Title: Reproducibility of a knee and hip proprioception test in healthy older adults

Authors: Mina Arvin MSc a, Marco J.M. Hoozemans PhD a,c,* , Bart J. Burger MD, PhD c, Sabine M. P. Verschueren PhD b, Jaap H. van Dieën PhD a,d, Mirjam Pijnappels PhD a

a MOVE Research Institute Amsterdam, Faculty of Human Movement Sciences, VU University Amsterdam, The Netherlands

b Department of Rehabilitation Sciences, Faculty of Kinesiology and Rehabilitation Sciences, KU Leuven, Belgium

c CORAL – Centre for Orthopaedic Research Alkmaar, Orthopaedic Outpatient Department, Medical Centre Alkmaar, The Netherlands

d King Abdulaziz University, Jeddah, Saudi Arabia

Corresponding author: Marco Hoozemans. Address: VU University Amsterdam , Van der Boechorststraat 9, 1081 BT Amsterdam, the Netherlands. Tel: +31 20 5988561

E-mail address: m.j.m.hoozemans@vu.nl
ABSTRACT

Background: Proprioception can be assessed by measuring joint position sense (JPS). Most studies have focused on JPS of the knee joint while literature for other joints especially for hip JPS is scarce. Although some studies have evaluated proprioception of the knee joint, the reproducibility of methods has rarely been investigated.

Aims: To estimate intra-session reliability and agreement of an active-active JPS test for hip flexion/abduction and knee flexion in healthy older adults.

Methods: Nineteen healthy older adults participated in this study. The proprioception of the hip (flexion and abduction) and knee (flexion) were assessed in both legs using the “active-active” reproduction technique. Intraclass Correlation Coefficient (ICC), standard error of measurement (SEM) and limits of agreement (LOA) were estimated for relative-angular-error (RE), absolute-angular-error (AE) and variable-angular-error (VE).

Results: Reliability of our JPS test was substantial to almost perfect for the RE for both joints and legs (ICC values ranging from 0.75 to 0.93). We also found that the ICC values for AE were substantial for knee flexion and hip abduction of the left and right leg. The ICC results of VE showed poor reliability for hip and knee joints. SEM and LOA values for hip abduction were generally lower than for hip and knee flexion, indicating lower measurement error or more precise scores for the proprioception test of hip abduction.

Conclusion(s): Proprioceptive acuity of the knee and hip joints in healthy older adults can be reliably assessed with an active-active procedure in a standing position with respect to relative and absolute error.

Key words: reliability, agreement, position sense, lower limb, standing, elderly
INTRODUCTION

Proprioception is defined as the perception of limb position, orientation and movement [1], and is essential for muscular control, the precision of motion, balance, and joint stability [2-4]. A large body of evidence suggests that declined proprioceptive function during the aging process impacted on motor coordination and balance [5-9]. Proprioception can be assessed by measuring joint position sense (JPS), and it usually involves a procedure where a target joint position is presented to a participant who is required to replicate that position, either simultaneously with the contralateral joint or after memorization with the same joint. The difference between the presented target and replicated position is used as a measure of accuracy. JPS of lower extremity joints can be assessed by various methods. Some studies have used a non weight-bearing (WB) posture such as side-lying [10] or sitting [11], whereas assessment of proprioception in a standing posture might be more functionally relevant, especially in relation to falling [12,13]. In addition, the JPS can be measured with various protocols such as passive-passive [14], active-active [15] or passive-active [16] testing paradigms, in which the limb is passively (by an examiner) or actively (by the participant) moved to the target angle when presenting the target and when replicating the target angle. Active positioning and repositioning require muscle contraction, which enhances proprioceptive acuity in comparison with passive positioning or repositioning [17], and also it has been shown to be more accurate and repeatable [18,19]. Furthermore, it is likely more representative of the sensory experience during the natural movement patterns of daily life activity. With respect to the lower extremities, most studies have focused on JPS of the knee joint [20-23] while literature for other joints especially for hip JPS is scarce [10]. In addition, although some studies have evaluated proprioception of the knee joint, the reproducibility of methods has rarely been investigated.
Reproducibility concerns the degree to which the same results are obtained on repeated measurements using the same methodology in a study [24]. Reproducibility is used as an umbrella term for the concepts of reliability and agreement. Reliability parameters assess how well subjects in a group can be distinguished from each other, despite measurement errors and are mainly expressed using an intraclass correlation coefficient (ICC). Agreement parameters assess how far apart the repeated measurements of a test-retest design within subjects are by estimating the measurement error [24].

In order to quantify proprioception in the context of balance control during standing and walking, in relation to falling in older adults, we developed an active-active JPS test in a standing position for hip flexion, hip abduction and knee flexion. In this study, we aimed to estimate the intra-session reliability and agreement of this JPS test in healthy older adults. Our findings may indicate whether these JPS tests are suitable for detecting short-term (experimental) effects in future studies with older populations.

**METHODS**

**Participants**

Nineteen healthy older adults (14 women and 5 men; mean age 73.5 (SD 7.8) years; height 167.3 (SD 9.7) cm; weight 70.6 (SD 11.3) kg) voluntarily participated in this study. They were recruited from a database and were excluded if they had a history of cardiovascular problems, joint disorders, neurological deficits, lower extremity injuries within the last 6 months, postural instability, and vestibular or visual problems. The local ethics committee approved the procedure in accordance with the declaration of Helsinki and all participants gave their written, informed consent before participation.
**Procedure of joint position sense (JPS) testing**

Participants wore their shoes and were dressed in shorts and shirts during all tests. They stood blindfolded upright on one leg on a 10 cm high block, with their other leg unsupported but aligned with the supported leg. By standing on the block with one leg, the other leg was allowed to freely flex the knee and flex or abduct the hip. Participants were allowed to touch a horizontal bar located in front of them at hip height for further support with both hands throughout the measurement. This reference position was the starting position for each of the JPS tasks described below.

Proprioception of the hip (flexion and abduction) and knee (flexion) were assessed in both legs using the “active-active” reproduction technique, which means that (1) from the reference starting position, the participants were verbally instructed to slowly, but at a self selected velocity, and actively move the segment (i.e. lower leg for knee flexion and entire straight-leg for hip flexion and abduction) to a target angle, identified by the examiner on the command “STOP”. The participants responded “YES” when they completely stopped the leg and the examiner pressed a switch; (2) the participants had to memorize the target angle while holding their leg at that position for approximately 4 s; (3) the participants were instructed to slowly return their leg to the reference starting position on the command “RETURN”, and to keep it there for 3 s; and finally (4) the participants were asked to actively reproduce the previous target angle using the same limb within 5 seconds. The switch was pressed by the examiner again when the participants completely stopped the leg and answered “YES”.

Four trials were performed for each of the 3 tests on both limbs. The target angles of the four trials were visually determined by the examiner, and randomly varied within the range from 40° to 90° for knee flexion, 10° to 40° for hip abduction, and 10° to 45° for hip flexion (see data
analysis section for the definition of the angles). To minimize the impact of potential sources of error, the protocol was standardized to a high degree, the examiner was trained to standardize subject positioning and instruction, and subjects practiced in two trials before the actual measurements. The same examiner was used for all subjects as well as all sessions.

In total, each participant had to perform 24 trials (4 trials × 3 movements of knee flexion, hip flexion and abduction × 2 legs) during 20 minutes. All participants were tested in two sessions (of 24 trials each) and they had a rest period of about 15 minutes between the first and the second session.

For each trial, movement kinematics of 8 LED markers, bilaterally attached to the apex of the iliac crest, greater trochanter, lateral femur epicondyle and lateral malleolus were captured using the Optotrak system (Northern Digital Inc, Waterloo, Ontario, Canada) at a sample frequency of 100 Hz.

**Data analysis**

Time series of the LED marker positions were used to calculate 1) the knee angle during knee flexion, defined as the relative angle between the lower segment (vector between the knee and ankle marker) and upper segment (vector between the knee and greater trochanter marker), 2) the hip angle during hip flexion and hip abduction, defined as the relative angle between the vector of greater trochanter and knee marker and the vector of greater trochanter and iliac crest marker. Time series of the joint angles were synchronized with the switch data. The switch was pressed for about half a second, thus the average angle of these 40-50 samples was defined as the angle of the knee and hip joints. The JPS was determined from the differences between the target and reproduced angle and expressed using the following parameters: (1) the absolute angular error
(AE), which is defined as the absolute difference between the target and the reproduced angle. The AE represents the overall accuracy in performance without directional bias, since the direction of the error (positive or negative) is ignored in the AE calculation; (2) the relative angular error (RE), which is the signed arithmetic difference between the target and reproduced angle and accounts for accuracy with the direction of the error. The mean value of the four trials was used to calculate the AE and RE, and (3) the variable angular error (VE), which is determined as the standard deviation of the RE. The VE represents the consistency of the RE within the set of 4 trials. Data were analyzed using MATLAB (The MathWorks, Inc., Massachusetts, USA).

**Statistical analysis**

To express reproducibility of the JPS measurements, intra-session reliability and agreement parameters were estimated. The Intraclass Correlation Coefficient (ICC(2,k), absolute agreement) suggested by Shrout and Fleiss [25] was calculated. As a general guideline, Landis and Koch (1998) suggested the ICC value between 0.2 and 0.4 as fair, 0.4 and 0.6 as moderate, 0.6 and 0.8 as substantial and from 0.8 to 1 as almost perfectly reliable [26]. In addition, the standard error of measurement (SEM) and the mean and standard deviation (SD) of the differences between the first and second sessions were used to estimate the limits of agreement (LOA) [27]. LOA indicates that 95% of the differences between the two session scores are within these intervals. SPSS version 20 was used to calculate all relative and absolute reliability values.
RESULTS

Mean and SD of two sessions for relative errors, absolute errors and variable errors are presented in Table 1. Mean reposition errors ranged from 0.36° (SD 1.2) for hip abduction to 4.66° (SD 2.45) for knee flexion and the mean difference between sessions ranged from 0.01° (SD 0.96) for hip flexion to 0.79° (SD 2.01) for knee flexion.

Table 1: Means and standard deviation (SD) of the two sessions of the joint position sense tests.

<table>
<thead>
<tr>
<th>RE variable*</th>
<th>left</th>
<th></th>
<th></th>
<th>right</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean1(SD)</td>
<td>Mean2(SD)</td>
<td>Mean-d(SD-d)</td>
<td>Mean1(SD)</td>
<td>Mean2(SD)</td>
<td>Mean-d(SD-d)</td>
</tr>
<tr>
<td>Hip abduction</td>
<td>0.64 (1.00)</td>
<td>0.84 (1.11)</td>
<td>-0.19 (0.67)</td>
<td>0.36(1.2)</td>
<td>0.63(1.51)</td>
<td>-0.27(1.36)</td>
</tr>
<tr>
<td>Hip flexion</td>
<td>0.50(1.50)</td>
<td>0.63(1.23)</td>
<td>-0.12(1.46)</td>
<td>0.85(0.81)</td>
<td>0.92(1.12)</td>
<td>-0.07(1.01)</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>2.38(2.62)</td>
<td>3.11(3.17)</td>
<td>-0.73(3.1)</td>
<td>2.53(3.69)</td>
<td>2.14(3.47)</td>
<td>0.38(3.82)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AE variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip abduction</td>
</tr>
<tr>
<td>Hip flexion</td>
</tr>
<tr>
<td>Knee flexion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VE variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip abduction</td>
</tr>
<tr>
<td>Hip flexion</td>
</tr>
<tr>
<td>Knee flexion</td>
</tr>
</tbody>
</table>

* RE=relative angular error; AE=absolute angular error; VE=variable angular error; Mean1(SD)=mean and standard deviation (SD) of first session in degrees; Mean2(SD)=mean and SD of second session in degrees; Mean-d (SD-d)= mean and SD differences between two sessions in degrees

Relative error

Table 2 shows the reliability (as expressed by the ICC) and agreement (as expressed by the SEM and LOA) for the RE of JPS of both joints and legs. Reliability was substantial to almost perfect
(0.75 to 0.93) for each of the joints in both legs. The value of the SEM ranged from 0.48° for left hip abduction to 2.71° for right knee flexion. The interval of the LOA was the largest for right knee flexion (-7.12° to 7.89°,) and the smallest for left hip abduction (-1.51° to 1.13°).

**Table 2: intra-session reliability and agreement for relative angular error (RE)**

<table>
<thead>
<tr>
<th>RE variable*</th>
<th>left</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>SEM</td>
</tr>
<tr>
<td>Hip abduction</td>
<td>0.93</td>
<td>0.48°</td>
</tr>
<tr>
<td>Hip flexion</td>
<td>0.77</td>
<td>1.03°</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.75</td>
<td>2.20°</td>
</tr>
</tbody>
</table>

* RE=relative angular error; ICC = intraclass correlation coefficient; SEM = standard error of measurement; LOA = limits of agreement

**Absolute error**

ICC values between 0.68 and 0.81 indicate a substantial reliability for the AE for hip abduction and knee flexion (Table 3). However, low ICC values were found for hip flexion, especially for the left leg (ICC 0.11). The SEM varied between 0.39° for left hip abduction and 1.68° for right knee flexion. The smallest and largest intervals of LOA for AE were found for right hip flexion and knee flexion, respectively.

**Table 3: intra-session reliability and agreement for absolute angular error (AE)**

<table>
<thead>
<tr>
<th>AE variable*</th>
<th>left</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>SEM</td>
</tr>
<tr>
<td>Hip abduction</td>
<td>0.81</td>
<td>0.39°</td>
</tr>
<tr>
<td>Hip flexion</td>
<td>0.11</td>
<td>0.72°</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.68</td>
<td>1.61°</td>
</tr>
</tbody>
</table>

* AE=absolute angular error; ICC = intraclass correlation coefficient; SEM = standard error of measurement; LOA = limits of agreement
Variable error

The reliability and agreement results for VE are presented in Table 4. The reliability of the VE ranged from very low (ICC 0.06) for right hip flexion to substantial (ICC 0.73) for right knee flexion. The SEM was between 0.54° and 2.08° for right hip abduction and left knee flexion, respectively. The largest interval for the LOA was for left knee flexion (-5.79° to 5.73°) and the smallest one was for left hip abduction (-0.89° to 2.09°).

Table 4: intra-session reliability and agreement for variable angular error (VE)

<table>
<thead>
<tr>
<th>VE variable*</th>
<th>left</th>
<th></th>
<th>right</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>SEM</td>
<td>LOA</td>
<td>ICC</td>
</tr>
<tr>
<td>Hip abduction</td>
<td>0.70</td>
<td>0.68°</td>
<td>-2.14°, 1.62°</td>
<td>0.36</td>
</tr>
<tr>
<td>Hip flexion</td>
<td>0.59</td>
<td>0.66°</td>
<td>-1.83°, 1.81°</td>
<td>0.06</td>
</tr>
<tr>
<td>Knee flexion</td>
<td>0.20</td>
<td>2.08°</td>
<td>-5.79°, 5.73°</td>
<td>0.73</td>
</tr>
</tbody>
</table>

* VE=variable angular error; ICC = intraclass correlation coefficient; SEM = standard error of measurement; LOA = limits of agreement

DISCUSSION

The objective of the present study was to estimate the intra-session reliability and agreement of an active-active JPS test for hip and knee joints in a standing position in healthy older adults. As a relative index for reliability, the ICC results indicated that reliability of the JPS test was substantial to almost perfect for the RE for both joints and legs (ICC values ranging from 0.75 to 0.93). The ICC results of VE showed poor reliability for hip and knee joints and the ICC values were different between left and right leg for both joints. With respect to the level of agreement of the JPS test, SEM and LOA values for hip abduction were generally lower than for hip and knee flexion.
Reliability

Direct comparisons of the ICC values observed in the present study with ICC values reported in previous studies cannot be made for knee position sense, since published ICC data were derived from different types of JPS tests [28,29,11,19,30]. Olsson and co-workers (2004) found fair to good reliability of knee position sense, but tested knee flexion in a sitting and a prone position with an active-active technique [11]. Mendelson and co-workers (2004) used an electrogoniometer to measure knee extension and hip flexion proprioception using a passive-active reproduction technique in a sitting and a prone position. Based on their results of intersession reliability, the ICC values varied from 0.51 to 0.69 for the hip and from 0.53 to 0.61 for the knee [31]. To our knowledge, only Benjaminse and co-workers [10] recently studied the intra- and intersession reliability of hip proprioception for hip flexion/extension, abduction/adduction and internal/external rotation, but their study showed poor intra-session reliability (ICCs <0.32) for their JPS measurements. They included Threshold to Detect Passive Motion (TTDPM), force sense (FS) and active JPS methods. Perhaps the large number of measurements caused fatigue or reduced participants’ concentration, with negative effects on reliability, although they allowed 10 minutes rest-periods between proprioception tests. The authors further suggested that the poor results might be due to the difficulty to maintain balance in a single-legged standing position with the eyes blindfolded. In the present study this is not considered as a factor that may have biased the results as our blindfolded participants indicated themselves that they were able to easily keep their balance since they were allowed to touch a horizontal bar in front of them with both hands.

Despite the substantial reliability of AE for knee flexion and hip abduction, it did not show good reliability for hip flexion. Moreover, the AE for hip flexion appeared not to be
comparable between the left and right hip. The difference observed in reliability of the AE between left and right hip flexion might be explained by exploring the limits of agreement (Table 3) and mean differences (Table 1), which illustrate the differences of systematic and random errors of AE respectively. The systematic error (mean differences) of AE for the left hip appeared less than for the right hip, but the random error for left hip flexion was higher than for right hip flexion as the interval of LOA for left hip flexion was wider in comparison with right hip flexion. Apparently, especially for the left hip, the participants may have responded too similarly in both sessions in relation to a relatively large between session variation, resulting in a low ICC for AE because of lower variance between participants rather than higher measurement error. The comparison of RE and AE variables for left hip flexion confirms that the low ICC value for left hip AE is not related to measurement error, since the systematic and random errors of left hip RE is comparable to left hip AE, while tables 2 and 3 show almost perfect reliability for RE but poor reliability for AE for left hip flexion.

The results of VE showed that neither the ICC values of the hip joint, nor the ICC values of the knee joint were comparable for the left and right leg. For instance, the reliability of right knee flexion was considerable (ICC 0.73), whereas the reliability of left knee flexion was poor (ICC 0.2). Thus, the VE is not a suitable parameter to characterize the proprioception with any of our knee and hip joint tests.

**Agreement**

The LOA is useful to determine when a change in an individual’s performance as a result of an intervention is “real”. If the change between two measurements is outside the LOA, this change is likely to be the cause of the intervention and the result indicates a true change in the subject’s
performance, beyond measurement error [32]. Compared to previous studies on knee and hip joint position sense [33-35,20], the interval of the LOA in our study was rather large, especially for knee position sense. This indicates that our test cannot detect small changes of knee proprioception over time in individuals. Rankin and Stocks [32] suggested that studies on LOAs need a sample set of at least 50, otherwise the LOA will be overestimated; so the LOA reported here might therefore be negatively affected by our sample size of 19 individuals.

In contrast with the ICC, which relates within subject variance to the total variance, agreement parameters (e.g. SEM and LOA) are absolute indices of repeatability of participants’ scores [36,37,29]. According to our results, the SEM value for the hip joint was lower than for the knee joint, especially for hip abduction, indicating a lower measurement error or more precise scores for the proprioception test for the hip joint. Our findings are comparable with results of previous study that also found smaller absolute matching errors in hip joint position sense compared to knee joint position sense [33].

**Limitations**

Standing position, movement velocity of the leg, environmental circumstances, subjects’ attention, and learning effects may have negatively affected the reproducibility of our test. However, our results showed that these potential sources of error did not have a major impact on the reproducibility of our test. This is probably because we standardized the protocol and trained the examiner to standardize subject positioning and instruction and because subjects were able to practice before the actual measurements. In addition, we measured JPS over a relatively large range of angles during four trials which were not equal over subjects and the target angles of first session were not the same as in the second session, which may have affected our results,
especially for the VE. As stated above, in this study JPS was assessed in a healthy older population, which limits the external validity of this study. The JPS tests may be used in groups in which JPS is likely to be affected, such as in the present sample of older adults. However, the reproducibility of the JPS tests may be different in other groups with affected JPS such as people with knee OA, peripheral neuropathy and Parkinson’s disease, which are patients groups that also suffer from balance loss and falls [38-40]. This needs to be further examined.

Participants were instructed to hold the target and reposition angles constantly for approximately four seconds. Only 40 to 50 ms were used for the analyses. Despite the ability to maintain a rather constant angle, there was some small variation in angle during holding the target or reproduction angle and it can be questioned what part of the four seconds for the target angle is remembered and whether this is reflected in the 40 to 50 ms of the reproduction angle. As we expect this part of the protocol to increase the random measurement error, the actual reproducibility of the presented JPS tests might be underestimated.

We tested reproducibility for two sessions that shortly followed each other at one day. This means that between day reproducibility still needs to be explored in future if the effects of interventions on JPS are evaluating over a period of time.

Finally, a limitation of the present study with respect to feasibility may be the use of the Optotrak system to assess segment movement. In clinical settings this advanced system is often not available. Although the protocol can be performed using inclinometers or electrogoniometers [41], the reproducibility is likely to become worse because of a larger random measurement error and should be further investigated.
CONCLUSION

The presented active-active procedure for the assessment of knee and hip proprioception in a standing position in a group of healthy older adults showed substantial intra-session reliability and acceptable agreement for RE and the majority of AE, but not for VE. Even though it can be used as a reliable method for future studies to investigate the effects of short-term intervention, we advocate further research aiming at between day reliability with larger sample sizes and patient populations to optimize the presented repositioning tests as tools for clinical evaluation.

ACKNOWLEDGEMENTS

This research was funded by the European Commission through MOVE-AGE, an Erasmus Mundus Joint Doctorate programme (2011-0015). Mirjam Pijnappels was financially supported by a TOP-NIG grant (#91209021) from the Dutch Organization for Scientific Research (NWO). The authors would like to thank Arjan de Zwart for his valuable assistance on data collection.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

REFERENCES