Physical Therapy in Sport 16 (2015) 107–113

Contents lists available at ScienceDirect

Physical Therapy in Sport

journal homepage: www.elsevier.com/ptsp

Original research

Test-retest reliability of seven common clinical tests for assessing lower extremity muscle flexibility in futsal and handball players

Antonio Cejudo [a], Pilar Sainz de Baranda [b], Francisco Ayala [c, d, *], Fernando Santonja [e, f]

[a] INACUA Sports Centre, Murcia, Spain
[b] Faculty of Sport Sciences and Physical Activity, University of Murcia, Spain
[c] Sports Research Centre, Miguel Hernández University of Elche, Alicante, Spain
[d] ISEN University Formation, Center Affiliate to the University of Murcia, Spain
[e] Faculty of Medicine, University of Murcia, Spain
[f] Department of Traumatology, V. de la Arrixaca University Hospital, Murcia, Spain

ARTICLE INFO

Article history:
Received 10 May 2013
Received in revised form
30 April 2014
Accepted 26 May 2014

Keywords:
Range of motion
Reproducibility
Muscle strain
Injury

ABSTRACT

Objective: To determine the inter-session reliability of seven lower limb muscle flexibility measures obtained from the passive hip extension test (PHE), passive hip flexion test (PHF), passive hip abduction test (PHA), passive straight leg raise test (PSLR), modified Thomas test (MTh), the ankle dorsi-flexion with knee extended (ADF_{KE}) and flexed (ADF_{KF}) tests.
Design: Repeated measures design.
Setting: Controlled laboratory environment.
Participants: 60 futsal and 30 handball players.
Main outcome measures: Reliability was examined through the change in the mean (ChM), standard error of measurement expressed in absolute values (SEM) and as a percentage of the mean score (%SEM), minimal detectable change (MDC_{95}), and intraclass correlation coefficient ($ICC_{2,k}$).
Results: The findings showed negligible ChM values for all the flexibility measures analysed (<1°). Furthermore, the SEM and MDC_{95} for each flexibility measure ranged from 1.3° to 2.5° and from 3.8° to 6.9°, respectively, with %SEM scores lower than 10% and ICC scores higher than 0.88.
Conclusions: Clinicians can be 95% confident that an observed change between 2 measures larger than 3.7°, 6.2°, 5.5°, 6.1°, 6.9°, 4.7°, and 5.0° for the flexibility measures obtained from the PHE, PHF, PHA, PSLR, MTh, ADF_{KE} and ADF_{KF}, respectively, likely indicates a real change in muscle flexibility.

1. Introduction

Clinicians and sports medicine practitioners routinely assess and monitor the flexibility of the major lower limb muscles, because it has been postulated that an inadequate level of flexibility is one of the most important risk factors for lower limb muscle injuries, particularly in competitive athletes (Witvrouw, Danneels, Asselman, D'Have, & Cambier, 2003; Worrel & Perrin, 1992). The passive hip extension test (PHE), passive hip flexion with knee flexed test (PHF), passive hip abduction test (PHA), passive straight leg raise test (PSLR), modified Thomas test (MTh) and the ankle dorsi-flexion with knee extended (ADF_{KE}) and flexed (ADF_{KF}) tests are probably the most widely used measurement methods to assess iliopsoas, gluteus, adductor, hamstring, rectus femoris, gastrocnemius and soleus muscle flexibility, respectively (Clarkson, 2000; Palmer & Epler, 2002). However, before these measurement methods can be used to identify athletes at increased risk of injury and establish progress from training and/or rehabilitation programmes, the validity and reliability of their outcomes must be determined (Hopkins, 2000).

Although the PHE, PHF, PHA, PSLR, MTh, ADF_{KE}, and ADF_{KF} are indirect measures of the major lower limb muscles flexibility, all of these tests have been considered appropriate by the most important American medical organizations (American Academy of Orthopaedic Surgeons, 1965; American Medical Association, 2001). Further, all of these tests have been included in many

* Corresponding author. Sports Research Centre, Miguel Hernández University of Elche, Avda. de la Universidad s/n., 03202 Elche, Alicante, Spain. Tel.: +34 965222401; fax: +34 965222409.
E-mail address: Franciscoayalarodriguez@gmail.com (F. Ayala).

http://dx.doi.org/10.1016/j.ptsp.2014.05.004
1466-853X/

prominent sports medicine textbooks (Kendall, McCreary, Provance, Rodgers, & Romani, 2005; Magee, 2002; Prentice, 2003).

In clinical assessments, the reliability of a measure is determined by human factors (the tester's experience or training in administering the test, variations in assessment methodology and participant-related variability) and/or the instrument used. Based on the fact that the most widely used instruments to estimate muscle flexibility (goniometer and inclinometer) have been shown to be reliable (Clapper & Wolf, 1988; Tousignant, Boucher, Bourbonnais, Gravelle, Quesnel, & Brosseau, 2001), then the reliability of the lower limb muscle flexibility measures depend mainly on human factors (Hayen, Dennis, & Finch, 2007). Two different aspects of human reliability-related factors should be discussed before considering a measure appropriate for research and clinical purposes: inter-tester reliability and intra-tester reliability (Kottner et al., 2011).

Inter-tester reliability provides information regarding the degree to which measurements taken by different testers are similar. The inter-tester reliability of the lower limb muscle flexibility measures has been examined, yielding moderate to high scores (Aalto, Airaksinen, Härkönen, & Arokoski, 2005; Bennell, Talbot, Wajswelner, Techovanich, & Kelly, 1998; Boland & Adams, 2000; Clapis, Davis, & Davis, 2007; Gabbe, Bennell, Wajswelner, & Finch, 2004; Piva, Fitzgerald, Irrgang, Jones, Hando, Browder, & Childs, 2006).

Intra-tester reliability provides information regarding the degree to which several measurements taken at different times for the same test by the same tester are similar (Hayen et al., 2007). The intra-tester reliability can be determined using short (generally less than 3 h: intra-session) or long (generally more than 24 h: inter-session) time intervals to separate the testing sessions. Few studies have assessed the intra-tester reliability using short time intervals (<3 h) of the major lower limb muscle flexibility measures, showing high reliability scores (Aalto et al., 2005; Dennis, Finch, Elliott, & Farhart, 2008; Gnat, Kuszewski, Koczar, & Dziewońska, 2010; O'Shea & Grafton, 2013; Simondson, Brock, & Cotton, 2012).

Longer periods of time between assessments (e.g., two weeks) are very important in clinical and sports contexts, because these enable physical therapists to monitor the performance or health of their athletes and to make knowledgeable decisions regarding whether a real change has occurred between testing sessions after application of a training treatment (Hopkins, 2000). Nevertheless, very few studies have assessed the intra-tester reliability of the measures obtained from the PHE, PSLR and MTh using long time intervals (>24 h) in non-athletes (Aalto et al., 2005; Ayala, Sainz de Baranda, De Ste Croix, & Santonja, 2012; Bozic, Pazin, Berjan, Planic, & Cuk, 2010, Gabbe et al., 2004; Peeler & Anderson, 2008). Surprisingly, no studies have examined the reliability for the PHF, PHA, ADF_{KF} and ADF_{KE} tests. In addition, based on the fact that the reliability scores of a measure are population-specific (Kottner et al., 2011), the moderate to high inter-session reliability scores reported for the PHE, PSLR and MTh tests may not be extrapolated to athletic populations.

Futsal and handball have enjoyed a spectacular increase in popularity in recent years and are played worldwide in both professional and amateur leagues (Barbero-Álvarez, Soto, Barbero-Álvarez, & Granda-Vera, 2008). Thus, this is a population where the application of these flexibility measurement tests may be useful to identify players at a high risk for injury as well as to monitor and verify the effectiveness of training and/or rehabilitation programmes. Therefore, the main purpose of this study was to determine the intra-tester reliability (inter-session) of seven lower limb muscles flexibility measures obtained from the PHE, PHF, PHA, PSLR, MTh, ADF_{KE}, and ADF_{KF} tests in futsal and handball players.

2. Method

2.1. Participants

A sample of 60 (30 male and 30 female) futsal and 30 (male) handball players with more than 8 years of playing practice completed this study. The mean height and body mass for females was 166.2 ± 4.9 cm and 59.4 ± 7.5 kg, respectively, and for males, it was 178.0 ± 6.1 cm and 75.5 ± 9.6 kg, respectively. The mean age was 22.4 ± 5.3 years for females and 22.1 ± 2.8 years for males. Futsal players were recruited from four men's and three women's teams that participated in the Spanish first (women's teams) and second (men's teams) divisions. Handball players were recruited from three teams participating in the Spanish second division. Both futsal and handball players had similar training loads: 3–4 training sessions per week, 1.5 h per session, and one game at the weekend. The study was carried out during the competitive phase of their year.

The exclusion criteria were: (a) history of orthopaedic problems, such as episodes of muscle injury, fractures, surgery or pain in the spine or lower limb muscles over the previous six months; (b) missing one testing session during the data collection phase; and (c) presence of delayed onset muscle soreness at any testing session. Female participants were not in the ovulation phase (days 10–14) of their menstrual cycle during testing, as fluctuating concentrations of oestrogen throughout the menstrual cycle affect musculotendinous stiffness and joint laxity (Bell, Myrick, Blackburn, Shultz, Guskiewicz, & Padua, 2009; Eiling, Bryant, Petersen, Murphy, & Hohmann, 2007). The participants were verbally informed about the study procedures before testing, and they provided written informed consent. This study was approved by the University of Murcia Research Ethics Committee (Spain).

2.2. Procedure

The intra-tester (inter-session) reliability was analysed for each muscle flexibility measure using a repeated measures design. A week before testing sessions commenced, the participants were familiarised with how to undergo the tests to reduce the influence of learning on the measurement.

After the familiarisation protocol was concluded, each participant completed each of the seven tests twice on three different occasions, with a 2-week interval between testing sessions. The rationale for using three testing sessions to determine the reliability in our study (instead of the two testing sessions that have been typically used in previous reliability studies) was based on the simulations run by Hopkins (2000), who stated that to achieve an accurate reliability score, a minimum of three testing sessions and 50 participants are needed. In futsal and handball, during the competitive phase of a year, each team plays one week at home and the following week travel away to another city to play against other team. When a team travels to play against other team, players spend long period of time sitting in buses, planes or trains. Because of this, a 2-week interval period was chosen to avoid the possible influence on the passive muscle-tendon properties of these long periods of time spent sitting, and to keep the same conditions between testing sessions. Thus, all players were assessed 72 h after their last game played at home.

Two physical therapists with greater than 10 years' experience (one conducted the tests and the other ensured proper testing position of the participants throughout the assessment manoeuvre) conducted each of the three testing sessions at the same time of the day under the same environmental conditions. The physical