Concurrent validity and reliability of a novel wireless inertial measurement system to assess trunk movement

C.M. Bauer a,b,⇑, F.M. Rast a,1, M.J. Ernst a,1, J. Kooi a,1, S. Oetiker a,1, S.M. Rissanen c,3, J.H. Suni d,4, M. Kankaanpää b,e,1,2

a Zurich University of Applied Sciences, Department of Health, Institute of Physiotherapy, Technikumstrasse 71, 8400 Winterthur, Switzerland
b University of Tampere, School of Medicine, Kalevantie 4, FI-33014 University of Tampere, Finland
c University of Eastern Finland, Department of Applied Physics, University of Eastern Finland, P.O. Box 1627, 70211 Kuopio, Finland
d UKK Institute for Health Promotion Research, Raapinpuistonkatu 1, 33500 Tampere, Finland
e Tampere University Hospital, Department of Physical and Rehabilitation Medicine, Tampere University Hospital, P.O. Box 2000, 33521 Tampere, Finland

A R T I C L E   I N F O

Article history:
Received 18 November 2014
Received in revised form 3 June 2015
Accepted 4 June 2015

Keywords:
Generalizability-theory
Movement disorders
Reproducibility of results
Biomechanical phenomena

A B S T R A C T

Introduction: Assessment of movement dysfunctions commonly comprises trunk range of motion (ROM), movement or control impairment (MCI), repetitive movements (RM), and reposition error (RE). Inertial measurement unit (IMU)-systems could be used to quantify these movement dysfunctions in clinical settings. The aim of this study was to evaluate a novel IMU-system when assessing movement dysfunctions in terms of concurrent validity and reliability. Methods: The concurrent validity of the IMU-system was tested against an optoelectronic system with 22 participants. The reliability of 14 movement dysfunction tests were analysed using generalizability theory and coefficient of variation, measuring 24 participants in seven trials on two days. Results: The IMU-system provided valid estimates of trunk movement in the primary movement direction when compared to the optoelectronic system. Reliability varied across tests and variables. On average, ROM and RM were more reliable, compared to MCI and RE tests. Discussion: When compared to the optoelectronic system, the IMU-system is valid for estimates of trunk movement in the primary movement direction. Four ROM, two MCI, one RM, and one RE test were identified as reliable and should be studied further for inter-subject comparisons and monitoring changes after an intervention.

1. Introduction

Movement dysfunctions in patients suffering from diseases such as low back pain (LBP), stroke and Parkinson’s disease can be clinically assessed by measuring their trunk range of motion (ROM) and their reaction to specific movement control tasks (Laird et al., 2014; Verheyden et al., 2007; Cole et al., 2010). Specifically, these assessments are comprised of (1) ROM (Laird et al., 2014), (2) movement control impairment (MCI) (Sahrmann, 2002; Luomajoki et al., 2007), (3) repetitive movement (RM) tests (Dideriksen et al., 2014), and (4) tests for proprioception deficits such as reposition error tests (RE) (Rausch Osthoff et al., 2015).

Optoelectronic measurement systems are accepted as gold-standards for non-invasive analysis of trunk movement within research settings (Cuesta-Vargas et al., 2010, McGinley et al., 2009). However they are not applicable in daily clinical practice due to their high cost, required installation space, specific marker placement and subsequent data capture, analysis and processing. These factors limit the analysis to some standard procedures, which cannot be extended to clinics (Wong and Wong, 2009). Alternative objective, valid, and reliable measurement systems are needed to allow clinicians to assess and monitor individual patient changes and compare between different population groups.

To overcome these limitations, new wireless movement analysis systems using body-worn sensors have recently been developed (e.g. Valedo from Hocoma AG, ViMove from dorsaVi, or Reabio from Corehab). These clinical systems comprise of multiple small
light weight inertial measurement units (IMU) which measure the angular tilt and velocity of body segments with respect to magnetic fields and gravity (Roetenberg et al., 2007). By combining the output of multiple IMUs and post processing algorithms into an IMU-system it is possible to estimate joint angles in a non-invasive way.

Using concurrent validation, the output of an IMU system can be correlated to the gold-standard, whilst simultaneously measuring with both systems (Steiner and Norman, 2008). Recent research examined concurrent validity of a wired IMU system and found a high correlation to the gold-standard (Wong and Wong, 2009; Wong et al., 2007). However correlation studies between two systems should provide both a measure of random error, or precision, as well as accuracy of the devices in their units of measurement (e.g. degrees). (de Vet et al., 2006). In a systematic review of the literature, Cuesta-Vargas and colleagues found that IMU systems can be concurrent to optoelectronic analysis of trunk measurements, but the degree of concurrent-validity is specific to the IMU system and anatomical site (Cuesta-Vargas et al., 2010).

Reliable measures of trunk movement and control are needed to monitor individual changes over time and to compare between different individuals. Reliability concerns the degree to which repeated measures provide similar results (de Vet et al., 2006). Reliability is affected by interrater, intrasession, and intersession variability (Corriveau et al., 2006). Interrater variability is unlikely to be a concern for measurements with an IMU system, except for sensor placement. Variability of sensor placement can be minimised by using a standardised protocol (Ernst et al., 2013). Intra- and intersession variability depend on biological variability, hence they are test specific. Reliable tests can be identified by estimating the magnitude of intra- and intersession variability. Furthermore, recommendations can be made for the number of trials needed to be averaged from one or more sessions in order to improve reliability (Santos et al., 2008).

This study assesses concurrent validity of a novel wireless IMU system, by using an optoelectronic system as a gold standard. Second, it investigates the reliability of commonly used trunk movement and control tests, when measured with a wireless IMU system.

2. Methods

This study was divided into two sub-studies: A concurrent validity study (study V) and a reliability study (study R).

2.1. Participants

Twenty-two and twenty-four asymptomatic participants volunteered for studies V and R respectively. The participant’s characteristics are presented in Table 1. Detailed exclusion criteria for both studies are described elsewhere (Schelldorfer et al., 2015). For study R, the sample size was calculated according to Walter et al. (1998). Twenty participants and five trials allow reliability estimations of 0.95 with a type I error of 0.05 and a type II error of 0.20. The studies were approved by the local ethics commission and participants provided their informed consent.

2.2. Marker and sensor placement

Four IMUs were placed on the right thigh (THI), over the sacrum (S2), and at the level of L1 (L1), and T1 (T1), as described elsewhere (Ernst et al., 2013; Schelldorfer et al., 2015). The IMUs were mounted on a plastic frame and attached to the skin with hydrogel tape (KCI Medical GmbH 8153 Rümlang, CH). Reflective markers were placed above and below every IMU with a third marker attached to the stiletto on the plastic frame. Thus it was possible to build virtual segments corresponding to the IMU plane, and to compare the two systems (Fig. 1). The IMU and optoelectronic systems were synchronised using digital signals generated from a Labjack U3® data acquisition device (Labjack Corporation, USA).

2.3. Measurement systems and data processing

Trunk movements were measured by the IMU system in both studies and additionally with an optoelectronic motion capture system (VICON, Oxford UK) in study V. In study V, a fourth-order zero-phase low-pass Butterworth filter (6 Hz cut-off frequency) was used to filter the raw data of both systems. In study R, an eighth-order zero-phase low-pass Butterworth filter (6 Hz cut-off frequency) was used since we analysed acceleration and jerk, which are noisy measures and require smoothing to obtain interpretable estimates.

2.3.1. Optoelectronic system

The optoelectronic system consisted of twelve infrared cameras. Data was sampled at 200 Hz and processed using VICON Nexus® software. The coordinate system of each segment, defined by three reflective markers, was aligned to the coordinate system of the IMU. The difference signal between two segments was calculated and transformed into tilt/twist angles according to Crawford and colleagues (Crawford et al., 1999). We adopted the following sign convention: flexion, lateral flexion towards the right, and axial rotation towards the left were assigned positive values; movements in the opposite directions were assigned negative values. We termed the angle between the L1 and T1 segment “Thoracic Spine”, the angle between S2 and L1 “Lumbar Spine,” and the angle between thigh and S2 “Hip angle”.

2.3.2. Inertial measurement units

The Valedo® system (Hocoma AG) is a professional medical system used for low back pain therapy. The Valedo IMU’s contain a tri-axillar gyroscope, magnetometer, and accelerometer, as well as wireless antenna and signal processing unit. The specifications of the IMU’s indicate they are able to record ±0.1° over a range of 360° around all axes (Valedo® User Manual, Hocoma AG). IMU sensor data was transmitted to a recording computer with a 200 Hz sampling frequency. Custom data acquisition and synchronisation software (Valedo® Research) was provided by Hocoma AG. The raw IMU sensor data was transformed into quaternions according to Madgwick and colleagues (Madgwick et al., 2010). The angular difference between two IMU’s placed above the body segments was calculated and transformed into tilt/twist angles. A complete description of the data processing from raw data to tilt/twist angles is documented in Supplementary File 1.

2.4. Data analysis

Intra- and intersession variability depend on biological variability, hence they are test specific. Reliable tests can be identified by estimating the magnitude of intra- and intersession variability. Furthermore, recommendations can be made for the number of trials needed to be averaged from one or more sessions in order to improve reliability (Santos et al., 2008).

This study assesses concurrent validity of a novel wireless IMU system, by using an optoelectronic system as a gold standard. Second, it investigates the reliability of commonly used trunk movement and control tests, when measured with a wireless IMU system.

Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>All (n = 22)</th>
<th>Women (n = 11)</th>
<th>Men (n = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>41.18 ± 11.14</td>
<td>38.27 ± 10.44</td>
<td>44.09 ± 11.53</td>
</tr>
<tr>
<td>Body mass index</td>
<td>22.99 ± 2.89</td>
<td>22.67 ± 3.02</td>
<td>23.32 ± 2.85</td>
</tr>
</tbody>
</table>

n: Number of participants.
2.4. Procedures

2.4.1. Study V

Participants attended one measurement session and performed four ROM tests in randomized order, as described in Table 2. They were tutored by a video showing the correct movement. Additionally, they were instructed to move as far as possible at their preferred speed. Each test was performed three times.

2.4.2. Study R

Participants attended two identical measurement sessions, separated by a 1 week period. Both measurement sessions took place at the same time of day. All participants performed 14 tests, which were grouped into four categories according to their purposes: (1) ROM, (2) MCI, (3) RM and (4) RE. Test (1) measures the flexibility of the participant’s spine within their comfort zone. Test (2) evaluates the participant’s ability to control and differentiate movement between two body segments and to stabilize their spine. The former parameter was analysed by calculating the ratio of the ROM of the respective body segments, while the latter was investigated using the ROM of the respective segment. Furthermore, the root mean squared jerk (RMSJ), as described by Slaboda et al. (2005), was calculated as indication of movement control. Test (3) measured the variability of angular displacement and acceleration during repeated movements. Variability was examined by calculating percentage of recurrence (%REC) and determinism (%DET) using recurrence quantification analysis (RQA) (Webber and Zbilut, 1994). Test (4) evaluates the participant’s proprioceptive deficits within the spine, analysed using constant error (CE) (Rausch Osthoff et al., 2015).

Participants performed four ROM, six MCI, two RM, and two RE tests as described in Table 2. Each test was performed seven times, except for those in four point kneeling (4pk) which was reduced to 5 repetitions to minimise loading through their wrists. The order of the tests was randomized between participants but not between days.

2.5. Statistical analysis

2.5.1. Study V

The coefficient of determination ($r^2$), a measure of precision, and root mean squared error (RMSE), a measure of accuracy, were used to test the concurrent validity of the IMU system:

$$r^2 = 1 - \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

$$\text{RMSE} = \sqrt{\frac{\sum(y_i - x_i)^2}{n}}$$

where $x$ and $y$ are the two time based movement signals, and $\bar{y}_i$ being the predicted value obtained by linear regression. The values of $r^2$ ranged from 0 to 1. A high value of $r^2$ implies that angles measured by IMUs and the optoelectronic system have the same characteristic. RMSE is the measure of the average difference between the two signals. Systematic differences between the systems were analysed using the Wilcoxon rank sum-test with $p$ set at <0.05.

2.5.2. Study R

The generalizability theory (Brennan, 2001) with the design $p \times t \times d$ (participants $\times$ trials $\times$ days) was used as a framework to estimate reliability of trunk movement measures, based on the linear model.

$$X_{ptd} = \mu + \nu_p + \nu_t + \nu_d + \nu_{pt} + \nu_{td} + \nu_{ptd}$$

with $\mu$ representing the global mean and $\nu$ any one of the seven components.

The index of dependability $\Phi$ was calculated as:

$$\Phi = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_n^2 + \sigma_n^2 + \sigma_n^2 + \sigma_n^2 + \sigma_n^2 + \sigma_n^2}$$

with $\sigma$ being the variance, and $n$ the number of the corresponding component (with $n_p$, $n_q$, and $n_r$ being the number of trials, participants, and days, respectively). $\Phi$ was interpreted as: <0.25 very low, 0.26–0.49 – low, 0.50–0.69 – moderate, 0.70–0.89 – high, and >0.90 – very high reliability (Carter et al., 2005). $\Phi \geq 0.70$ was interpreted as sufficient to compare between different individuals. Subsequently, $\Phi$ coefficients were calculated for alternative measurement strategies, where $n_p$ was varied up to ten trials, and $n_d$ varied across two days, which represent acceptable measurement strategies. Thereby, the number of required trials per day to achieve high reliability was evaluated.

The coefficient of variation (CV) (Hopkins, 2000) was calculated as

$$CV = \frac{\sigma_{\text{std}}}{\sqrt{n_p \times \bar{x}}} \times 100$$

with $\bar{x}$ being the grand mean and $\sigma_{\text{std}}$ being the standard deviation of the differences between days and calculated from the mean of seven trials per day. The CV values were rated as follows: $>$10% not reliable, 6–10% adequately reliable and 5% highly reliable. CV’s $\leq 10\%$ were construed as sufficient to monitor changes over time (Suni et al., 2014).

The diagnostic value of a variable was assessed by $\Phi$ whereas the ability to detect changes over time was evaluated by the CV.

3. Results

3.1. Study V

In general, trunk movements in the sagittal plane were overestimated by the IMU system compared to the optoelectronic system (angular values between 1.3° and 6.5°). In contrast, frontal plane movements of the trunk were underestimated (angular values between 0.7° and 3.1°). Movements of the hip were measured...
<table>
<thead>
<tr>
<th>Test</th>
<th>Starting position</th>
<th>Movement</th>
<th>BS</th>
<th>Variable (unit)</th>
<th>Description of variable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROM tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROM flexion</td>
<td>Standing upright</td>
<td>Maximal flexion of the LS</td>
<td>LS</td>
<td>ROM_FLEX (°)</td>
<td>ROM LS</td>
</tr>
<tr>
<td>ROM extension</td>
<td>Standing upright</td>
<td>Maximal extension of the LS</td>
<td>LS</td>
<td>ROM_EXT (°)</td>
<td>ROM LS</td>
</tr>
<tr>
<td>ROM lateral flexion right</td>
<td>Standing upright</td>
<td>Maximal lateral flexion of the LS</td>
<td>LS</td>
<td>ROM_RIGHT (°)</td>
<td>ROM LS</td>
</tr>
<tr>
<td>ROM lateral flexion left</td>
<td>Standing upright</td>
<td>Maximal lateral flexion of the LS</td>
<td>LS</td>
<td>ROM_LEFT (°)</td>
<td>ROM LS</td>
</tr>
<tr>
<td><strong>MCI tests (Luomajoki et al., 2007)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pelvic tilt</td>
<td>Standing upright</td>
<td>Anterior pelvic tilt without moving</td>
<td>LS</td>
<td>RATIO_PT</td>
<td>ROM LS/ROM TS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the trunk or knees</td>
<td>TS</td>
<td>RMSJ_PT_LS (°/s²)</td>
<td>Smoothness of movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSJ_PT_TS (°/s²)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waiters bow</td>
<td>Standing upright</td>
<td>Hip flexion without moving the LS</td>
<td>LS</td>
<td>RATIO_WB</td>
<td>ROM LS/ ROM Hip</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hip</td>
<td>RMSJ_WB_LS (°/s²)</td>
<td>Smoothness of movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSJ_WB_Hip (°/s²)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitting knee extension</td>
<td>Sitting upright</td>
<td>Knee extension without moving the LS</td>
<td>LS</td>
<td>RATIO_FE</td>
<td>ROM LS/ROM Hip</td>
</tr>
<tr>
<td></td>
<td>Hips at 90°</td>
<td></td>
<td></td>
<td>RMSJ_SKE_LS (°/s²)</td>
<td>Smoothness of movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSJ_SKE_Hip (°/s²)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rocking backwards</td>
<td>4pk</td>
<td>Hip flexion and shoulder extension</td>
<td>LS</td>
<td>RATIO_RF</td>
<td>ROM LS/ROM Hip</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without moving the LS</td>
<td>Hip</td>
<td>RMSJ_RF_LS (°/s²)</td>
<td>Smoothness of movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSJ_RF_Hip (°/s²)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rocking forwards</td>
<td>4pk</td>
<td>Hip extension and shoulder flexion</td>
<td>LS</td>
<td>RATIO_PKB</td>
<td>ROM LS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without moving the LS</td>
<td></td>
<td>RMSJ_PKB_LS (°/s²)</td>
<td>Smoothness of movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prone knee bend</td>
<td>Lying prone</td>
<td>Knee flexion without moving the LS</td>
<td>LS</td>
<td>RATIO_PKB</td>
<td>ROM LS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RMSJ_PKB_LS (°/s²)</td>
<td>Smoothness of movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RM tests (Dideriksen et al., 2014)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picking up a box</td>
<td>Standing upright</td>
<td>Lifting a box (5% body weight) four</td>
<td>LS</td>
<td>%REC_PU_AD, %DET_</td>
<td>Percentage of recurrence points within a recurrences plot (%REC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>times in a row at 60bpm</td>
<td></td>
<td>PU_AD (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>%REC_PU_AA, %DET_</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FE_AD (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>%REC_FE_AA, %DET</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FE_AA (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flexion and extension</td>
<td>Sitting upright</td>
<td>Repeated flexion and extension of the</td>
<td>LS</td>
<td>%REC_FE_AD, %DET_</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hips at 60°</td>
<td>trunk, five times in a row at 80bpm</td>
<td></td>
<td>FE_AD (%)</td>
<td>And percentage of recurrences points forming diagonal line structures in this plot (%DET)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>%REC_FE_AA, %DET</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FE_AA (%)</td>
<td>(Webber and Zbilut, 1994; Marwan et al., 2002; Rissanen et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RE Tests (Rausch Osthoff et al., 2015)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reposition Error Sitting</td>
<td>Sitting upright</td>
<td>Flexion of the trunk and reproducing</td>
<td>LS</td>
<td>CE_SIT (°)</td>
<td>Angular difference between starting and final position</td>
</tr>
<tr>
<td></td>
<td>Hips at 60°</td>
<td>the starting position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4pk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reposition Error 4pk</td>
<td>Extension of the</td>
<td>Extension of the LS and reproducing</td>
<td>LS</td>
<td>CE_4PK (°)</td>
<td>Angular difference between starting and final position</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>the starting position</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4pk: four point kneeling; %DET: percentage of determinism; %REC: percentage of recurrence; AA = angular acceleration; AD: angular displacement; bpm: beats per minute; BS: Body segment; CE: constant error; EXT: Extension; FE: Flexion and Extension; FLEX: Flexion; LS: lumbar spine; MCI: Movement control impairment; PKB: prone knee bend; PT: Pelvic Tilt; PU: Picking Up a Box; RB: rocking backwards; RE: Reposition Error; RF: rocking forwards; RM: repetitive movement; RMSJ: root mean squared jerk; ROM: range of motion; SKE sitting knee extension; SIT: sitting; TS: Thoracic Spine; WB: waiting bow.
almost equally with both systems. A summary of the results is presented in Table 3.

No significant systematic differences were found in the primary movement direction, except for sagittal and frontal plane movement of the thoracic spine (flexion and lateral flexion to the right).

The measurement systems showed acceptable agreement and small measurement errors in the primary movement direction. The \( r^2 \) coefficients ranged between 0.94 and 0.99, except for hip movement during the lateral flexion tests (0.85–0.87) and the RMSE ranged between 1.1° and 6.8°. Flexion of the lumbar spine and the hip, as well as lateral flexion of the thoracic and lumbar spine, revealed very high agreement with an \( r^2 \) coefficient of 0.99 and RMSE ranging between 1.8° and 6.1°. In the non-primary movement directions, \( r^2 \) coefficients were lower (0.36–0.87) while RMSE were similar (1.2°–6.8°) compared to the primary movement direction (Supplementary File 2).

3.2. Study R

Table 4 summarises the grand mean, \( \Phi \)-coefficients, and the number of trials averaged from one or two measurement days which are needed to gain \( \Phi \geq 0.70 \), and the CV for each variable. On average, ROM and RM tests needed a smaller number of trials to reach high reliability and had smaller CVs compared to MCI and RE tests.

Measured values from single trial tests of trunk ROM revealed high to very high reliability except for extension of the lumbar spine. All CVs were smaller than 10%. The MCI tests differed in their reliability with \( \Phi \)-coefficients of a single measurement ranging from low to high, and CVs from 8 to 22%. The RM tests showed CVs smaller than 10%, with the “Picking up a Box” test being more reliable than the “Flexion and Extension” test. The RE tests showed a relatively low reliability for a single measurement with CVs greater than 10%.

4. Discussion

The main findings of the present study were that the use of a wireless IMU system is a valid alternative to measure trunk movement in the primary movement direction when compared to the golden standard (i.e. an optoelectronic system). Secondly, on average, the ROM and RM tests needed a smaller number of trials to reach high reliability and had smaller CVs when compared to MCI and RE tests.

4.1. Study V

The measured ROM falls well within the range of previously published results, although comparability is hampered by a large variety of measurement approaches, including measurement systems and participants selection (Laird et al., 2014). Both our optoelectronic and IMU systems measured similar ROM, whilst sagittal plane movement was slightly overestimated, and frontal plane movement underestimated, by the IMU systems.

This study showed that trunk ROM in the primary movement direction can be accurately measured by using a wireless IMU system; however, the system appears less valid for assessing movements in non-primary directions. Although RMSE were similar in magnitude compared to the primary movement direction, they were higher relative to the total ROM. The agreement could be affected by the noise, and limited resolution of the IMU system, a nonlinear correlation between both systems, and constraints on mathematical calculations.

The present study improves upon previous work (Ha et al., 2013; Wong and Wong, 2009) with a more detailed analysis of ROM measures which includes thoracic spine and hip ROM. Furthermore, the concurrent validity of the novel wireless IMU system compares well to other studies validating different IMU systems against a gold-standard (Dunne et al., 2008; Ha et al., 2013; Wong and Wong, 2008, 2009).

4.2. Study R

The index of dependability \( \Phi \) of a single trial varied across different tests and variables, ranging from 0.19 to 0.90. The CV varied considerably as well, ranging from <1% to 37%. Reliability can be improved by increasing the number of trials/days and using the mean value. While, for some variables, averaging over days affected reliability more than averaging over trials on one day, this is not necessarily a practical solution, especially in clinical settings. If one attempts to increase the number of trials, care should be taken that a learning-effect or fatigue does not influence the participants’ performance (Santos et al., 2008).

4.2.1. Range of motion

Three out of the four lumbar ROM variables reached high reliability with a single trial on one day, whereas the extension ROM only had moderate reliability. Averaging two single trials over two days increased the reliability of ROM extension more than averaging several trials on one day, indicating that it is affected more by sources of variance between days rather than within one day. The decreased reliability of ROM extension could be explained by biological variability between days, the test-setup, or the slightly lower concurrent validity of the IMU system (Table 3).

The low CVs (3–9%) indicate high reliability for measuring changes in ROM over time. These results are in accordance with other studies reporting high reliability of ROM measures (Al Zoubi and Preuss, 2013). The measured ROM is almost identical to study V and within the range of previously published results (Laird et al., 2014).

4.2.2. Movement control impairment

The MCI tests differed in their reliability. “Waiters Bow” and “Sitting Knee Extension” reached high reliability when averaging a maximum of six trials on one day, or two trials on two days. The magnitude of the between-day variance is also shown by the CV, ranging between 8% and 22%. Nonetheless, the mean ROM in “Sitting Knee Extension” was approaching zero, with about 25% of participants moving into extension, hampering the interpretation of the CV (22%) for this variable. “Pelvic Tilt”, “Rocking Forwards”, “Rocking Backwards,” and “Prone Knee Bend” showed little to moderate reliability. The reliability might be affected by the complexity or the standardisation of the MCI tests or because segment movement ranges, duration, and speed were not controlled. Standardising the MCI tests for one of these factors might decrease within-day and between-days variance.

Our results are somewhat contradictory in regard to previous research, where the reliability of MCI tests was reported as substantial based on a dichotomous variable (positive or negative indication) (Luomajoki et al., 2007). Although a growing body of research investigates MCI of the trunk and hip (Luomajoki et al., 2007; Saner et al., 2015), no normative values have been published aside from this study. Additionally, the different approaches to quantify MCI tests make it difficult to compare our results.

4.2.3. Repeated movement tests

The “Picking Up a Box” test had high reliability by averaging a maximum of four trials on one measurement day, with low CVs (<3%). Our descriptive results for %DET of angular displacement
Table 3

<table>
<thead>
<tr>
<th>Study</th>
<th>Results for trunk range of motion measures, primary movement direction.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROM thoracic spine (°)</td>
</tr>
<tr>
<td>flexion left</td>
<td>32.6 ± 9.2</td>
</tr>
<tr>
<td>flexion right</td>
<td>32.6 ± 9.2</td>
</tr>
<tr>
<td>extension left</td>
<td>36.0 ± 16.1</td>
</tr>
<tr>
<td>extension right</td>
<td>36.0 ± 16.1</td>
</tr>
</tbody>
</table>

- 2 = Significance of their measures (Rausch Osthoff et al., 2015).
- Significantly performed by flexing the spine and hips, while the second test is based on flexion and extension. This study measured extension was less reliable and had lower concurrent validity, which might explain the lower \( r \) values. Both tests were highly standardised, possibly explaining the small standard deviations of these variables.

4.2.4. Reposition error

Reposition error, CE (Rausch Osthoff et al., 2015), reached high reliability after averaging six trials on one day (4pk) or eight trials across two days (sitting). The CE can have positive and negative values and a score of zero implies a good performance. These characteristics result in an expected grand mean around zero and, therefore, huge CVs. Consequently, the CV should not be interpreted for these two variables. In such situations \( r \) gives a better indication of reliability. The magnitude of the measured RE is well within the range of previously published data on pain-free participants (Rausch Osthoff et al., 2015). Data on reliability of RE measures is discouraging. Several studies report poor reliability of RE tests, use an inadequate numbers of trials, or do not report reliability of their measures (Rausch Osthoff et al., 2015).

4.3. Limitations of this study

The IMU system is a valid tool when measuring flexion of the lumbar spine and hip, as well as lateral flexion of the thoracic and lumbar spine. On the other hand, measurements of thoracic spine flexion and hip lateral flexion should be viewed with caution. Some of the differences between the two systems can be characterised as errors in the optoelectronic system. These errors could be triggered by camera noise, limited sight of markers, or vibrations of the marker frame (Ehara et al., 1997). Additionally both systems are affected by skin surface artefacts caused by contraction of the muscles or prominent spinal processes (Yang et al., 2008).

The sample size was calculated for an Intraclass-Correlation-C coefficient model (Walter et al., 1998). We assume this to be appropriate as both models share similarities while generalizability theory is regarded as an expansion of classical reliability theory (Brennan, 2001). RMSJ was calculated as a measure of movement control that has been shown to be reliable and discriminative between populations (Slaboda et al., 2005). However, RMSJ is sensitive to movement duration, amplitude, and arrest (Hogan and Sternad, 2009). Other indices of movement control could be investigated in future studies. This study has focused on pain-free participants. Although reliability is affected by the heterogeneity of study populations (Lariviere et al., 2013), the inclusion of pain-free participants was reasonable to evaluate the usability of an IMU system to measure trunk kinematics.

4.4. Suggestions for future research

The evaluated wireless IMU system is appropriate as a more affordable alternative to an optoelectronic system within the demonstrated boundaries regarding secondary movement directions. The IMU system’s concurrent validity might be enhanced by investigating the technical validity of the IMU components and subsequently improving these components.

Future studies should address reliability on different populations and assess diagnostic value and the ability to detect changes of the presented measures over time in more detail.
Differences between populations and treatment effects of interventions aiming at improving movement control have to be investigated. Measures of RQA in repeated movement tests are highly dependent on the input parameters (Rissanen et al., 2008; Webber and Zbilut, 1994). Other choices for input parameters, apart from the ones used in our study (Table 5), are possible, and optimal input parameters have to be investigated in future studies.

4.5. Clinical implications and recommendations

Clinicians commonly use range of motion and movement control tests of the trunk and hip to assist in identifying patterns of dysfunction and to monitor change (Laird et al., 2014). This paper presents a measurement tool which enables the clinicians to do this objectively. To identify dysfunctions and changes in performance, high reliability is important. Based on our results, we recommend the use of four ROM tests, selected MCI tests (“Waiters Bow” and “Sitting Knee Extension”), RE in 4pk, and “Picking up a Box” for RM, using an adequate number of trials for each test (Table 4).

5. Conclusion

The usage of a wireless IMU system led to valid estimates of trunk movement in the primary movement directions. A number of tests to assess movement dysfunctions and their corresponding variables were identified as reliable and should be studied further for intersubject comparisons and monitoring changes after an intervention.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

The Swiss Commission of Technical Innovation funded this study. Saara Rissanen was supported by the Academy of Finland.
We thank Dr. Melissa Lawrence for language editing. We thank Mrs Rausch-Osthoff, Hausherr, Aegterer and Mr Heimgartner for assisting with data-acquisition.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jelekin.2015.06.001.

References


Sahrmann S. Diagnosis and treatment of movement impairment syndromes. St. Louis, Missouri 63146 USA: Mosby; 2002.


Christoph Bauer received his BSc in physiotherapy from the Hogeschool of Amsterdam, the Netherlands, and his MSc in physiotherapy from Philips-University Marburg, Germany. He is currently a PhD student at Tampere University, School of Medicine, Finland and vice head of research at the institute of Physiotherapy, Zurich University of Applied Sciences.

Fabian Rast received his MSc in Movement Science from the Federal Technical University Switzerland, Zurich, Switzerland. He is currently researcher at the movement laboratory of the Institute of Physiotherapy, Zurich University of Applied Sciences.

Markus Ernst received his MSc in Physiotherapy from Maastricht University, the Netherlands. He is currently researcher at the Institute of Physiotherapy, Zurich University of Applied Sciences.
Jan Kool received his MSc and PhD in Physiotherapy from Maastricht University, the Netherlands. He was head of research and development at the institute of physiotherapy, Zurich University of Applied Sciences from 2006 until 2013. He is currently researcher and physiotherapist at the Clinic Valens, Switzerland.

Sarah Oetiker received her MSc in Movement Science from the Federal Technical University Switzerland, Zurich, Switzerland. She is currently researcher at the movement laboratory of the Institute of Physiotherapy, Zurich University of Applied Sciences.

Saara Rissanen received her MSc and PhD in Physics from University of Eastern Finland, Finland. She is currently researcher at the Department of Applied Physics, University of Eastern Finland, Finland.

Jaana Suni received her MSc and PhD in Exercise Therapy from University of Jyväskylä, Finland. She is a senior researcher at the UKK Institute, Tampere Finland.

Markku Kankaanpää received his MD and PhD in Medicine from University of Eastern Finland, Finland. He is currently associated professor of rehabilitation medicine at University of Tampere, Finland.