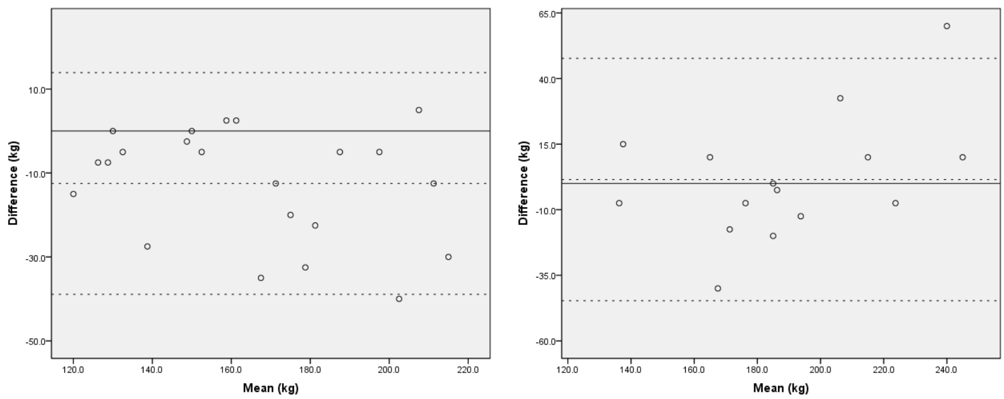


Figure 9*: Bland Altman plot for maximum leg press for the football players (left) and the handball players (right). The difference between measurements from the two test sessions is plotted against the mean of the two test occasions for each subject. Dotted lines shows upper- and lower limit of agreement, while the central dotted line shows the mean score.*

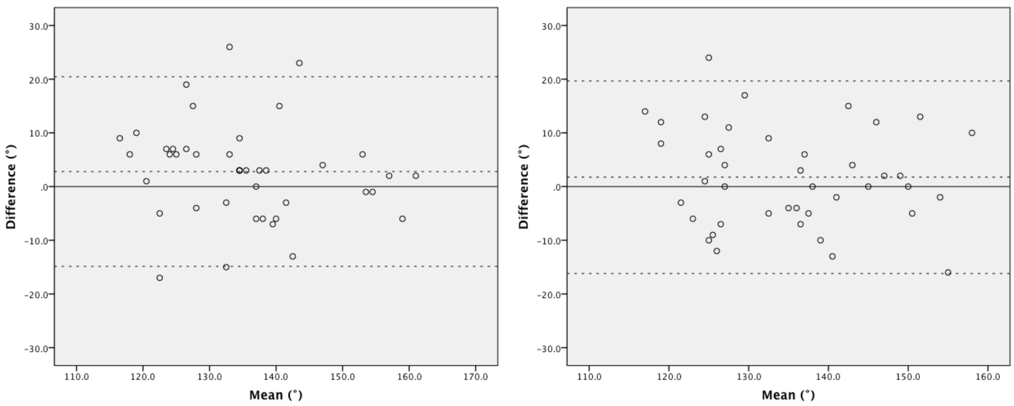


Figure 10*: Bland Altman plot for the maximum hamstrings flexibility left leg (left) and right leg (right). The difference between measurements from the two test sessions is plotted against the mean of the two test occasions for each subject. Dotted lines shows upper- and lower limit of agreement, while the central dotted line shows the mean score.*

4.5.2 Visual presentation of the absolute reliability for the SEBT

Figure 11-13 shows that for all directions, both legs, the SEBT had a mean score bellow the zero line, indicating a systematic error were the players achieving higher score on test session two.

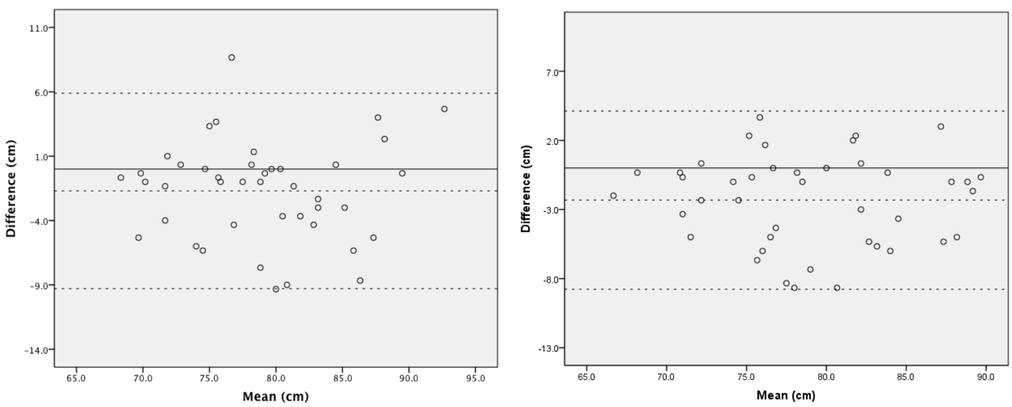


Figure 11*: Bland Altman plot for the mean score of the SEBT anterolateral direction left leg (left) and right leg (right). The difference between measurements from the two test sessions is plotted against the mean of the two test occasions for each subject. Dotted lines shows upper- and lower limit of agreement, while the central dotted line shows the mean score.*

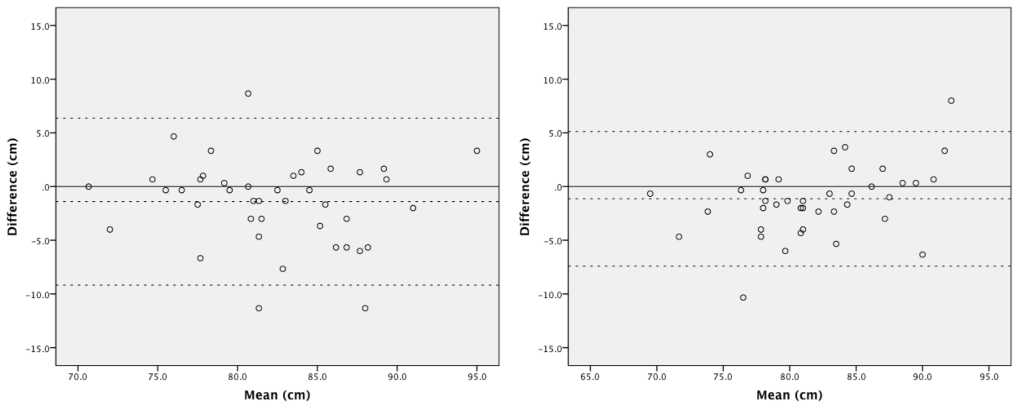


***Figure 12**: Bland Altman plot for the mean score of the SEBT mediolateral direction left leg (left) and right leg (right). The difference between measurements from the two test sessions is plotted against the mean of the two test occasions for each subject. Dotted lines shows upper- and lower limit of agreement, while the central dotted line shows the mean score.*

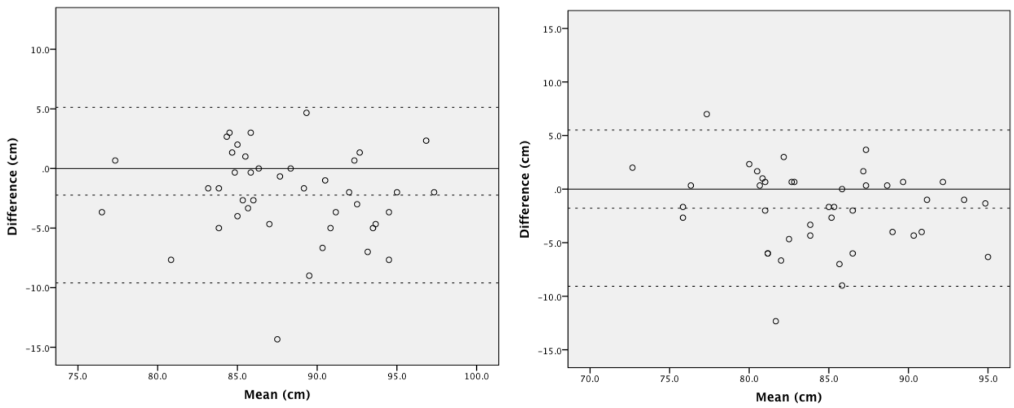


***Figure 13**: Bland Altman plot for the mean score of the SEBT posterolateral direction left leg (left) and right leg (right). The difference between measurements from the two test sessions is plotted against the mean of the two test occasions for each subject. Dotted lines shows upper- and lower limit of agreement, while the central dotted line shows the mean score.*

5.0 Discussion

The aim of this study was to investigate the inter-rater reliability of tests assessing the isometric hip abductor strength, 1RM in a leg press machine, the hamstrings flexibility and the Star Excursion Balance Test evaluating risk factors for ACL injury in elite female football- and handball players.

5.1 Main findings

We found excellent inter-rater reliability of the measurements between two testers in the SEBT and the 1RM leg press. For the hamstrings flexibility- and the isometric hip abductor strength measures there were moderate to good inter-rater reliability between two testers.

The absolute reliability showed a greater individual variation in the strength measures while the SEBT measures showed the smallest individual variation. For the SEBT the absolute reliability was quite comparable in each direction, with the mediolateral reach (right leg) showing smallest individual variation. For the strength measures the individual variation was quite comparable. The exception was the leg press for the football players showing the lowest individual variation, while the handball players showed the greatest individual variations.

With exception of the hamstrings flexibility, both legs, and the leg press for the handball players all the tests showed a small systematic error with higher scores on test two. A systematic error can derive from several different factors, like learning effect among raters/subjects, malfunction or lack of calibration of instruments (Portney and Watkins 2009). These different factors will be discussed in this section.

5.1.1 Strength- and hamstrings flexibility measures

Isometric hip abductor strength test

For the hip abductor strength we presented moderate to good inter-rater reliability on both legs with $ICC_{(2.k)}$ value of 0.67-0.69. The sizes of SEM were acceptable when compared to the mean score on each leg. There was a significant difference from the first test session to the second test session, with exception of the right leg maximum measures. This indicates a small systematic error with an increase in score from the first test session of 0.3-1.3 kg. The SEM, an expression of the individual variations of the measurements, was 1.2-1.5kg. This indicates that both the systematic- and the random errors in these measurements were small. It also indicates that the amount of random error might be slightly higher than the amount of systematic error.

For the hip abductor strength it seems like learning effect from test to retest was the major factor affecting the small systematic error of the measurement. In our study one sub-maximal effort was followed by two maximal efforts on each leg. More trials for familiarization or practice with the exercise, positioning and stabilization of the participant, instruction and encouragement are important to reduce risk of measurement errors in every strength measures (Verdijk, van Loon et al. 2009). More maximal trials will also give a mean value of the measurements more close to the true value when the random error cancel each other out (Portney and Watkins 2009). Compared to other studies assessing inter-rater reliability of the hip abductor strength our study has fewer practice trials and fewer maximal trials. By increasing number of practice trials, or add a familiarization test prior to ordinary testing, the players would most likely not have shown the increase from the first to the second test session. In future testing in the cohort it would be advisable that the players had more practice trials. That would increase the chance to achieve a measurement more close to the players' real score.

The individual variation of the measurements was 6.9-8.7 % (SEM%). There are no reference values for the SEM scores, but compared to the other strength measures in this study the values were comparable and acceptable.

We were measuring an isolated hip abduction movement, but a small change in movement from the players could increase muscle strength by activating the hip flexors. The hip flexors are a stronger muscle group compared to the hip abductors (Dahl and Rinvik 2010) and activation of the hip flexors can potentially increase the score for the players. Our protocol for the hip abductor strength gave an accurate and exact description of the performance of the test and we believe that our experienced testers assured that this activation did not occur. In a similar way as activation of hip flexor could affect the results, the positioning of the hand-held dynamometer may affect the results. If the dynamometer was placed anterior or posterior of the lateral malleolus this could reduce the players score. An ideal performance was a lateral movement of the hip, but an oblique placement of the hand-held dynamometer would lead to a potential underestimation of the force exerted. Although the testers were experienced and the protocol well known for them, a difference in performance of the player or the placement of the dynamometer can happen. The testers motivation or tiredness might reduce their vigilance and reduce the focus on the performance of the test. In this way different placement of the hand-held dynamometer and activation of the hip flexors may be a potential factor for both systematic- and random error.

There are also other factors that can explain our ICC value of 0.67-0.69. Hip strength assessment using hand-held dynamometer may be subject to systematic inter-rater error when testers of different sex and upper- extremity strength perform the measurements (Thorborg, Bandholm et al. 2012). Strong athletes can easily have hip abductor strength up to 20 kg (Thorborg, Serner et al. 2011). In these situations there is a possibility that the athletes are stronger than upper extremity strength of the tester and the tester may underestimate the athletes real strength (Wikholm and Bohannon 1991; Wadsworth, Nielsen et al. 1992; Thorborg, Bandholm et al. 2011). The players in our study had mean hip abductor strength from 16.4-18.2 kg and there is a possibility that the players could be stronger than the testers in this study. This might add upper arm strength of the tester as a confounding factor in this study. Inconsistency of the measurements can affect the ICC value negatively, and could be a possible factor decreasing the ICC scores. A recent study has shown that external fixation of the hand-held dynamometer exclude the arm strength to the tester as a confounding factor (Thorborg, Bandholm et al. 2012). Both our testers were females, which may reduces the risk of systematic measurements error due to gender differences of testers. However with the high strength of the participants in this study it would probably be advantageous with external fixation of the hand-held dynamometer. It would also be a useful modification of the test protocol in future testing in the cohort.

In this study the SD for the hip abductor strength was 1.8-2.6 kg. The small SD indicates that the sample group in this study was homogenous. A homogenous sample group may affect the ICC score negatively due to the chance of reduced variance of measurement between the subjects (Portney and Watkins 2009). This variability must be large to demonstrate high reliability with ICC (Portney and Watkins 2009). The homogenous sample group may therefore be a factor affecting the ICC value in this study. There are some methods of increasing the variability between subjects in the sample size, which will be discussed later. However an increase of the variability would not increase the consistency of the measurement (Portney and Watkins 2009).

Compared with other studies, examining inter-rater reliability, our $ICC_{(2.k)}$ of 0.67-0.69 is lower than Kelln et al. who showed $ICC_{(2.k)}$ of 0.87. Thorborg et al. showed $ICC_{(2.1)}$ of 0.84 in their first study (Thorborg, Bandholm et al. 2011) and $ICC_{(2.1)}$ 0.85 in their most recent study (Thorborg, Bandholm et al. 2012). A comparing of the results from other studies assessing inter-rater reliability is however difficult due to differences in the sample groups. In addition to the smallest SD presented, our sample group have higher mean hip strength compared to the other studies. This shows that comparing the results in the different studies is difficult due to our unique and more homogenous sample group.

1RM leg press test

We presented excellent relative inter-rater reliability for the 1RM leg press test. Split by sports the football players presented a greater $ICC_{(2.1)}$ value of 0.84 compared to the handball players with $ICC_{(2.1)}$ value of 0.78. While the handball players showed a SEM % on 8.8 %, which was comparable with the other strength measures in this study, the football players showed a considerable lower value of 5.7 %. There is no reference value for these measurements, but the football players showed a lower individual variation of the measurements. The MD between the tests on -1.5 kg for the handball players is very small, indicating an increased score from the first test session of 0.8 %. For the football players however the MD between the tests was -12.5 kg, indicating an increased score from the first test session of 7.2 %. For the football players this means a significant increase from test session one to two, indicating a potential learning effect. This is a systematic error in the measurements of the football players, while for the handball players the errors of measurements most likely come from random factors. Since the execution of the tests follows the same procedure for both sports the differences between the sports most probably derives from the differences in the two groups.

As presented earlier sixteen of the football players came from a team recently relegated from the top league. Many of these players were young and new on the team and therefore lacked experience from elite level. For several of the players the leg press was an unknown exercise and during the test sessions some of the youngest football players did not reach the start weight of 150 kg. This made us differ from test protocol a few times with the youngest football players. When assessing the test scores, several of the players from this club had a great increase in maximum leg press from test session one to session two. We did not take into account that so many of the players were unfamiliar with strength testing and this may have affected the results. Familiarization with the exercise, positioning and stabilization of the participant, and instruction and encouragement is important to reduce risk of measurement errors in every strength measures (Verdijk, van Loon et al. 2009). To ensure reliability of the measurements it is suggested that a person new to the exercise should have conducted at least two to five sessions with practicing on the same exercise as the testing is to be conducted on (Jacobs, Schilling et al. 2009). The combination of many young players and the lack of familiarization of the test can explain why the football players had greater increase from test session one to session two compared to the handball players. This supports the possibility that the systematic error can derive from a learning effect from the first to the second test session. It is likely that our results would have been even better if the players had performed a practice session prior the ordinary testing. It would also be advisable to add a practice session for young players prior to the annual testing in the cohort.

Hamstrings flexibility test

I this study we presented moderate to good relative inter-rater reliability for hamstrings flexibility for both legs with $ICC_{(2.1)}$ of 0.71-0.72. The sizes of the SEM for both legs are acceptable when compared to the mean. In the same way as for the strength measures the small SD of 11.3-13.0° for the hamstrings flexibility indicates that our sample group was homogenous. As discussed earlier in this section a homogenous sample group may affect the variance of the measurements. The reduced variance affects the calculation of ICC, which is dependant of variance to calculate a high ICC (Streiner and Norman 2008; Portney and Watkins 2009).

It is difficult to compare our results with other studies due to differences in the test procedure and our unique sample group of female elite football- and handball players. Atamaz et al. (Atamaz, Ozcaldiran et al. 2011) has 20 athletes in their sample group, however only three females. Shultz et al. (Shultz, Nguyen et al. 2006) gave no information regarding the activity level of their participants, but their mean score on each test was 123,5-129.4° which was lower than in our group with 133.8-136.6°. This indicates that a comparing of the results might be inappropriate due to differences in the sample groups. Nevertheless, if we look at our $ICC_{(2.1)}$ value of 0.71-0.72, it was lower compared to Shultz et al. (Shultz, Nguyen et al. 2006) who presented $ICC_{(2.1)}$ value of 0.89-0.97 while our value was close to Atamaz et al. who presented $ICC_{(2.k)}$ value of 0.68-0.73.

In the statistical analyses there are factors, which affect the calculation of the ICC. We used $ICC_{(2.1)}$ with the single measurement we collected for each leg. Shultz et al.

presented their value as $ICC_{(2.1)}$, but used the mean of three measurements and this was basically a calculation of $ICC_{(2.k)}$. Within each model of ICC single ratings will yield a lower correlation than a calculation based on mean ratings (Portney and Watkins 2009). If Shultz et al. have calculated their ICC with a mean value and interpret it as an $ICC_{(2.1)}$ their ICC was an overestimation, and in practice their value was probably much closer to our value.

Compared with the other studies that assessed inter-rater reliability of the hamstrings flexibility there were some differences in starting position. Both Shultz et al. (Shultz, Nguyen et al. 2006) and Atamaz et al. (Atamaz, Ozcaldiran et al. 2011) examined the flexibility of hamstrings during an active extension of the knee with the hip placed in 90° flexion (Atamaz, Ozcaldiran et al. 2011) and 120° flexion (Shultz, Nguyen et al. 2006). In this test set-up the results of the measurements can be affected by the players subjective feel of end-point. By reducing the subjective factors, measurements will be more accurate compared to methods dependant of the subjects' feedback (Ben and Harvey 2010). We believe that our set-up with an applied standardized load to the active extension of the knee is preferable. This reduces the subjective influence as a cofounding factor. Even though we were unable to present the same high ICC score as Shultz et al., we believe that some of the reason was due to our more homogeneous sample group and the calculation of their ICC value which probably was wrong.

5.1.2 Star Excursion Balance Test

We presented excellent relative inter-rater reliability for the SEBT with $ICC_{(2.k)}$ values of 0.81-0.90. The sizes of SEM for the different directions on both legs were acceptable when compared to the mean in each direction. In the different directions on both legs there was a significant difference from test one to test two. The MD between the test sessions was from -1.4 to -2.4 cm, indicating a small systematic error. The SEM for the different directions was 2.2-3.4 cm and the SEM % was 3.0-4.2 %. This indicates that the amount of random error could be greater then the systematic error. The systematic error with increased score on test session two was most likely a learning effect during the testing.

Our protocol with only one practice trial may not be sufficient to eliminate the learning effect. By performing more practice trials the players would have reduced the learning effect during the test session and our results probably would have been even better. Other studies measuring reaching distance on SEBT concluded that familiarization or training was essential to increase the reliability (Hertel, Miller et al. 2000; Plisky, Gorman et al. 2009; Munro and Herrington 2010). The studies presented the highest ICC value also had the most practice trials; four (Munro and Herrington 2010) and six (Plisky, Gorman et al. 2009).

It is difficult to compare the results from the different studies assessing inter-rater reliability of the SEBT. There are for example differences in attributes of the sample group, gender or lacking information on activity level. However, when comparing the mean of the measurements from the different studies it seems like Hertel et al. (Hertel, Miller et al. 2000) and Munro et al. (Munro and Herrington 2010) had sample groups most comparable to ours. Our $ICC_{(2.k)}$ values of 0.81-0.90 were even better than Hertel et al. (Hertel, Miller et al. 2000) who presented an $ICC_{(2.k)}$ score of 0.35-0.93 using the same protocol as us. Plisky et al. (Plisky, Gorman et al. 2009) and Munro et al. (Munro and Herrington 2010) respectively presented $ICC_{(2.1)}$ score of 0.99-1.00 and 0.87-0.92. Munro et al. score was close to ours, but was probably even better than our score. Because a single measurement calculation will give a greater ICC value compared to a mean measurement calculation of ICC of the same set of data (Portney and Watkins 2009). Our SEM values of 2.2–2.8 cm was almost equal to Munro et al. with 2.2–2.9 cm (Munro and Herrington 2010) and somewhat smaller than Hertel et al. (Hertel, Miller et al. 2000) who showed 2.3–5.0 cm. While Pliskey et al. presented SEM values 0.7–0.9 cm and was much better than all the other studies (Munro and Herrington 2010). Neither of the other studies examining inter-rater reliability of the SEBT presented SEM % or other CV.

Although our ICC values for inter-rater reliability were excellent they were lower compared to other studies examining inter-rater reliability. The SD in this study is smaller in every direction on both legs compared to the other studies, indicating that our sample group is more homogenous with the consequences as discussed earlier. In future testing in the cohort, it would be advisable to implement more practice trials prior to testing to reduce learning effect during testing.

5.1.3 Summary

We found excellent relative inter-rater reliability for the SEBT and the 1RM leg press test. The hamstrings flexibility test and the hip abduction strength test showed moderate to good inter-rater reliability. Modifications of the protocol with fixation belt for the hand-held dynamometer in the isometric hip abduction strength test, practice session for young players in the 1RM leg press test and more practice trials in the SEBT would probably increase our ICC values. Compared with other studies our SD from the different measurements was considerably smaller. This indicates that the sample group in this study was more homogenous

5.2 Methods

5.2.1 Sample group

The included participants in this study were recruited voluntarily among handball and football players playing at the elite level and participating in the yearly testing in the cohort study. This created a homogenous group, which we must take into account when comparing our results with other studies and to whom we can generalize the results.

There was no control of the activity the day before testing or between the two days of testing. This was a potential confounding factor as the football players were in the middle of their season playing matches. While the handball players were in their pre-season, usually a tough training period. This could have influenced the performance of the players in different ways. They could be more fatigued after match or hard training the previous day. Or they could have changed their motivation. It is however, most likely that the amount of training was stable during the testing period. The football players had normal in-season training, usual four to six training session and one match a week. While the handball players were in pre-season training, with stable amount of training, during the whole test period. This indicates that the preparation prior to testing most likely was not changed between the test sessions for either of the sports.

However it is likely that match one or two days prior to testing may have affected the players' potential to perform maximal. Studies have shown that participation in a match leads to acute fatigue characterized by a decline in physical performance over the following hours and days, both in elite female- (Andersson, Raastad et al. 2008) and elite male football players (Ispirlidis, Fatouros et al. 2008). The studies reported that, with individual differences, more than 72 hours are required to achieve pre-match values for physical performance, and to heal muscle damage and inflammation (Andersson, Raastad et al. 2008; Ispirlidis, Fatouros et al. 2008). Since the players in this study performed maximal strength and balance tasks, the performance could have reduced for players with match prior to testing. This can affect the results of the measurements if a player prior to the first test session participated in a match, while in the following retest session tested without having played match day before. For this study it would have been an advantage if the players had the same preparations the last days before both test sessions. In future testing of the cohort it would be advisable that the players did not participate in match or very hard training in the days before testing. This would reduce the risk of the players not performing maximal due to inadequate restitution.

5.2.2 Test team

The test team in this study had varied backgrounds as both sports sciences students and physiotherapists. They were all experienced tester and had all conducted testing earlier. Experienced testers though are not a guarantee for the quality of the measurements (Portney and Watkins 2009; Thomas, Silverman et al. 2011). In an ideal situation we should have examined the testers intra-rater reliability before we examined the inter-rater reliability. By examining the intra-rater reliability we can say something about the quality of each testers measurements on the specific test (Portney and Watkins 2009; Thomas, Silverman et al. 2011). Through adding one or more round of tests, with the same testers, on the first or the second test session we could have calculated intra-rater reliability for each of the testers. This, however, would become a very time consuming addition to the already long test sessions for the players and testers.

Even though we did not examine intra-tester reliability other actions were made to ensure the quality of the measurements. Prior to testing a pilot test was conducted to ensure consistency across the tester. A pilot test is a valuable opportunity to verify that instruments and procedures will function as specified on the group of participants that the research is intended to (Thomas, Silverman et al. 2011). In addition we teamed up the different testers on each screening test to try simultaneous assessment during the pilot testing, which helped us practicing the measurements.

Despite the lack of intra-tester assessment we believe that we have taken actions to ensure that our measurements are correct and the results are generalizable to similar groups of participant and testers.

5.2.3 Study design

The purpose of this study was to assess the inter-rater reliability of some of the tests included in the cohort. This is useful information for both the quality of the measurements in the cohort and for the possibility to generalize the results to other similar groups.

The conduction of the screening tests in this study was equal to the procedures in the cohort. It can be described as a semi standardized approach which describes the reliability, with a moderate level of standardization, which is possible to obtain in a clinical setting (Carter, Lubinsky et al. 2011). A design with one screening test only and better control on confounding factors like fatigue, food, match or training would possible have given an even better estimate of the reliability. However, there is positive aspect with a study design with a partially standardized approach. Even though it does not describe reliability as it is in a typical clinical setting, as a non-standardized approach does, it can still be achievable in a clinical setting (Carter, Lubinsky et al. 2011). If we had better control of the potential confounding factors our internal validity would have been increased, while our external validity and generalizability would have been decreased.

A test-retest session is the preferred method of assessing inter-rater reliability (Portney and Watkins 2009). Our study consisted of two testers assessing the same group in two different occasions. Unfortunately, we did not separate the scores to the different testers. This reduces the ability to compare results from each tester and detect any systematic error between the testers. However the testers had random selections of measurements on test session one and test session two. In this way eventually differences in motivation, fatigue or learning effects was randomly distributed between the testers. The majority of other studies assessing inter-rater reliability of isometric hip abductor strength, hamstrings flexibility or the SEBT used two testers as in our study (Hertel, Miller et al. 2000; Krause, Schlagel et al. 2007; Plisky, Gorman et al. 2009; Atamaz, Ozcaldiran et al. 2011; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012). Many of these studies have isolated the score for each tester and can therefore more precisely determine if any systematic error can be attributed to the testers (Shultz, Nguyen et al. 2006; Krause, Schlagel et al. 2007; Kelln, McKeon et al. 2008; Plisky, Gorman et al. 2009; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012). In this study it would have been an advantage if we had stored the measurements separately for each tester as in other studies examining inter-rater reliability of the isometric hip abductor strength, the hamstrings flexibility and the SEBT. This could have reviled any systematic differences of the measurements between the testers. The problem occurred when we had to carry out a second round of reliability testing four to five weeks later. Some of the testers were unable to participate in this testing, and replacements familiar with the procedures took their places.

5.2.4 Order of the screening tests

The screening tests assessed in this study were four out of a series of several tests used in the cohort. In the first test session the players completed all the tests in the cohort, while they in the re-test session only completed the tests included in this study. Even though the order of the screening tests were randomly, it is possible that one test could affect the performance in the following test. The tests included in this study places demands for the players to active perform maximum reach or maximum strength. The need for maximum performance may lead to a potential problem. For example when the players had conducted the maximal leg press and after a short break started the SEBT. The maximal leg press will likely increase the tension or stiffness in the muscle, which can affect the ability to perform maximal on the SEBT. In studies assessing the effect of combined training the authors concluded that due to lack of restitution from the previous training the subjects might be unable to perform maximal on the next training (Leveritt, Abernethy et al. 1999; Leveritt, Abernethy et al. 2003). Although the studies have assessed strength and endurance training, is not unlikely that factors like restitution and muscle properties could possible influence the results when the players perform several of these tests during the same day. The players could possible, be able to perform the tests with more consistency if this study only was designed to test and assess the inter-rater reliability of *one test only*. In future testing of the cohort this factor can be reduced if the players' order of the screening tests is randomized.

Another difference between the two test sessions was the time used during testing. The first test session was, as described, a seven-hour session. The retest session however, was a three-hour session. That can possible affect the results as the players on day two used considerably less time for testing. Potentially the players could have increased fatigued or reduced motivation late on the first test session compared to the retest session. This can potentially contribute to the systematic error with higher scores on the retest session, due to increased motivation in the retest session, reduced motivation or increased fatigue during first test session.

Our study is however not the only study which have assessed several different tests in one single study. Other studies examining inter-rater reliability of isometric hip abductor strength and the hamstrings flexibility were also a part of larger set of tests (Shultz, Nguyen et al. 2006; Krause, Schlagel et al. 2007; Kelln, McKeon et al. 2008; Atamaz, Ozcaldiran et al. 2011; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012). In the same way as in our study, the performances of one screening test can possible affect the performance of the following screening test.

5.2.5 Data collection

The experimental mortality may be a confounding factor, but is beyond the direct control of the researcher (Thomas, Silverman et al. 2011). As shown in figure 2, the drop out prior test one was twelve players with injuries or sickness and three players without reason. Between the two test-sessions another eight players dropped out due to illness or injury and four players were away on international matches. Potentially this could have led to that the remaining players may be unique from the standpoint of health, interest, motivation or other factors (Berg and Latin 2004). Since we were in no position of monitoring the players' health during testing we were unable to verify if they really were injured or sick. Potentially the players who performed poor on the first test session dropped out because they felt their performance were not good enough compared to the other players. These conditions may have created a sample group, which is not representative for elite female football- and handball players. This would affect both the results and the generalizability of the results.

5.2.6 Test- retest interval

There was a range of two days to seven weeks between the test and retest sessions. However, a predominance of the players did both test sessions within six days. Ideally the interval in a test-retest session examining reliability of a measurement cannot be so long that an actual change has happened (Thomas, Silverman et al. 2011). On the other hand a short interval could make the players tired after the first test session (Thomas, Silverman et al. 2011). Although the majority of the players in our study performed the tests with the same interval, potentially the players with the longest interval between tests could have had an actual change in the attributes measured. This factor might also explain some of the systematic change from the first test session to the retest session. Sixteen of the players had their retest six-seven weeks after the first test session. The football players were in season and trained normally, with focus on maintaining physical attributes. During season there is unlikely that the players will increase significantly in physical attributes, thus the football players with long intervals between tests will probably not have affect the results too much. The handball players were in a pre-season training phase involving intensive training. Since we measured maximal strength there is a potential for the players to have obtained greater strength during the test-retest session. Adaptation to resistance exercise varies between subjects and with the strength status of the individuals (Raastad, Paulsen et al. 2010). In general a trained person will need more time to gain muscle strength and the progress is smaller than in untrained person (Kraemer, Adams et al. 2002; Hubal, Gordish-Dressman et al. 2005). In our study we did not have any information of the players strength status, however we knew that the handball players had played at the elite level for three years, on average. If we assume that players who had played at the elite level for several seasons had a good basis of strength, there is an even smaller chance for the players to have obtained an actual change in muscle strength during the test-retest session. Therefore we believe that the gap between tests for some of the handball players will not affect our results considerably.

Of other studies which have examined the inter-rater reliability of the isometric hip abductor strength, the hamstrings flexibility or the SEBT four studies had the test-retest session within seven to ten days (Hertel, Miller et al. 2000; Shultz, Nguyen et al. 2006; Munro and Herrington 2010; Atamaz, Ozcaldiran et al. 2011) four studies had both sessions the same day (Krause, Schlagel et al. 2007; Plisky, Gorman et al. 2009; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012) while Kelln et al. did not provide any information on the interval between the sessions (Kelln, McKeon et al. 2008). Both leg press and isometric hip abduction strength were assessed with 1RM. After a traditional strength exercise training the muscle needs one to seven days of recovery before restoring muscle function (Raastad, Risoy et al. 2003), while explosive or maximal strength exercise training needs hours to three days (Linnamo, Hakkinen et al. 1998). Our design with the majority of the players performing both test session within six days is therefore likely more preferable of measuring maximal muscle strength compared to the studies that conducted two sessions the same day (Krause, Schlagel et al.

2007; Thorborg, Bandholm et al. 2011; Thorborg, Bandholm et al. 2012). The most advantageous interval to ensure no change between the sessions or to short time for restitution is probably seen for the studies performing the tests within seven to ten days.

5.3 Choice of statistical methods

5.3.1 Relative reliability

In measures where the systematic error is large, the choice of ICC model will affect the size of the ICC with model 1 > model 2 > model 3 (Portney and Watkins 2009). In these situations a comparison between different models must be done with caution. In our measurements we reviled small amounts of systematic error in most of the tests. The systematic error varied from 0.8-7.9 % for the strength measures, 1.3-2.0 % for the hamstrings flexibility and 1.7-3.0 % for the SEBT. There is no reference value for the accepted size of the systematic error before it affects the size of the ICC. We believe that our ICC values most likely were comparable to other ICC values, independent of which of the three models being used due to the relative low amount of systematic error in the measurements.

The use of single measure or average measures also affects the ICC value. An ICC value calculated with the mean of two or more single measures, $ICC_{(2.k)}$ or $ICC_{(3.k)}$ gives a larger ICC value compared to a calculation using a single measure like $ICC_{(1.1)}$, $ICC_{(2.1)}$ or $ICC_{(3.1)}$ (Shrout and Fleiss 1979; Portney and Watkins 2009). In this study we presented ICC values for both single measures $ICC_{(2.1)}$ and average measures $ICC_{(2.k)}$. To ensure no misinterpretation this was clearly marked in our calculation. Since we have presented both $ICC_{(2.k)}$ and $ICC_{(2.1)}$ for most of our measurements, comparing of the ICC values is possible. Our ICC values were consistently lower than most of the other studies assessing inter-rater reliability of the isometric hip abductor strength, the hamstrings flexibility and the SEBT. An important factor could be the homogeneity of the sample group, which has already been discussed as one of the factors affecting the ICC value negatively.

An other way of increasing the reliability is to reduce the amount of systematic- and random error of the measurements (Streiner and Norman 2008; Portney and Watkins 2009). As discussed in section 5.1 there are some possible modifications we could have done to reduce the risk of both types of error.

5.3.2 Absolute reliability

Our calculation of SEM was based on the random error $SEM = \sqrt{residual\ mean\ square}$. This calculation is independent of the choice of ICC model, the range of the variation and can easier be compared across studies (Weir 2005). Some of the other studies assessing inter-rater reliability of the isometric hip abductor strength, hamstrings flexibility and the SEBT calculated $SEM = SD\sqrt{1 - ICC}$ with the use of the reliability coefficient and SD (Shultz, Nguyen et al. 2006; Munro and Herrington 2010; Thorborg, Bandholm et al. 2012). This calculation of SEM is more influenced by the heterogeneity of the sample group because characterizes within-subject variability, not between subject variation as in the other equation (Weir 2005). For other studies examining inter-rater reliability of the hip abductor strength and the SEBT there lack information in which method SEM was calculated. In both cases this means comparing of the SEM values must not be don uncritically due to uncertainty of the influence from systematic errors.

In our study we presented the absolute reliability with SEM and visually in Bland Altman plots. This is however not necessary since they are closely linked and show approximately the same (Hopkins 2000; Portney and Watkins 2009). Regardless we wanted to present the plots to add a visual demonstration in addition to the SEM values presented in the tables.

5.4 Clinical implication of the findings

In this study we have examined the reliability of a series of test used in the screening of risk factors for ACL injury. This has given us the opportunity to more closely investigate the test procedures in the cohort study. By conducting this reliability testing we can present in which degree the measurements are correct/reliable. In a larger context this study will give an indication of the quality of the measurements in the on-going cohort. Hopefully the on-going cohort will be able to present results of any risk factors for ACL injury. By conducting this study we were able examine the inter-rater reliability of the isometric hip abductor strength test, 1RM leg press, the hamstrings flexibility test and the SEBT on a similar population as in the cohort. We may than use the screening tests to reveal any potential risk factors found in the cohort and target more specific prevention programs for the players in risk of injuries, knowing the measures are reliable. It is proposed that more specified and targeted prevention programs may be important to increase the compliance of the prevention programs (Renstrom, Ljungqvist et al. 2008). An increased compliance of the prevention programs can potentially reduce the incidence of ACL injuries (Alentorn-Geli, Myer et al. 2009b).

From our results it seems like the greatest variation between tests was found in the isometric hip abductor strength measures and the hamstrings flexibility measures. For the isometric hip abductor strength tests more trials and some form of familiarization for the players prior testing could reduce the risk of systematic error. It also would be useful with external fixation of the hand-held dynamometer to reduce the risk error from upper extremity strength among testers.

The SEBT and 1RM leg press showed the highest inter-rater reliability with low absolute reliability and high relative reliability. For both tests there was a small systematic error with higher score on test two. For the SEBT a change in the protocol with more practice trials prior maximal reach would probably decrease the systematic error. For the 1RM leg press it seems important to conduct trial sessions prior testing, especially with the youngest players and players who are not familiar with leg press exercise.

However any possible modifications of the protocols used in the cohort cannot be implemented uncritically. Modifications of the protocols in the cohort may influence the already collected data when comparing with data from the modified tests. A possibility is to compare the tests, before and after modifications, to examine if they show the same measurements.

Although we presented reliable results for strength-, hamstrings flexibility measures and the SEBT in a selected sample group of elite handball- and football players, these results must not uncritically be transferred to other sample groups. A measuring method is not reliable, but is closely linked to the sample group and conditions surrounding the testing (Streiner and Norman 2008; Portney and Watkins 2009).

6.0 Conclusion

The isometric hip abductor strength test, the 1RM leg press test, the hamstring flexibility test and the SEBT proved to be reliable on our cohort of female handball- and football players. The ICC score varied from moderate to excellent, with SEBT and leg press showing excellent inter-rater reliability. The isometric hip abductor strength test and the hamstrings flexibility test showed moderate to good inter-rater reliability. Compared to the mean of the different measurements the SEM was acceptable and SEM% showed the greatest individual variations were in the strength measures.

References

Abernethy, P., G. Wilson, et al. (1995). "Strength and power assessment. Issues, controversies and challenges." Sports Med **19**(6): 401-417.

Adachi, N., K. Nawata, et al. (2008). "Relationship of the menstrual cycle phase to anterior cruciate ligament injuries in teenaged female athletes." Arch Orthop Trauma Surg **128**(5): 473-478.

Alentorn-Geli, E., G. D. Myer, et al. (2009a). "Prevention of non-contact anterior cruciate ligament injuries in soccer players. Part 1: Mechanisms of injury and underlying risk factors." Knee Surg Sports Traumatol Arthrosc **17**(7): 705-729.

Alentorn-Geli, E., G. D. Myer, et al. (2009b). "Prevention of non-contact anterior cruciate ligament injuries in soccer players. Part 2: a review of prevention programs aimed to modify risk factors and to reduce injury rates." Knee Surg Sports Traumatol Arthrosc **17**(8): 859-879.

Andersson, H., T. Raastad, et al. (2008). "Neuromuscular fatigue and recovery in elite female soccer: effects of active recovery." Med Sci Sports Exerc **40**(2): 372-380.

Arendt, E. and R. Dick (1995). "Knee injury patterns among men and women in collegiate basketball and soccer. NCAA data and review of literature." Am J Sports Med **23**(6): 694-701.

Arendt, E. A., B. Bershadsky, et al. (2002). "Periodicity of noncontact anterior cruciate ligament injuries during the menstrual cycle." J Gend Specif Med **5**(2): 19-26.

Askling, C. M., J. Nilsson, et al. (2010). "A new hamstring test to complement the common clinical examination before return to sport after injury." Knee Surg Sports Traumatol Arthrosc **18**(12): 1798-1803.

Atamaz, F., B. Ozcaldiran, et al. (2011). "Interobserver and intraobserver reliability in lower-limb flexibility measurements." J Sports Med Phys Fitness **51**(4): 689-694.

Bahr, R. and T. Krosshaug (2005). "Understanding injury mechanisms: a key component of preventing injuries in sport." Br J Sports Med **39**(6): 324-329.

Ben, M. and L. A. Harvey (2010). "Regular stretch does not increase muscle extensibility: a randomized controlled trial." Scand J Med Sci Sports **20**(1): 136-144.

Bent, N. P., C. C. Wright, et al. (2009). "Selecting outcome measures in sports medicine: a guide for practitioners using the example of anterior cruciate ligament rehabilitation." Br J Sports Med **43**(13): 1006-1012.

Berg, K. E. and R. W. Latin (2004). "Essentials of research methods in health, physical education, exercise science, and recreation." (2nd ed.): XVI, 292 s. : ill.

Besier, T. F., D. G. Lloyd, et al. (2001). "Anticipatory effects on knee joint loading during running and cutting maneuvers." Med Sci Sports Exerc **33**(7): 1176-1181.

Beyer, N. and P. Magnusson (2003). Målemetoder i fysioterapi. K¯benhavn, Munksgaard.

Blackburn, J. T. and D. A. Padua (2008). "Influence of trunk flexion on hip and knee joint kinematics during a controlled drop landing." Clin Biomech (Bristol, Avon) **23**(3): 313-319.

Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement." Lancet **1**(8476): 307-310.

Boden, B. P., G. S. Dean, et al. (2000). "Mechanisms of anterior cruciate ligament injury." Orthopedics **23**(6): 573-578.

Brandon, M. L., P. T. Haynes, et al. (2006). "The association between posterior-inferior tibial slope and anterior cruciate ligament insufficiency." Arthroscopy **22**(8): 894-899.

Brophy, R., H. J. Silvers, et al. (2010). "Gender influences: the role of leg dominance in ACL injury among soccer players." Br J Sports Med **44**(10): 694-697.

Brophy, R. H., H. J. Silvers, et al. (2010). "Anterior cruciate ligament injuries: etiology and prevention." Sports Med Arthrosc **18**(1): 2-11.

Brukner, P. and K. A. A. Khan (2012). Brukner & Khan's clinical sports medicine. Sydney, Mc Graw-Hill.

Carter, R. E., J. Lubinsky, et al. (2011). Rehabilitation research: principles and applications. St. Louis, Miss., Elsevier Saunders.

Chandrashekar, N., J. Slauterbeck, et al. (2005). "Sex-based differences in the anthropometric characteristics of the anterior cruciate ligament and its relation to intercondylar notch geometry: a cadaveric study." Am J Sports Med **33**(10): 1492-1498.

Chappell, J. D., R. A. Creighton, et al. (2007). "Kinematics and electromyography of landing preparation in vertical stop-jump: risks for noncontact anterior cruciate ligament injury." Am J Sports Med **35**(2): 235-241.

Chappell, J. D., D. C. Herman, et al. (2005). "Effect of fatigue on knee kinetics and kinematics in stop-jump tasks." Am J Sports Med **33**(7): 1022-1029.

Conley, S., A. Rosenberg, et al. (2007). "The female knee: anatomic variations." J Am Acad Orthop Surg **15 Suppl 1**: S31-36.

Cowling, E. J. and J. R. Steele (2001). "Is lower limb muscle synchrony during landing affected by gender? Implications for variations in ACL injury rates." J Electromyogr Kinesiol **11**(4): 263-268.

Dahl, H. A. and E. Rinvik (2010). Menneskets funksjonelle anatomi: med hovedvekt på bevegelsesapparatet. [Oslo], Cappelen akademisk.

Dallinga, J. M., A. Benjaminse, et al. (2012). "Which screening tools can predict injury to the lower extremities in team sports?: a systematic review." Sports Med **42**(9): 791-815.

Dienst, M., G. Schneider, et al. (2007). "Correlation of intercondylar notch cross sections to the ACL size: a high resolution MR tomographic in vivo analysis." Arch Orthop Trauma Surg **127**(4): 253-260.

Duthon, V. B., C. Barea, et al. (2006). "Anatomy of the anterior cruciate ligament." Knee Surg Sports Traumatol Arthrosc **14**(3): 204-213.

Ekstrand, J., T. Timpka, et al. (2006). "Risk of injury in elite football played on artificial turf versus natural grass: a prospective two-cohort study." Br J Sports Med **40**(12): 975-980.

Elliot, D. L., L. Goldberg, et al. (2010). "Young women's anterior cruciate ligament injuries: an expanded model and prevention paradigm." Sports Med **40**(5): 367-376.

Engebretsen, L. and R. Bahr (2009). Sports injury prevention. Chichester, Wiley-Blackwell.

Faude, O., A. Junge, et al. (2006). "Risk factors for injuries in elite female soccer players." Br J Sports Med **40**(9): 785-790.

Fauno, P. and B. Wulff Jakobsen (2006). "Mechanism of anterior cruciate ligament injuries in soccer." Int J Sports Med **27**(1): 75-79.

Fening, S. D., J. Kovacic, et al. (2008). "The effects of modified posterior tibial slope on anterior cruciate ligament strain and knee kinematics: a human cadaveric study." J Knee Surg **21**(3): 205-211.

Ferretti, A., P. Papandrea, et al. (1992). "Knee ligament injuries in volleyball players." Am J Sports Med **20**(2): 203-207.

Ford, K. R., G. D. Myer, et al. (2003). "Valgus knee motion during landing in high school female and male basketball players." Med Sci Sports Exerc **35**(10): 1745-1750.

Franchignoni, F., L. Tesio, et al. (1998). "Reliability of four simple, quantitative tests of balance and mobility in healthy elderly females." Aging (Milano) **10**(1): 26-31.

Fredriksen, H., H. Dagfinrud, et al. (1997). "Passive knee extension test to measure hamstring muscle tightness." Scand J Med Sci Sports **7**(5): 279-282.

Fuller, C. W., R. W. Dick, et al. (2007). "Comparison of the incidence, nature and cause of injuries sustained on grass and new generation artificial turf by male and female football players. Part 1: match injuries." Br J Sports Med **41 Suppl 1**: i20-26.

Fuller, C. W., R. W. Dick, et al. (2007). "Comparison of the incidence, nature and cause of injuries sustained on grass and new generation artificial turf by male and female football players. Part 2: training injuries." Br J Sports Med **41 Suppl 1**: i27-32.

Giuliani, J. R., K. G. Kilcoyne, et al. (2009). "Anterior cruciate ligament anatomy: a review of the anteromedial and posterolateral bundles." J Knee Surg **22**(2): 148-154.

Goodfellow, J. and J. O'Connor (1978). "The mechanics of the knee and prosthesis design." J Bone Joint Surg Br **60-B**(3): 358-369.

Gribble, P. A., J. Hertel, et al. (2012). "Using the Star Excursion Balance Test to assess dynamic postural-control deficits and outcomes in lower extremity injury: a literature and systematic review." J Athl Train **47**(3): 339-357.

Griffin, L. Y., M. J. Albohm, et al. (2006). "Understanding and preventing noncontact anterior cruciate ligament injuries: a review of the Hunt Valley II meeting, January 2005." Am J Sports Med **34**(9): 1512-1532.

Heiderscheit, B. C., J. Hamill, et al. (2000). "Influence of Q-angle on lower-extremity running kinematics." J Orthop Sports Phys Ther **30**(5): 271-278.

Heidt, R. S., Jr., S. G. Dormer, et al. (1996). "Differences in friction and torsional resistance in athletic shoe-turf surface interfaces." Am J Sports Med **24**(6): 834-842.

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

Hertel, J., R. A. Braham, et al. (2006). "Simplifying the star excursion balance test: analyses of subjects with and without chronic ankle instability." J Orthop Sports Phys Ther **36**(3): 131-137.

Hertel, J., S. Miller, et al. (2000). "Intratester and intertester reliability during the Star Excursion Balance Tests." J Sport Rehabil **9**: 104-116.

Hewett, T. E. (2000). "Neuromuscular and hormonal factors associated with knee injuries in female athletes. Strategies for intervention." Sports Med **29**(5): 313-327.

Hewett, T. E., G. D. Myer, et al. (2004). "Decrease in neuromuscular control about the knee with maturation in female athletes." J Bone Joint Surg Am **86-A**(8): 1601-1608.

Hewett, T. E., G. D. Myer, et al. (2005). "Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study." Am J Sports Med **33**(4): 492-501.

Hewett, T. E., A. L. Stroupe, et al. (1996). "Plyometric training in female athletes. Decreased impact forces and increased hamstring torques." Am J Sports Med **24**(6): 765-773.

Hewett, T. E., B. T. Zazulak, et al. (2007). "Effects of the menstrual cycle on anterior cruciate ligament injury risk: a systematic review." Am J Sports Med **35**(4): 659-668.

Hohmann, E., A. Bryant, et al. (2011). "Is there a correlation between posterior tibial slope and non-contact anterior cruciate ligament injuries?" Knee Surg Sports Traumatol Arthrosc **19 Suppl 1**: S109-114.

Hopkins, W. G. (2000). "Measures of reliability in sports medicine and science." Sports Med **30**(1): 1-15.

Hubal, M. J., H. Gordish-Dressman, et al. (2005). "Variability in muscle size and strength gain after unilateral resistance training." Med Sci Sports Exerc **37**(6): 964-972.

Ispirlidis, I., I. G. Fatouros, et al. (2008). "Time-course of changes in inflammatory and performance responses following a soccer game." Clin J Sport Med **18**(5): 423-431.

Jacobs, C. A., T. L. Uhl, et al. (2007). "Hip abductor function and lower extremity landing kinematics: sex differences." J Athl Train **42**(1): 76-83.

Jacobs, I., B. Schilling, et al. (2009). "American College of Sports Medicine position stand. Progression models in resistance training for healthy adults." Med Sci Sports Exerc **41**(3): 687-708.

Kelln, B. M., P. O. McKeon, et al. (2008). "Hand-held dynamometry: reliability of lower extremity muscle testing in healthy, physically active,young adults." J Sport Rehabil **17**(2): 160-170.

Kinzey, S. J. and C. W. Armstrong (1998). "The reliability of the star-excursion test in assessing dynamic balance." J Orthop Sports Phys Ther **27**(5): 356-360.

Koga, H., A. Nakamae, et al. (2010). "Mechanisms for noncontact anterior cruciate ligament injuries: knee joint kinematics in 10 injury situations from female team handball and basketball." Am J Sports Med **38**(11): 2218-2225.

Kraemer, W. J., K. Adams, et al. (2002). "American College of Sports Medicine position stand. Progression models in resistance training for healthy adults." Med Sci Sports Exerc **34**(2): 364-380.

Kramer, L. C., C. R. Denegar, et al. (2007). "Factors associated with anterior cruciate ligament injury: history in female athletes." J Sports Med Phys Fitness **47**(4): 446-454.

Krause, D. A., S. J. Schlagel, et al. (2007). "Influence of lever arm and stabilization on measures of hip abduction and adduction torque obtained by hand-held dynamometry." Arch Phys Med Rehabil **88**(1): 37-42.

Krosshaug, T., J. R. Slauterbeck, et al. (2007). "Biomechanical analysis of anterior cruciate ligament injury mechanisms: three-dimensional motion reconstruction from video sequences." Scand J Med Sci Sports **17**(5): 508-519.

LaPrade, R. F. and Q. M. Burnett, 2nd (1994). "Femoral intercondylar notch stenosis and correlation to anterior cruciate ligament injuries. A prospective study." Am J Sports Med **22**(2): 198-202; discussion 203.

Law, R. Y., L. A. Harvey, et al. (2009). "Stretch exercises increase tolerance to stretch in patients with chronic musculoskeletal pain: a randomized controlled trial." Phys Ther **89**(10): 1016-1026.

Lephart, S. M., C. M. Ferris, et al. (2002). "Gender differences in strength and lower extremity kinematics during landing." Clin Orthop Relat Res(401): 162-169.

Leveritt, M., P. J. Abernethy, et al. (2003). "Concurrent strength and endurance training: the influence of dependent variable selection." J Strength Cond Res **17**(3): 503-508.

Leveritt, M., P. J. Abernethy, et al. (1999). "Concurrent strength and endurance training. A review." Sports Med **28**(6): 413-427.

Lexell, J. E. and D. Y. Downham (2005). "How to assess the reliability of measurements in rehabilitation." Am J Phys Med Rehabil **84**(9): 719-723.

Linnamo, V., K. Hakkinen, et al. (1998). "Neuromuscular fatigue and recovery in maximal compared to explosive strength loading." Eur J Appl Physiol Occup Physiol **77**(1-2): 176-181.

Liu, W. and M. E. Maitland (2003). "Influence of anthropometric and mechanical variations on functional instability in the ACL-deficient knee." Ann Biomed Eng **31**(10): 1153-1161.

Lohmander, L. S., P. M. Englund, et al. (2007). "The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis." Am J Sports Med **35**(10): 1756-1769.

Loudon, J. K., W. Jenkins, et al. (1996). "The relationship between static posture and ACL injury in female athletes." J Orthop Sports Phys Ther **24**(2): 91-97.

Ly, L. P. and D. J. Handelsman (2002). "Muscle strength and ageing: methodological aspects of isokinetic dynamometry and androgen administration." Clin Exp Pharmacol Physiol **29**(1-2): 37-47.

Malinzak, R. A., S. M. Colby, et al. (2001). "A comparison of knee joint motion patterns between men and women in selected athletic tasks." Clin Biomech (Bristol, Avon) **16**(5): 438-445.

McGraw, K. and S. P. Wong (1996). "Forming inferences about some intraclass correlation coefficients." Psychological Methods **1**(1): 30-46.

McLean, S. G., R. E. Fellin, et al. (2007). "Impact of fatigue on gender-based high-risk landing strategies." Med Sci Sports Exerc **39**(3): 502-514.

McNair, P. J., R. N. Marshall, et al. (1990). "Important features associated with acute anterior cruciate ligament injury." N Z Med J **103**(901): 537-539.

Meunier, A., M. Odensten, et al. (2007). "Long-term results after primary repair or non-surgical treatment of anterior cruciate ligament rupture: a randomized study with a 15-year follow-up." Scand J Med Sci Sports **17**(3): 230-237.

Mizuno, Y., M. Kumagai, et al. (2001). "Q-angle influences tibiofemoral and patellofemoral kinematics." J Orthop Res **19**(5): 834-840.

Moller Nielsen, J. and M. Hammar (1991). "Sports injuries and oral contraceptive use. Is there a relationship?" Sports Med **12**(3): 152-160.

Monnier, A., J. Heuer, et al. (2012). "Inter- and intra-observer reliability of clinical movement-control tests for marines." BMC Musculoskelet Disord **13**: 263.

Munro, A. G. and L. C. Herrington (2010). "Between-session reliability of the star excursion balance test." Phys Ther Sport **11**(4): 128-132.

Myer, G. D., J. L. Brent, et al. (2011). "Real-time assessment and neuromuscular training feedback techniques to prevent ACL injury in female athletes." Strength Cond J **33**(3): 21-35.

Myer, G. D., K. R. Ford, et al. (2009). "The relationship of hamstrings and quadriceps strength to anterior cruciate ligament injury in female athletes." Clin J Sport Med **19**(1): 3-8.

Myer, G. D., K. R. Ford, et al. (2008). "The effects of generalized joint laxity on risk of anterior cruciate ligament injury in young female athletes." Am J Sports Med **36**(6): 1073-1080.

Myklebust, G. and R. Bahr (2005). "Return to play guidelines after anterior cruciate ligament surgery." Br J Sports Med **39**(3): 127-131.

Myklebust, G., L. Engebretsen, et al. (2003). "Prevention of anterior cruciate ligament injuries in female team handball players: a prospective intervention study over three seasons." Clin J Sport Med **13**(2): 71-78.

Myklebust, G., S. Maehlum, et al. (1997). "Registration of cruciate ligament injuries in Norwegian top level team handball. A prospective study covering two seasons." Scand J Med Sci Sports **7**(5): 289-292.

Myklebust, G., S. Maehlum, et al. (1998). "A prospective cohort study of anterior cruciate ligament injuries in elite Norwegian team handball." Scand J Med Sci Sports **8**(3): 149-153.

Nyland, J. A., D. N. Caborn, et al. (1999). "Crossover cutting during hamstring fatigue produces transverse plane knee control deficits." J Athl Train **34**(2): 137-143.

Odensten, M. and J. Gillquist (1985). "Functional anatomy of the anterior cruciate ligament and a rationale for reconstruction." J Bone Joint Surg Am **67**(2): 257-262.

Olsen, O. E., G. Myklebust, et al. (2004). "Injury mechanisms for anterior cruciate ligament injuries in team handball: a systematic video analysis." Am J Sports Med **32**(4): 1002-1012.

Olsen, O. E., G. Myklebust, et al. (2003). "Relationship between floor type and risk of ACL injury in team handball." Scand J Med Sci Sports **13**(5): 299-304.

Orchard, J., H. Seward, et al. (1999). "Rainfall, evaporation and the risk of non-contact anterior cruciate ligament injury in the Australian Football League." Med J Aust **170**(7): 304-306.

Orchard, J., H. Seward, et al. (2001). "Intrinsic and extrinsic risk factors for anterior cruciate ligament injury in Australian footballers." Am J Sports Med **29**(2): 196-200.

Orchard, J. W., I. Chivers, et al. (2005). "Rye grass is associated with fewer non-contact anterior cruciate ligament injuries than bermuda grass." Br J Sports Med **39**(10): 704-709.

Petersen, W. and T. Zantop (2007). "Anatomy of the anterior cruciate ligament with regard to its two bundles." Clin Orthop Relat Res **454**: 35-47.

Plisky, P. J., P. P. Gorman, et al. (2009). "The reliability of an instrumented device for measuring components of the star excursion balance test." N Am J Sports Phys Ther **4**(2): 92-99.

Portney, L. G. and M. P. Watkins (2009). "Foundations of clinical research : applications to practice." (3rd ed.): XIX, 892 s. : ill.

Posthumus, M., A. V. September, et al. (2009). "The COL5A1 gene is associated with increased risk of anterior cruciate ligament ruptures in female participants." Am J Sports Med **37**(11): 2234-2240.

Prodromos, C. C., Y. Han, et al. (2007). "A meta-analysis of the incidence of anterior cruciate ligament tears as a function of gender, sport, and a knee injury-reduction regimen." Arthroscopy **23**(12): 1320-1325 e1326.

Renstrom, P., A. Ljungqvist, et al. (2008). "Non-contact ACL injuries in female athletes: an International Olympic Committee current concepts statement." Br J Sports Med **42**(6): 394-412.

Ruedl, G., P. Ploner, et al. (2009). "Are oral contraceptive use and menstrual cycle phase related to anterior cruciate ligament injury risk in female recreational skiers?" Knee Surg Sports Traumatol Arthrosc **17**(9): 1065-1069.

Raastad, T., G. Paulsen, et al. (2010). Styrketrening: i teori og praksis. Oslo, Gyldendal undervisning.

Raastad, T., B. A. Risoy, et al. (2003). "Temporal relation between leukocyte accumulation in muscles and halted recovery 10-20 h after strength exercise." J Appl Physiol **95**(6): 2503-2509.

Schickendantz, M. S. and G. G. Weiker (1993). "The predictive value of radiographs in the evaluation of unilateral and bilateral anterior cruciate ligament injuries." Am J Sports Med **21**(1): 110-113.

Scranton, P. E., Jr., J. P. Whitesel, et al. (1997). "A review of selected noncontact anterior cruciate ligament injuries in the National Football League." Foot Ankle Int **18**(12): 772-776.

Sell, T. C., C. M. Ferris, et al. (2007). "Predictors of proximal tibia anterior shear force during a vertical stop-jump." J Orthop Res **25**(12): 1589-1597.

Shambaugh, J. P., A. Klein, et al. (1991). "Structural measures as predictors of injury basketball players." Med Sci Sports Exerc **23**(5): 522-527.

Shelbourne, K. D., T. Gray, et al. (2009). "Incidence of subsequent injury to either knee within 5 years after anterior cruciate ligament reconstruction with patellar tendon autograft." Am J Sports Med **37**(2): 246-251.

Shrout, P. E. and J. L. Fleiss (1979). "Intraclass correlations: uses in assessing rater reliability." Psychol Bull **86**(2): 420-428.

Shultz, S. J., A. D. Nguyen, et al. (2006). "Intratester and intertester reliability of clinical measures of lower extremity anatomic characteristics: implications for multicenter studies." Clin J Sport Med **16**(2): 155-161.

Shultz, S. J., R. J. Schmitz, et al. (2010). "ACL Research Retreat V: an update on ACL injury risk and prevention, March 25-27, 2010, Greensboro, NC." J Athl Train **45**(5): 499-508.

Silvers, H. J. and B. R. Mandelbaum (2007). "Prevention of anterior cruciate ligament injury in the female athlete." Br J Sports Med **41 Suppl 1**: i52-59.

Slauterbeck, J. R., S. F. Fuzie, et al. (2002). "The Menstrual Cycle, Sex Hormones, and Anterior Cruciate Ligament Injury." J Athl Train **37**(3): 275-278.

Soderman, K., H. Alfredson, et al. (2001). "Risk factors for leg injuries in female soccer players: a prospective investigation during one out-door season." Knee Surg Sports Traumatol Arthrosc **9**(5): 313-321.

Sole, G., J. Hamren, et al. (2007). "Test-retest reliability of isokinetic knee extension and flexion." Arch Phys Med Rehabil **88**(5): 626-631.

Souryal, T. O. and T. R. Freeman (1993). "Intercondylar notch size and anterior cruciate ligament injuries in athletes. A prospective study." Am J Sports Med **21**(4): 535-539.

Stark, T., B. Walker, et al. (2011). "Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review." PM R **3**(5): 472-479.

Steffen, K., T. E. Andersen, et al. (2007). "Risk of injury on artificial turf and natural grass in young female football players." Br J Sports Med **41 Suppl 1**: i33-37.

Stratford, P. W. and C. H. Goldsmith (1997). "Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data." Phys Ther **77**(7): 745-750.

Streiner, D. L. and G. R. Norman (2008). Health measurement scales: a practical guide to their development and use. Oxford, Oxford University Press.

Sutton, K. M. and J. M. Bullock (2013). "Anterior cruciate ligament rupture: differences between males and females." J Am Acad Orthop Surg **21**(1): 41-50.

Tegnander, A., O. E. Olsen, et al. (2008). "Injuries in Norwegian female elite soccer: a prospective one-season cohort study." Knee Surg Sports Traumatol Arthrosc **16**(2): 194-198.

Teitz, C. C., B. K. Lind, et al. (1997). "Symmetry of the femoral notch width index." Am J Sports Med **25**(5): 687-690.

Terauchi, M., K. Hatayama, et al. (2011). "Sagittal alignment of the knee and its relationship to noncontact anterior cruciate ligament injuries." Am J Sports Med **39**(5): 1090-1094.

Thomas, J. R., S. J. Silverman, et al. (2011). Research methods in physical activity. Champaign, Ill., Human Kinetics.

Thorborg, K., T. Bandholm, et al. (2012). "Hip- and knee-strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable." Knee Surg Sports Traumatol Arthrosc.

Thorborg, K., T. Bandholm, et al. (2011). "Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength." Scand J Med Sci Sports.

Thorborg, K., J. Petersen, et al. (2010). "Clinical assessment of hip strength using a hand-held dynamometer is reliable." Scand J Med Sci Sports **20**(3): 493-501.

Thorborg, K., A. Serner, et al. (2011). "Hip adduction and abduction strength profiles in elite soccer players: implications for clinical evaluation of hip adductor muscle recovery after injury." Am J Sports Med **39**(1): 121-126.

Todd, M. S., S. Lalliss, et al. (2010). "The relationship between posterior tibial slope and anterior cruciate ligament injuries." Am J Sports Med **38**(1): 63-67.

Uhorchak, J. M., C. R. Scoville, et al. (2003). "Risk factors associated with noncontact injury of the anterior cruciate ligament: a prospective four-year evaluation of 859 West Point cadets." Am J Sports Med **31**(6): 831-842.

Verdijk, L. B., L. van Loon, et al. (2009). "One-repetition maximum strength test represents a valid means to assess leg strength in vivo in humans." J Sports Sci **27**(1): 59-68.

Vincent, W. J. and J. Weir (2012). "Statistics in kinesiology." (4th ed.): xiv, 378 s. : ill.

Wadsworth, C., D. H. Nielsen, et al. (1992). "Interrater reliability of hand-held dynamometry: effects of rater gender, body weight, and grip strength." J Orthop Sports Phys Ther **16**(2): 74-81.

Walden, M., M. Hagglund, et al. (2006). "High risk of new knee injury in elite footballers with previous anterior cruciate ligament injury." Br J Sports Med **40**(2): 158-162; discussion 158-162.

Walden, M., M. Hagglund, et al. (2011). "The epidemiology of anterior cruciate ligament injury in football (soccer): a review of the literature from a gender-related perspective." Knee Surg Sports Traumatol Arthrosc **19**(1): 3-10.

Weir, J. P. (2005). "Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM." J Strength Cond Res **19**(1): 231-240.

Wikholm, J. B. and R. W. Bohannon (1991). "Hand-held Dynamometer Measurements: Tester Strength Makes a Difference." J Orthop Sports Phys Ther **13**(4): 191-198.

Withrow, T. J., L. J. Huston, et al. (2006). "The relationship between quadriceps muscle force, knee flexion, and anterior cruciate ligament strain in an in vitro simulated jump landing." Am J Sports Med **34**(2): 269-274.

Wojtys, E. M., L. J. Huston, et al. (2002). "The effect of the menstrual cycle on anterior cruciate ligament injuries in women as determined by hormone levels." Am J Sports Med **30**(2): 182-188.

Wojtys, E. M., L. J. Huston, et al. (1996). "Neuromuscular adaptations in isokinetic, isotonic, and agility training programs." Am J Sports Med **24**(2): 187-192.

Wright, R. W. (2009). "Knee injury outcomes measures." J Am Acad Orthop Surg **17**(1): 31-39.

Xerogeanes, J. W., Y. Takeda, et al. (1995). "Effect of knee flexion on the in situ force distribution in the human anterior cruciate ligament." Knee Surg Sports Traumatol Arthrosc **3**(1): 9-13.

Yu, B., C. F. Lin, et al. (2006). "Lower extremity biomechanics during the landing of a stop-jump task." Clin Biomech (Bristol, Avon) **21**(3): 297-305.

Zazulak, B. T., P. L. Ponce, et al. (2005). "Gender comparison of hip muscle activity during single-leg landing." J Orthop Sports Phys Ther **35**(5): 292-299.

Tables

Figures

Appendix

Appendix 1

Lag: Example

	Marker placement		3D		Anthropometric		Isokinetic
15:00	Player 1	15:00		15:00	Player 11	15:00	Player 9
	Player 2		Player 1		Player 12		Player 10
15:55	Player 3	15:55	Player 2	15:55	Player 13	15:55	Player 11
	Player 4		Player 3		Player 14		Player 12
16:50	Player 5	16:50	Player 4	16:50	Player 1	16:50	Player 13
17:17	BREAK		Player 5		Player 2		Player 14
17:40	Player 6	17:45	BREAK	17:45	BREAK	17:45	BREAK
18:05	Player 7	18:05	Player 6	18:05	Player 3	18:05	Player 1
	Player 8		Player 7		Player 4		Player 2
19:00	Player 9	19:00	Player 8	19:00	Player 5	19:00	Player 3
	Player 10		Player 9		Player 6		Player 4
19:55	Player 11	19:55	Player 10	19:55	Player 7	19:55	Player 5
	Player 12		Player 11		Player 8		Player 6
20:50	Player 13	20:50	Player 12	20:50	Player 9	20:50	Player 7
	Player 14		Player 13		Player 10		Player 8
		21:45	Player 14	21:45		21:45	

Subjective		Anatomic		Legpress/ balance
Player 7	15:00	Player 5	15:00	Player 3
Player 8		Player 6		Player 4
Player 9	15:55	Player 7	15:55	Player 5
Player 10		Player 8		Player 6
Player 11	16:50	Player 9	16:50	Player 7
Player 12		Player 10		Player 8
BREAK	17:45	BREAK	17:45	BREAK
Player 13	18:05	Player 11	18:05	Player 9
Player 14		Player 12		Player 10
Player 1	19:00	Player 13	19:00	Player 11
Player 2		Player 14		Player 12
Player 3	19:55	Player 1	19:55	Player 13
Player 4		Player 2		Player 14
Player 5	20:50	Player 3	20:50	Player 1
Player 6		Player 4		Player 2
	21:45		21:45	

Appendix 2

Date and team:

	Marker placement		3D		Anthropometric
09:00	Player A	09:00		09:00	Player B
09:30	Player B	09:30	Player A	09:30	Player C
10:00	Player C	10:00	Player B	10:00	Player D
10:30	Player D	10:30	Player C	10:30	Player E
11:00	Player E	11:00	Player D	11:00	Player F
11:30	Player F	11:30	Player E	11:30	Player A
12:00		12:00	Player F	12:00	

	Balance		Anatomic		Legpress
09:00		09:00	Player E	09:00	Player D
09:30	Player D	09:30	Player F	09:30	Player E
10:00	Player E	10:00	Player A	10:00	Player F
10:30	Player F	10:30	Player B	10:30	Player A
11:00	Player A	11:00	Player C	11:00	Player B
11:30	Player B	11:30	Player D	11:30	Player C
12:00	Player C	12:00		12:00	

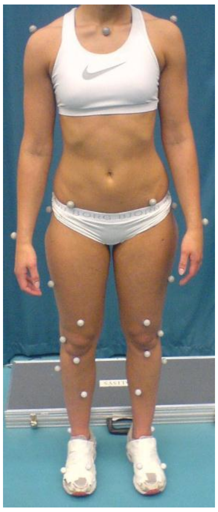
Appendix 3

SENTER FOR
Idrettsskadeforskning
KLOKE AV SKADE

Forskningsprosjekt blant fotballspillere i Toppserien 2012

Senter for idrettsskadeforskning ved Norges idrettshøgskole gjennomfører et forskningsprosjekt der vi undersøker risikofaktorer for korsbåndskader blant kvinnelige elitefotballspillere. Vi har derfor hver sesong siden 2009 testet alle spillere i Toppserien, og har nå totalt testet 320 spillere. Spillerne følges deretter opp de kommende sesongene i form av å registrere eventuelle korsbåndskader som oppstår.

Vi har nå satt av tid til testing av spillere fra **Stabæk onsdag 15.februar kl.15.00**. Testingen foregår på Norges idrettshøgskole, og dere kan møte opp i resepsjonen ved hovedinngangen. Vi vil da ha et kort informasjonsmøte først, hvor vi også ber alle om å skrive under på en samtykkeerklæring for prosjektdeltakelsen.



Vi har totalt 7 teststasjoner som innebærer 3D bevegelsesanalyse av finter/vendinger og fallhopp/spenst, styrketester av forside/bakside lår og hofte, bevegelighet, balansetester, anatomiske målinger og en blodprøve. Testingen vil totalt ta ca. 6-7 timer, og dere vil selvfølgelig få en pause og mat og drikke underveis.

Under testingen har dere på treningstøy og de skoene dere vanligvis bruker til innetrening. For å gjøre testingen lettere bør dere bruke en shorts og t-skjorte. To av testene krever at hofter/hoftekam er tilgjengelig for markører (se bilde), så ta gjerne på en boksershorts, bikinitruse eller eventuelt en kort sykkelshorts til disse testene. Markørene vi bruker til bevegelsesanalysen festes med teip – **unngå derfor å bruke bodylotion** på testdagen.

For å se bilder fra testingen, kan dere finne dette på hjemmesiden til Senter for idrettskadeforskning under følgende link;
http://www.klokavskade.no/no/Nyhetsarkiv/Nyhetsarkiv-2009/Hvorfor-skader-fotballjenter-fremre-korsband/

Vi ser frem til å møte dere 15.februar.

Dersom dere har spørsmål i mellomtiden kan dere ta kontakt på telefon (99 22 44 69) eller e-post (agnethe.nilstad@nih.no).

Vennlig hilsen

Agnethe Nilstad
Fysioterapeut MSc, PhD-kandidat
Prosjektleder

Appendix 4

NORGES IDRETTSHØGSKOLE

SENTER FOR Idrettsskadeforskning KLOKE AV SKADE

FORESPØRSEL OM DELTAKELSE I PROSJEKTET: ”Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndball- og fotballspillere - En prospektiv kohortstudie”

Bakgrunn for undersøkelsen

Korsbåndsskader i fotball og håndball har i det siste vært et svært aktuelt tema, både i media og i forskningssammenheng. Dette skyldes først og fremst den relativt store hyppigheten av denne alvorlige skaden, spesielt blant kvinnelige utøvere, som ser ut til å skade seg 3-7 ganger hyppigere enn menn. Problemet så langt er imidlertid at vi vet for lite om risikofaktorene og skademekanismene for korsbåndskader. Denne informasjonen er viktig når vi forsøker å forebygge skader, både for å kunne vite hvem som vil ha størst glede av forebyggende trening og for å kunne utvikle mest mulig effektive treningsmetoder.

Senter for idrettsskadeforskning er en forskningsgruppe bestående av fysioterapeuter, kirurger og biomekanikere med kunnskap innen idrettsmedisin. Vår hovedmålsetting er å forebygge skader i norsk idrett, med spesiell satsning på fotball, håndball, ski og snowboard. Denne studien er en viktig brikke i arbeidet med å finne ut hvorfor noen får en korsbåndskade. Vi ønsker nå å undersøke ulike mulige risikofaktorer for korsbåndskader, for deretter å kartlegge hvem som får korsbåndskader de påfølgende sesongene.

Gjennomføring av undersøkelsen

Vi ønsker at du som elitespiller deltar i denne studien, og deltakelsen er frivillig. Testingen vil finne sted på Norges idrettshøgskole. I løpet av en dag vil vi gjennomføre ulike styrke-, balanse- og bevegelighetstester, anatomiske målinger, samt gjennomføre en bevegelsesanalyse av hvordan du finter, vender, hopper og lander. Undersøkelsen starter med en kort oppvarming, deretter får du festet små refleksmarkører på kroppen (35 stk totalt). Du vil så bli bedt om å gjennomføre tre finter/vendinger og tre fallhopp. Under disse øvelsene vil det være 8 infrarøde kamera som filmer markørene, samtidig som kreftene fra underlaget blir målt. Dataene fra markører, kraftplattform og anatomiske mål benyttes i en matematisk modell som gir ut leddkrefter og momenter. Disse kreftene/momentene gir oss informasjon om hvordan muskler og passive strukturer som leddbånd belastes.

Bevegelsesanalysen vil ta ca. 1,5 time, inkludert anatomiske målinger og påsetting av markører. De andre testene gjennomføres resten av tiden laget er på NIH, og totalt vil testene ta om lag åtte timer. I tillegg til disse testene vil du få utdelt et skjema, der vi spør om treningserfaring, tidligere skader, skade i familien, treningsmengde, menstruasjonsstatus og knefunksjon. Spørreskjemaet besvares i løpet av testdagen, og det vil ta ca. 30 min.

Behandling av testresultatene

Vi vil de neste tre sesongene følge opp alle lag og spillere som har deltatt på testing hos oss for å registrere alle korsbåndskader som oppstår.

Vi er også interessert i å kunne kontakte deg senere med tanke på oppfølgningsstudier. Dette kan f.eks. skje ved at du får tilsendt et spørreskjema. Av den grunn vil vi lagre resultatene fra testene og svarene på spørreskjemaet fram til 1.6.2017. Etter dette vil dataene bli anonymisert. Dataene vil bli behandlet konfidensielt, og kun i forskningsøyemed. Alle som utfører testingen og forskere som benytter dataene er underlagt taushetsplikt. Dersom du ikke ønsker å være med på etterundersøkelser, kan du reservere deg mot dette i samtykkeerklæringen. I så fall vil alle dine data bli anonymisert etter fire år.

Vi vil underveis i testingen ta videoopptak av dere som vi senere kan ønske å bruke i undervisnings- og formidlingssammenheng. Opptakene inkluderer situasjoner der dere kun har på shorts og sports-BH. Dersom dere ikke vil at deres opptak skal være aktuelle for slik bruk krysser dere av for det i samtykkeerklæringen.

Hva får du ut av det?

Vi kan ikke tilby noe honorar for oppmøtet, men vil dekke eventuelle reise- og matutgifter. I tillegg vil du få kopi av dine resultater fra styrketestene som gjennomføres i løpet av testdagen.

Angrer du?

Du kan selvfølgelig trekke deg fra forsøket når som helst uten å måtte oppgi noen grunn. Alle data som angår deg vil uansett bli anonymisert.

Spørsmål?

Ring gjerne til Tron Krosshaug, tlf.: 45 66 00 46 hvis du har spørsmål om prosjektet, eller send e-post til tron.krosshaug@nih.no.

NORGES IDRETTSHØGSKOLE

SENTER FOR Idrettsskadeforskning KLOKE AV SKADE

"Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndball- og fotballspillere - En prospektiv kohortstudie"

SAMTYKKEERKLÆRING

Jeg har mottatt skriftlig og muntlig informasjon om studien *Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndball- og fotballspillere - En prospektiv kohortstudie*. Jeg er klar over at jeg kan trekke meg fra undersøkelsen på et hvilket som helst tidspunkt.

- ☐ Jeg ønsker ikke å bli kontaktet etter endt karriere med tanke på oppfølgningsstudier
- ☐ Jeg ønsker ikke at video av meg skal brukes i undervisningssammenheng

Sted ……………………….. Dato ………………………..

....
Underskrift

....
Navn med blokkbokstaver

....
Adresse

....
Mobiltelefon

....
E-postadresse

Appendix 5

NORGES IDRETTSHØGSKOLE

SENTER FOR Idrettsskadeforskning KLOKE AV SKADE

FORESPØRSEL OM DELTAKELSE I PROSJEKTET:

"Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndball- og fotballspillere - En prospektiv kohortstudie"

Bakgrunn for undersøkelsen

Korsbåndsskader i fotball og håndball har i det siste vært et svært aktuelt tema, både i media og i forskningssammenheng. Dette skyldes først og fremst den relativt store hyppigheten av denne alvorlige skaden, spesielt blant kvinnelige utøvere, som ser ut til å skade seg 3-7 ganger hyppigere enn menn. Problemet så langt er imidlertid at vi vet for lite om risikofaktorene og skademekanismene for korsbåndskader. Denne informasjonen er viktig når vi forsøker å forebygge skader, både for å kunne vite hvem som vil ha størst glede av forebyggende trening og for å kunne utvikle mest mulig effektive treningsmetoder.

Senter for idrettsskadeforskning er en forskningsgruppe bestående av fysioterapeuter, kirurger og biomekanikere med kunnskap innen idrettsmedisin. Vår hovedmålsetting er å forebygge skader i norsk idrett, med spesiell satsning på fotball, håndball, ski og snowboard. Denne studien er en viktig brikke i arbeidet med å finne ut hvorfor noen får en korsbåndskade. Vi ønsker nå å undersøke ulike mulige risikofaktorer for korsbåndskader, for deretter å kartlegge hvem som får korsbåndskader de påfølgende sesongene.

Gjennomføring av undersøkelsen

Vi ønsker at du som elitespiller deltar i denne studien, og deltakelsen er frivillig. Testingen vil finne sted på Norges idrettshøgskole. I løpet av en dag vil vi gjennomføre ulike styrke-, balanse- og bevegelighetstester, anatomiske målinger, samt gjennomføre en bevegelsesanalyse av hvordan du finter, vender, hopper og lander. Undersøkelsen starter med en kort oppvarming, deretter får du festet små refleksmarkører på kroppen (35 stk totalt). Du vil så bli bedt om å gjennomføre tre finter/vendinger og tre fallhopp. Under disse øvelsene vil det være 8 infrarøde kamera som filmer markørene, samtidig som kreftene fra underlaget blir målt. Dataene fra markører, kraftplattform og anatomiske mål benyttes i en matematisk modell som gir ut leddkrefter og momenter. Disse kreftene/momentene gir oss informasjon om hvordan muskler og passive strukturer som leddbånd belastes.

Bevegelsesanalysen vil ta ca. 1,5 time, inkludert anatomiske målinger og påsetting av markører. De andre testene gjennomføres resten av tiden laget er på NIH, og totalt vil testene ta om lag åtte timer. I tillegg til disse testene vil du få utdelt et skjema, der vi spør om treningserfaring, tidligere skader, skade i familien, treningsmengde, menstruasjonsstatus og knefunksjon. Spørreskjemaet besvares i løpet av testdagen, og det vil ta ca. 30 min.

Behandling av testresultatene

Vi vil de neste tre sesongene følge opp alle lag og spillere som har deltatt på testing hos oss for å registrere alle korsbåndskader som oppstår.

Vi er også interessert i å kunne kontakte deg senere med tanke på oppfølgningsstudier. Dette kan f.eks. skje ved at du får tilsendt et spørreskjema. Av den grunn vil vi lagre resultatene fra testene og svarene på spørreskjemaet fram til 1.6.2017. Etter dette vil dataene bli anonymisert. Dataene vil bli behandlet konfidensielt, og kun i forskningsøyemed. Alle som utfører testingen og forskere som benytter dataene er underlagt taushetsplikt. Dersom du ikke ønsker å være med på etterundersøkelser, kan du reservere deg mot dette i samtykkeerklæringen. I så fall vil alle dine data bli anonymisert etter fire år.

Vi vil underveis i testingen ta videoopptak av dere som vi senere kan ønske å bruke i undervisnings- og formidlingssammenheng. Opptakene inkluderer situasjoner der dere kun har på shorts og sports-BH. Dersom dere ikke vil at deres opptak skal være aktuelle for slik bruk krysser dere av for det i samtykkeerklæringen.

Hva får du ut av det?

Vi kan ikke tilby noe honorar for oppmøtet, men vil dekke eventuelle reise- og matutgifter. I tillegg vil du få kopi av dine resultater fra styrketestene som gjennomføres i løpet av testdagen.

Angrer du?

Du kan selvfølgelig trekke deg fra forsøket når som helst uten å måtte oppgi noen grunn. Alle data som angår deg vil uansett bli anonymisert.

Spørsmål?

Ring gjerne til Tron Krosshaug, tlf.: 45 66 00 46 hvis du har spørsmål om prosjektet, eller send e-post til tron.krosshaug@nih.no.

NORGES IDRETTSHØGSKOLE

SENTER FOR Idrettsskadeforskning KLOKE AV SKADE

"Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndball- og fotballspillere - En prospektiv kohortstudie"

SAMTYKKEERKLÆRING

Jeg har mottatt skriftlig og muntlig informasjon om studien *Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndball- og fotballspillere - En prospektiv kohortstudie*. Jeg er klar over at jeg kan trekke meg fra undersøkelsen på et hvilket som helst tidspunkt.

- ☐ Jeg ønsker ikke å bli kontaktet etter endt karriere med tanke på oppfølgingsstudier
- ☐ Jeg ønsker ikke at video av meg skal brukes i undervisningssammenheng

Sted ……………………….. Dato ………………………..

.... ……………………………………………..

Underskrift spiller Underskrift foresatt

....

Navn med blokkbokstaver

....

Adresse

....

Mobiltelefon

....

E-postadresse

Appendix 6

UNIVERSITETET I OSLO
DET MEDISINSKE FAKULTET

Forsker dr.scient. Tron Krosshaug
Norges idrettshøgskole
Pb. 4014 Ullevål Stadion
0806 Oslo

Regional komité for medisinsk og helsefaglig forskningsetikk Sør-Øst A (REK Sør-Øst A)
Postboks 1130 Blindern
NO-0318 Oslo

Telefon: 22 84 46 66
Telefaks: 22 85 05 90
E-post: jorgen.hardang@medisin.uio.no

Dato: 15.12.08
Deres ref.:
Vår ref.: S-07078a

S-07078a Risikofaktorer for fremre korsbåndskader hos kvinnelige elitehåndballspillere - en prospektiv kohortstudie [2.2007.511]

Vi viser til skjema for protokolltillegg og endringer datert 3.12.08 vedlagt revidert informasjonsskriv.

Prosjektleder ønsker å utvide prosjektpopulasjonen til kvinnelige elitefotballspillere fra Toppserien i Norge (ca 240 spillere).

Komiteen godkjenner endringen slik den er beskrevet i skjema for protokolltillegg og endringer og videresender kopi av informasjonsskriv, endringsskjema samt komiteens vedtak til Helsedirektoratet for behandling av endring av biobanken.

Med vennlig hilsen

Kristian Hagestad
Fylkeslege cand.med., spes. i samf.med
Leder

Jørgen Hardang
Sekretær

Kopi: Helsedirektoratet, Postboks 7000, St. Olavs plass, 0130 Oslo