Biomechanical measures of short-term maximal cycling on an ergometer: a test-retest study

Louise Burnie[a,b,c]; Paul Barratt[d]; Keith Davids[b]; Paul Worsfold[c,e]; Jon Wheat[f*].

[a] *Applied Sports, Technology, Exercise and Medicine Research Centre, Swansea University, Swansea, UK*

[b] *Academy of Sport and Physical Activity, Sheffield Hallam University, Sheffield, UK*

[c] *Biomechanics, English Institute of Sport, Manchester, UK*

[d] *Team INEOS, Manchester, UK*

[e] *Sport and Exercise Sciences, University of Chester, Chester, UK*

[f] *Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, UK*

*Corresponding author – Jon Wheat. E-mail: j.wheat@shu.ac.uk T: 0114 225 4330

Biomechanical measures of short-term maximal cycling on an ergometer: a test-retest study

An understanding of test-retest reliability is important for biomechanists, such as when assessing the longitudinal effect of training or equipment interventions. Our aim was to quantify the test-retest reliability of biomechanical variables measured during short-term maximal cycling. Fourteen track sprint cyclists performed 3 x 4 s seated sprints at 135 rpm on an isokinetic ergometer, repeating the session 7.6 ± 2.5 days later. Joint moments were calculated via inverse dynamics, using pedal forces and limb kinematics. EMG activity was measured for 9 lower limb muscles. Reliability was explored by quantifying systematic and random differences within- and between-session. Within-session reliability was better than between-sessions reliability. The test-retest reliability level was typically moderate to excellent for the biomechanical variables that describe maximal cycling. However, some variables, such as peak knee flexion moment and maximum hip joint power, demonstrated lower reliability, indicating that care needs to be taken when using these variables to evaluate biomechanical changes. Although measurement error (instrumentation error, anatomical marker misplacement, soft tissue artefacts) can explain some of our reliability observations, we speculate that biological variability may also be a contributor to the lower repeatability observed in several variables including ineffective crank force, ankle kinematics and hamstring muscles’ activation patterns.

Keywords: sprint cycling, kinematics, kinetics, emg, maximal power.

Introduction

The reliability of a clinical or sports science test is defined as the consistency or reproducibility of a performance when a test is performed repeatedly (Hopkins, Schabort, & Hawley, 2001). This is an important consideration for researchers, clinicians and applied sports scientists as the better the reliability of the measurement the easier it is to detect a real change in outcome (Hopkins, 2000). If the reliability of a test is low, then the outcome of a test may conceal the true effect of an intervention. Conversely, if the reliability of a test is not known then small random deviations may be misinterpreted as a meaningful change in performance (Yavuzer, Öken, Elhan, & Stam, 2008).

Applied biomechanics researchers are often interested in assessing the short- or long-term effects of interventions that aim to improve clinical or sports performance outcomes. In clinical gait analysis, for example, the results of biomechanical assessments are used to inform clinical decision making, by evaluating the effectiveness of interventions such as surgery, physical therapy, medication or orthotics on gait biomechanics (Kadaba et al., 1989; McGinley, Baker, Wolfe, & Morris, 2009; Yavuzer et al., 2008). Test-retest reliability studies of clinical gait have found that the sagittal plane kinematics and kinetics have high values of reliability in comparison to the data collected in the transverse and coronal planes (McGinley et al., 2009). Furthermore, knee abduction/adduction and hip, knee and foot rotation joint angles demonstrate the lowest reliability (McGinley et al., 2009), with the size of the measurement error the same order of magnitude as the real joint motion in these planes. In the context of clinical gait therefore, reliability studies have proved valuable by identifying those variables that need to be interpreted with particular caution in order to effectively inform clinical decision making (McGinley et al., 2009).

An understanding of test-retest reliability has similar relevance when assessing sporting movements, as biomechanical measures are often used to evaluate the effectiveness of longitudinal interventions such as changes to training programmes or equipment modification (Costa, Bragada, Marinho, Silva, & Barbosa, 2012; Milner, Westlake, & Tate, 2011). Cycling is a commonly used sporting movement for this purpose, as it is a relatively constrained movement that can be accurately manipulated (Neptune, Kautz, & Hull, 1997; Neptune & Kautz, 2001). Whilst the reliability of submaximal or "endurance" cycling is well reported (Bini & Hume, 2013; Hopkins et al., 2001; Jobson, Hopker, Arkesteijn, & Passfield, 2013; Laplaud, Hug, & Grélot, 2006), only a small amount by comparison is known about the reliability of short-term maximal cycling. This comparative deficit exists despite maximal cycling being an important paradigm for studying physiological capacity (Coso & Mora-Rodríguez, 2006), muscle coordination and motor control strategies, as well as having direct relevance to a range of competitive cycling performance environments (Martin, Davidson, & Pardyjak, 2007). Therefore, quantifying test-retest reliability in maximal cycling biomechanics is important. Test-retest reliability has been quantified for overall net crank power output on an inertial load cycling ergometer within- and between-session (Coso & Mora-Rodríguez, 2006; Hopkins et al., 2001; Mendez-Villanueva, Bishop, & Hamer, 2007), with trained cyclists producing reliable power within the first testing session (Martin, Diedrich, & Coyle, 2000). These studies demonstrated within-session reliability was better than between-sessions reliability for overall net crank power output (Coso & Mora-Rodríguez, 2006; Martin et al., 2000). There have been no studies quantifying the within- and between-session reliability of biomechanical variables (crank power and forces, joint angles, angular velocities, moments and powers and EMG activity) for short-term maximal cycling despite these measures being important descriptors of the outcome, technique and intermuscular coordination of a movement (Brochner Nielsen et al., 2018; Jacobs & van Ingen Schenau, 1992; Wakeling, Blake, & Chan, 2010). EMG activity can be used to determine muscle activation onset and offset times and level of activation (Dorel, Guilhem, Couturier, & Hug, 2012; Hug & Dorel, 2009). This is important when investigating intermuscular coordination in cycling as the timing and magnitude of muscle activation has to be coordinated appropriately to allow an efficient energy transfer from the muscles though the body segments to the pedal (Neptune & Kautz, 2001; Raasch, Zajac, Ma, & Levine, 1997). Joint kinetic measures (moments and powers) at the hip, knee and ankle throughout the pedal revolution describe the action and contribution of the joints to pedal power and can be used to identify different coordination strategies between cyclists (Elmer, Barratt, Korff, & Martin, 2011; Martin & Brown, 2009; McDaniel, Behjani, Brown, & Martin, 2014). Combining information on muscle activation from EMG and joint kinetics from inverse dynamics analysis provides a deeper understanding of the joint and muscle actions that produce the movement, and hence both are required to describe intermuscular coordination in maximal cycling and were chosen for measurement and analysis during maximal cycling (Brochner Nielsen et al., 2018; Dorel, 2018).

The aim of this study was to quantify the test-retest reliability of kinematic, kinetic, and muscle activation variables during maximal sprint cycling. We hypothesise that within-session reliability would be better than between-sessions reliability.

Methods

Participants

Fourteen track sprint cyclists participated in the study. Participants regularly competed at track cycling competitions at either Master’s international and national level (10), or Junior national level (4). Although the participants were varied in their anthropometrics (7 males and 7 females, age: 40.5 ± 17.7 yr, body mass: 72.5 ± 8.5 kg, height: 1.71 ± 0.06 m,), they were similar with respect to cycling performance level (flying 200 m personal best: 11.98 ± 0.90 s). Participants were provided with study details and gave written informed consent. The study was approved by the Sheffield Hallam University Faculty of Health and Wellbeing Research Ethics Sub-Committee.

Experimental protocol

An isokinetic ergometer was set up to replicate each participant's track bicycle position. All participants' crank lengths were set to 165 mm, which was what they rode on their track bicycles. Riders undertook their typical warm-up on the ergometer at self-selected pedalling rate and resistance for at least 10 minutes, followed by one familiarisation sprint (4 s at 135 rpm). Martin and colleagues demonstrated that trained cyclists can produce valid and reliable results for maximal cycling power from the first testing session (Martin et al., 2000), therefore one familiarisation sprint was deemed appropriate. Riders then conducted 3 x 4 s seated sprints at a pedalling rate of 135 rpm on the isokinetic ergometer with 4 minutes recovery between efforts. Participants undertook an identical session 7.6 ± 2.5 days apart, at approximately the same time of day (0.11 ± 2.18 h). A pedalling rate of 135 rpm was chosen as this is a typical pedalling rate during the flying 200 m event in track cycling and within the optimal pedalling rate range for track sprint cyclists (Dorel et al., 2005). The competitive level and typical training volume of our participants meant that it was not feasible to ask them to stop exercising 24 hours prior to the testing sessions, so instead they were instructed to undertake the same training in the preceding 24 hours before both sessions.

Isokinetic ergometer

A SRM Ergometer (Julich, Germany) cycle ergometer frame and flywheel were used to construct an isokinetic ergometer. The modified ergometer flywheel was driven by a 2.2-kW AC induction motor (ABB Ltd, Warrington, UK). The motor was controlled by a frequency inverter equipped with a braking resistor (Model: Altivar ATV312 HU22, Schneider Electric Ltd, London, UK). This set-up enabled the participants to start their bouts at the target pedalling rate, rather than expending energy in accelerating the flywheel. The ergometer was fitted with Sensix force pedals (Model ICS4, Sensix, Poitiers, France) and a crank encoder

(Model LM13, RLS, Komenda, Slovenia), sampling data at 200 Hz. Normal and tangential pedal forces were resolved using the crank and pedal angles into the effective (propulsive) and ineffective (applied along the crank) crank forces (Figure 1).

Kinematic and Kinetic Data Acquisition

Two-dimensional kinematic data of the participants' left side were recorded at 100 Hz using one high speed camera with infra-red ring lights (Model: UI-522xRE-M, IDS, Obersulm, Germany). The camera was perpendicular to the participant, centred on the ergometer and set about 3 m from the ergometer. The camera was in a very similar position for both sessions. Reflective markers were placed on the pedal spindle, lateral malleolus, lateral femoral condyle, greater trochanter and iliac crest. The same researcher attached the markers for all sessions. Kinematics and kinetics on the ergometer were recorded by CrankCam software (Centre for Sports Engineering Research, SHU, Sheffield, UK), which synchronised the camera and pedal force data (down sampled to 100 Hz to match the camera data) and was used for data processing, including auto-tracking of the marker positions.

EMG Data Acquisition

EMG signals were recorded continuously from nine muscles of the left leg: vastus lateralis (VL), rectus femoris (RF), vastus medialis (VM), tibialis anterior (TA), long head of biceps femoris (BF), semitendinosus (ST), lateralis gastrocnemius (GL), soleus (SO), and gluteus maximus (GMAX) with Delsys Trigno wireless surface EMG sensors (Delsys Inc, Boston, MA). The skin at electrode placement sites was prepared by shaving the area then cleaning it with an alcohol wipe. The EMG sensors were then placed in the centre of the muscle belly - with the bar electrodes perpendicular to the muscle fibre orientation, using the guidelines in (Konrad, 2005) and secured using wraps to reduce motion artefacts during pedalling. The same researcher attached the EMG sensors for all sessions. A Delsys wireless sensor containing an accelerometer (148 Hz sampling rate) was attached to the left crank arm to obtain a measure of crank angle synchronised with the EMG signals. The EMG system was operated and recorded in EMGworks Acquisition software (Delsys Inc, Boston, MA), sampling data at 1926 Hz. The Delsys trigno EMG system automatically applied a bandwidth filter of 20 ± 5 Hz to 450 ± 50 Hz (>80 dB/dec) to the raw signals.

Data Processing

All kinetic and kinematic data were filtered using a Butterworth fourth order (zero-lag) low pass filter with a cut off frequency of 14 Hz selected using residual analysis (Winter, 2009). The same cut off frequency was chosen for the kinematic and kinetic data as recommended by Bezodis and colleagues to avoid data processing artefacts in the calculated joint moments (Bezodis, Salo, & Trewartha, 2013). Instantaneous crank power was calculated from the product of the left crank torque and the crank angular velocity. The average left side crank power was calculated by averaging the instantaneous crank power over a complete pedal revolution. Owing to a technical fault with the force measurement in the right pedal, it was not possible to calculate total average crank power per revolution (sum of left and right crank powers). Joint angles were calculated using the convention shown in Figure 1. Joint moments were calculated via inverse dynamics (Elftman, 1939), using pedal forces, limb kinematics, and body segment parameters (de Leva, 1996). Joint extension moments were defined as positive and joint flexion moments as negative. The joint moments are presented from the internal perspective (Derrick et al., 2020). Joint powers at the ankle, knee and hip were determined by taking the product of the net joint moment and joint angular velocity.

Insert Figure 1

Data were analysed using a custom Matlab (R2017a, MathWorks, Cambridge, UK) script. Each sprint lasted for 4 s providing six complete crank revolutions which were resampled to 100 data points around the crank cycle. Crank forces and powers, joint angles, angular velocities, moments and powers were averaged over these revolutions to obtain a single ensemble-averaged time series for each trial.

The accelerometer data for the crank arm was filtered using a Butterworth fourth order low pass filter with a cut off frequency of 10 Hz. The minimum value of the acceleration of the sensor in the direction of the crank arm corresponded to top dead centre (TDC) crank position. To synchronise the EMG data with the kinematic and kinetic data, the TDC locations from the accelerometer on the crank arm were matched to the corresponding TDC measured by the crank encoder.

The raw EMG signals for the sprint efforts were high pass filtered (Butterworth second order, cut off frequency 30 Hz) to diminish motion artefacts (De Luca, Gilmore, Kuznetsov, & Roy, 2010), root mean squared (RMS, 25 ms window) and then low pass filtered (Butterworth second order, cut off frequency 24 Hz) (Brochner Nielsen et al., 2018). The data were then interpolated to 100 data points around the crank cycle and then averaged over 6 crank revolutions to create a linear envelope for each muscle. The EMG signals were normalised to the mean value in the linear envelope across the crank cycle for each muscle.

Statistical Analysis

In order to test for any systematic change in performance between-sessions (for example, due to learning or fatigue effects) paired *t*-tests were used to compare differences between discrete values. Paired *t*-tests only test if there is a statistically significant bias between-sessions (systematic change) but provide no indication of the random error due to biological or mechanical variation between-sessions (Atkinson & Nevill, 1998). Similarly, differences in time series data (instantaneous crank powers, crank forces, joint angles, angular velocities, moments, powers and normalised EMG linear envelopes) between-sessions were assessed using Statistical Parametric Mapping (SPM); paired *t*-tests were used for all variables except crank forces where Hotelling's paired T^2 test was used (Pataky, 2010). Crank force consists of two vector components (effective and ineffective crank force), therefore a multivariate statistical test was required (Pataky, 2010). The level of statistical significance was set to $p < 0.05$ for all tests.

The reliability of the discrete variables between sessions was assessed using intra-class correlation coefficient (ICC) tests. ICC's were calculated using IBM SPSS Statistics Version 24 (IBM UK Ltd, Portsmouth, UK), based on average measures, absolute agreement, two-way mixed effects model (ICC (3,*k*) - where *k* is equal to the number of trials in a session which in this study is three). The ICCs were interpreted using Koo and Li's guidelines: values less than 0.50 are indicative of poor reliability, between 0.50 and 0.75 indicates moderate reliability, 0.75 to 0.90 indicates good reliability and > 0.90 indicates excellent reliability (Koo & Li, 2016). For a variable to be considered as having excellent reliability, both upper and lower bounds of the 95% confidence intervals must fall within the excellent range (i.e. > 0.9) (Koo & Li, 2016).

Standard error of measurement (SEM) for between sessions was calculated using the formula (Weir, 2005), where SD is standard deviation of the mean difference:

$$SEM = SD\sqrt{1 - ICC}$$

Minimal detectable difference (MDD) was calculated for between sessions using the formula (Weir, 2005):

$$MDD = SEM \times 1.96 \times \sqrt{2}$$

The coefficient of variation (CV) was calculated for the average crank power over a complete revolution (Hopkins, 2000).

The standard error of measurement (SEM) was calculated for the kinematic and kinetic time series data to evaluate the reliability of these waveforms within- and between-session using the methods described in Pini, Markström, & Schelin, 2019. The mean and SD SEM for a complete revolution was calculated for each variable. The EMG data were visually inspected for signal quality and the frequency spectrum of the raw and filtered EMG signal calculated. EMG signals with a high frequency content below 20 Hz, indicates low frequency noise due to movement artefact (De Luca et al., 2010) and therefore, these trials were discarded. The SEM for within- and between-session for the EMG linear envelopes of the VL, VM, ST, and GMAX muscles were calculated using 13 participants. At least 2 trials for each muscle per session per participant were required to calculate SEM. The calculated reliability of the EMG data is therefore the upper bound, as very noisy trials were discarded.

The cross-correlation coefficient (R) was calculated to compare the temporal effects of within- and between-session EMG linear envelopes (Wren, Do, Rethlefsen, & Healy, 2006). The between-sessions cross-correlation coefficient was calculated comparing the session mean EMG linear envelope, and within-session the cross-correlation coefficient was calculated comparing the EMG linear envelope for two trials.

Results

Discrete variables

Discrete crank level variables demonstrated good to excellent between-sessions reliability ICC(3,*k*) > 0.756 (Table 1). Average crank power for a complete revolution for the left side only was 445.3 ± 95.7 and 438.8 ± 111.5 W for session 1 and 2 respectively (Table 1), which gives an indicative total power for a complete revolution, for both cranks, of 891 and 878 W. MDD between-sessions for peak crank power and forces was 21 W and between 9 to 72 N respectively (Table 1). Peak joint angle values typically demonstrated moderate to excellent reliability, with MDD between-sessions from 1.1 to 4.4° (Table 1). Peak joint angular velocity between-sessions reliability was typically moderate to excellent, except for peak knee flexion and hip extension angular velocity which had poor to good reliability (Table 1). MDD between-sessions for peak joint angular velocities ranged from 14 to 59°/s (Table 1). Peak joint moments demonstrated moderate to excellent between-sessions reliability, except for peak knee flexion moment which demonstrated poor to moderate reliability (Table 1). Maximum ankle and knee joint powers demonstrated good to excellent reliability between-sessions whereas, maximum hip power showed poor to good reliability (Table 1). MDD between-sessions for peak joint moments ranged from 2 to 26 N.m and for maximum joint powers 30 to 144 W.

Insert Table 1

CV for average crank power over a revolution was 3.0 ± 1.5% and 4.6 ± 1.9% for within- and between-session respectively.

Time Series Variables

Crank power demonstrated excellent within- and between-session reliability, with a mean SEM between-sessions over a complete revolution of 46.6 ± 9.4 W (Figure 2, Figure 3). Crank power was significantly different ($p < 0.05$) between sessions one and two, between crank angles 340 to 6° (7.2% of crank cycle) (Figure 2). The ineffective crank force was less repeatable (mean SEM = 31.6 ± 18.2 N) than effective crank force (mean SEM = 19.8 ± 4.0 N) within- and between-session, which was associated with a large SEM for ineffective crank force between crank cycles of 140° and 210° (Figure 4, Figure 5). The crank forces were significantly different ($p < 0.05$) between sessions one and two, between crank angles 191 to 199° (2.2% of crank cycle), and 347 and 1° (3.9% of crank cycle) (Figure 4).

Joint angles and angular velocities demonstrated excellent within- and between-session reliability (mean SEM ≥ 2.4° and 34.1°/s) (Figure 6). Ankle joint angles and angular velocities were less repeatable than those at the knee and hip joints. Ankle joint angular velocity was significantly different ($p < 0.05$) between sessions one and two, between crank angles 152 to 170° (5.0% of crank cycle) (Figure 6).

Joint moments and powers demonstrated reasonable within- and between-session reliability (mean SEM ≥ 15.5 N.m and 62.6 W) (Figure 6, Figure 7). Hip joint moments and powers were less repeatable than those at the knee and ankle joints, particularly around the location of maximum hip extension moment and power (Figure 7). Ankle joint moment was significantly different ($p < 0.05$) between sessions one and two, between crank angles 340 to 6° (7.2% of crank cycle) (Figure 6). Hip joint power was significantly different ($p < 0.05$) between session one and two between crank angles 340 to 2° (6.1% of crank cycle) (Figure 6).

EMG linear envelope normalised to the mean value in the signal demonstrated high within- and between-session reliability (Figure 8). Mean SEM values for EMG linear envelopes ranged between 0.14 to 0.16, and 0.16 to 0.20 proportion of the mean EMG signal, for within- and between-session respectively. The GMAX, TA, and BF muscles demonstrated the lowest reliability for EMG activity, and the VL and VM muscles the highest reliability (Figure 8). The cross-correlation coefficient (*R*) which compares timing of EMG linear envelopes between-sessions ranged from 0.976 to 0.990 (Figure 8).

Insert Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, Figure 7, Figure 8

Discussion and implications

The purpose of this study was to quantify the test-retest reliability of kinematic, kinetic, and EMG muscle activation variables measured during short-term maximal sprint cycling. Our main findings were that between-sessions test-retest reliability level was typically moderate to excellent for the biomechanical variables that describe maximal cycling, and furthermore that within-session reliability was better than between-sessions reliability. However, some variables, such as peak knee flexion moment and maximum hip joint power demonstrated lower reliability, indicating that care needs to be taken when using these variables to evaluate changes in maximal cycling biomechanics.

Within- and between-session values of SEM for joint angles and angular velocities demonstrated high reliability (Figure 6). We found that ankle joint kinematics (angle and angular velocity) were less repeatable than knee and hip joint kinematics, evidenced by the larger mean SEM values for the ankle joint kinematics. The source of the lower reliability in our ankle joint kinematics data is not clear, although it seems unlikely to be a measurement error, given that anatomical landmark marker placement errors for the lower limb are greatest at the hip, rather than the ankle joint (intra-examiner precision for the greater trochanter marker is 12.2 mm along the long axis of the femur, and 11.1 mm in the anterior-posterior direction, compared to lateral malleolus - 2.6 mm along the long axis fibula, 2.4 mm anterior-posterior direction) (Della Croce, Cappozzo, & Kerrigan, 1999; Della Croce, Leardini, Chiari, & Cappozzo, 2005). Furthermore, the soft tissue artefact (STA) of the lower limb markers in cycling is also largest for the hip rather than the ankle joint (greater trochanter marker displacement at 30 rpm submaximal cycling, 37.3 mm anterior-posterior and 10.3 mm proximal-distal, compared to the lateral malleolus 15.8 mm anterior-posterior and 8.6 mm proximal-distal) (Li et al., 2017). By comparison there are potential biological explanations for the lower reliability of the ankle joint kinematics. Martin and Nichols, for example, demonstrated that the ankle has a different role to the knee and hip joints in maximal cycling and acts to transfer - instead of maximise power (Martin & Nichols, 2018). More specifically, the ankle works in synergy with the hip joint to transfer power produced by the muscles surrounding the hip joint to the crank (Fregly & Zajac, 1996). Our results support this notion by suggesting that cyclists may regulate their ankle angle as part of this hip-ankle synergy, in order to maintain a stable effective crank force. A specially designed experiment would be required to test this hypothesis.

In terms of joint kinetics, joint moments and powers demonstrated lower reliability at more proximal compared to distal joints – with the largest values of SEM for the hip joint moment (Figure 6, Figure 7). This observation may be due to the STA and skin marker misplacement errors being largest at the hip joint, as discussed above (Della Croce et al., 1999; Li et al., 2017). It may also be due to the fact that measurement errors in general (STA, marker misplacement, force pedal measurement precision) will propagate through the inverse dynamics calculations (Myers, Laz, Shelburne, & Davidson, 2015). In either scenario, this indicates that the observed differences in proximal to distal joint reliability are likely to be due to measurement error, rather than biological variability.

The peak knee flexion moment showed poor to moderate between-sessions reliability, with the largest MDD of all joint moments (26 N.m). Error due to knee marker misplacement is dependent on knee flexion angle, with previous studies demonstrating that the greater the knee flexion, the larger error in the joint angle (Della Croce et al., 1999). Marker displacement could therefore explain the poor reliability of our peak knee flexion angular velocity and moment data. Further work is required, using more detailed marker sets and models of STA, to reduce the influence of STA and skin marker misplacement on the calculated kinematics and kinetic variables, which may improve the reliability of the calculated knee flexion and hip joint variables.

Average crank power output over a complete revolution was highly reliable both within- and between-session, supporting the findings of Martin and colleagues that trained cyclists are able to reproduce reliable maximal crank power within one testing session (Martin et al., 2000). Effective crank force exhibited similar reliability to crank power, whereas ineffective crank force demonstrated lower within- and between-session reliability which was associated with the large intra-participant variability and SEM in ineffective crank force between crank angles of 140° and 210° (Figure 4, Figure 5). It is unlikely that force pedals' measurement precision would provide an explanation for these observed differences in reliability between the effective and ineffective crank forces, given that the measurement precision values are the same for all components of force for the instrumented pedals we used (combined error - linearity and hysteresis 1% measuring range (MR) and crosstalk between the components (<1.5% MR) (Sensix, Poitiers, France)). Therefore, it seems probable that the reliability difference between effective and ineffective force may have a biological basis, a notion which can be expanded upon using our EMG results.

EMG linear envelopes generally demonstrated excellent reliability (Figure 8). However, the GMAX, BF and the TA muscles demonstrated the lowest reliability for EMG activity. Lower reliability of the EMG activity for the GMAX and TA muscles have been demonstrated in submaximal cycling (Jobson et al., 2013). The between-sessions reliability of the EMG activity of the GMAX muscle has been shown to decrease with increasing workload (between-sessions CV = 43.1% at 265 W compared to CV = 23.0 at 135 W) (Jobson et al., 2013) which might suggest greater biological variation in the GMAX muscle activity with increased workload, potentially explaining the lower reliability of the GMAX EMG activity. Jobson and colleagues suggested the lower reliability of the EMG activity for the TA muscle might be owing to the fact some cyclists have two bursts of muscle activity per crank revolution which may introduce more between crank revolution variability (Jobson et al., 2013). Measurement error could also be a potential source of the lower reliability of the EMG activity for the TA, as the location of the EMG sensor can strongly influence the pattern of EMG activity recorded owing to crosstalk from the peroneus longus muscle during dynamic movements (Campanini et al., 2007; Hug, 2011). Therefore, small changes in positioning of the EMG sensor between sessions could influence the EMG activity measured. Wren and colleagues suggested the lower reliability of the hamstrings may be due to measurement error reflecting the increased sensitivity of these muscles to electrode placement owing to muscle length and overlying fat mass (Wren et al., 2006). The lower reliability of EMG activity in the BF hamstring muscle may also have a biological basis however, given that our findings are consistent with other studies who suggest that this is related to their bi-articular function (Ryan & Gregor, 1992). Van Ingen Schenau and colleagues for example demonstrated that the bi-articular muscles are important for controlling the direction of the external force on the pedal (van Ingen Schenau, Boots, De Groot, Snackers, & Van Woensel, 1992). They identified that the paradoxical coactivation of the mono-articular agonists (vastii) with bi-articular antagonists (hamstrings) emerges so the bi-articular muscles can help control the desired direction of the force applied to the pedal by adjusting the relative distribution of net moments over the joints (van Ingen Schenau et al., 1992).

On a mechanical basis, the goal of maximal cycling is to maximise the effective crank force as this maximises the propulsive power and thus the speed of the bicycle. Taking our crank force and EMG data together therefore, our results allow us to speculate that cyclists may regulate bi-articular muscles activation to control the direction of the pedal force, with the aim of maximising effective crank force and maintaining a stable outcome at the expense of the ineffective force which does not directly affect the task outcome. The bi-articular muscles (BF, ST and GL) are active in the region of the crank cycle where the ineffective crank is more variable which could explain the biological mechanism underlying this finding. This principle has been observed in walking (Kadaba et al., 1989; Giakas & Baltzopoulos, 1997) and running (Kinoshita, Bates, & DeVita, 1985), where the propulsion and braking ground reaction forces (anterior-posterior and vertical direction) have been shown to have lower between-stride variability than the medio-lateral force. However, further, purposefully designed experiments are required to confirm or refute these speculations.

SPM indicated a significant between-session difference for small regions of the crank cycle, for crank power, crank forces, ankle angular velocity and moment, and hip power. These differences are unlikely to be meaningful changes as these are less than 7.2% of the crank cycle, and typically occur in regions of low magnitude in these variables.

The experimental protocol could have introduced some variability to the kinematics, as although the participants were instructed to remain seated during the sprints on the ergometer, they tended to hover slightly over the saddle (potentially with the aim to increase crank power), which increases pelvis movement. Also, the ergometer was set-up to match each participant's track bike. Therefore, saddle height was not standardised to percentage of inside leg length, which is often recommended (de Vey Mestdagh, 1998). Some of the participants had a relatively low saddle height compared to their leg length, which resulted in relatively large pelvis obliquity (rocking) and transverse rotation when they sprinted. This strategy may have introduced more within- and between-trial variability, particularly at the hip joint. We acknowledge that we measured 2D kinematics using a high-speed video camera, which is not considered the 'gold standard' for measuring kinematics which is 3D motion capture systems (Fonda, Sarabon, & Li, 2014). However, these methods were utilised because during cycling the movement is predominantly in the sagittal plane (Umberger & Martin, 2001; van Ingen Schenau, Van Woensel, Boots, Snackers, & De Groot, 1990) and therefore previous studies that have investigated maximal cycling have just considered the sagittal plane actions, as this is the plane where muscles produce power to generate effective crank force (Barratt, Korff, Elmer, & Martin, 2011; Elmer et al., 2011; Martin & Brown, 2009; McDaniel et al., 2014). Therefore, we measured 2D kinematics in the sagittal plane using a simple marker set which has the added benefit of reducing time required for data collection sessions which is an important ethical consideration when working with elite athletes.

Conclusion

Typically, the biomechanical variables that describe maximal cycling are reliable. However, some variables have lower reliability indicating that care needs to be taken when using these variables to evaluate changes in maximal cycling biomechanics. Our results allow us to speculate that biological variability is the source of the lower reliability of the ineffective crank force, ankle kinematics and hamstring muscles activation while measurement error is the source of the lower reliability in hip and knee joint kinetics. Further research using purposefully designed experiments is required to confirm or refute these speculations. We recognise that there were some data collection problems (noisy EMG data and no right force pedal data) which might indicate potentially lower reliability of our data collection method. These reliability data can be used to help understand the practical relevance of a longitudinal intervention on athletes' maximal cycling performance.

Acknowledgements

The authors are grateful to Todd Pataky for all his help and advice on using and implementing statistical parametric mapping.

References

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*, 217-238.

Barratt, P., Korff, T., Elmer, S. J., & Martin, J. C. (2011). Effect of crank length on joint-specific power during maximal cycling. *Medicine and Science in Sports and Exercise, 43,* 1689-1697.

Bezodis, N. E., Salo, A. I. T., & Trewartha, G. (2013). Excessive fluctuations in knee joint moments during early stance in sprinting are caused by digital filtering procedures. *Gait & Posture, 38*, 653-657.

Bini, R. R., & Hume, P. A. (2013). Between-day reliability of pedal forces for cyclists during an incremental cycling test to exhaustion. *Isokinetics and Exercise Science*, *21*, 203-209.

Brochner Nielsen, N. P., Hug, F., Guevel, A., Colloud, F., Lardy, J., & Dorel, S. (2018). Changes in motor coordination induced by local fatigue during a sprint cycling task. *Medicine and Science in Sports and Exercise, 50*, 1394-1404.

Campanini, I., Merlo, A., Degola, P., Merletti, R., Vezzosi, G., & Farina, D. (2007). Effect of electrode location on EMG signal envelope in leg muscles during gait. *Journal of Electromyography and Kinesiology, 17*, 515-526.

Coso, J. D., & Mora-Rodríguez, R. (2006). Validity of cycling peak power as measured by a short-sprint test versus the wingate anaerobic test. *Applied Physiology, Nutrition, and Metabolism, 31*, 186-189.

Costa, M. J., Bragada, J. A., Marinho, D. A., Silva, A. J., & Barbosa, T. M. (2012). Longitudinal interventions in elite swimming: A systematic review based on energetics, biomechanics, and performance. *The Journal of Strength & Conditioning Research, 26*, 2006-2016.

de Leva, P. (1996). Adjustments to zatsiorsky-seluyanov's segment inertia parameters. *Journal of Biomechanics, 29*, 1223-1230.

De Luca, C. J., Gilmore, L. D., Kuznetsov, M., & Roy, S. H. (2010). Filtering the surface EMG signal: Movement artifact and baseline noise contamination. *Journal of Biomechanics, 43*, 1573-1579.

de Vey Mestdagh, K. (1998). Personal perspective: In search of an optimum cycling posture. *Applied Ergonomics*, *29*, 325-334.

Della Croce, U., Cappozzo, A., & Kerrigan, D. C. (1999). Pelvis and lower limb anatomical landmark calibration precision and its propagation to bone geometry and joint angles. *Medical & Biological Engineering & Computing*, *37*, 155-161.

Della Croce, U., Leardini, A., Chiari, L., & Cappozzo, A. (2005). Human movement analysis using stereophotogrammetry: Part 4: Assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait & Posture, 21*, 226-237.

Derrick, T. R., van den Bogert, Antonie J, Cereatti, A., Dumas, R., Fantozzi, S., & Leardini, A. (2020). ISB recommendations on the reporting of intersegmental forces and moments during human motion analysis. *Journal of Biomechanics, 99*, doi: https://doi.org/10.1016/j.jbiomech.2019.109533

Dorel, S. (2018). Mechanical effectiveness and coordination: New insights into sprint cycling performance. In J. Morin, & P. Samozino (Eds.), *Biomechanics of training and testing* (pp. 33-62). Switzerland: Springer.

Dorel, S., Guilhem, G., Couturier, A., & Hug, F. (2012). Adjustment of muscle coordination during an all-out sprint cycling task. *Medicine and Science in Sports and Exercise*, *44*, 2154-2164.

Dorel, S., Hautier, C. A., Rambaud, O., Rouffet, D., Praagh, E. V., Lacour, J. R., & Bourdin, M. (2005). Torque and power-velocity relationships in cycling: Relevance to track sprint performance in world-class cyclists. *International Journal of Sports Medicine, 26*, 739-746.

Elftman, H. (1939). Forces and energy changes in the leg during walking. *American Journal of Physiology, 125*, 339-356.

Elmer, S. J., Barratt, P., Korff, T., & Martin, J. C. (2011). Joint-specific power production during submaximal and maximal cycling. *Medicine and Science in Sports and Exercise, 43,* 1940-1947.

Fonda, B., Sarabon, N., & Li, F. (2014). Validity and reliability of different kinematics methods used for bike fitting. *Journal of Sports Sciences*, *32*, 940-946.

Fregly, B. J., & Zajac, F. E. (1996). A state-space analysis of mechanical energy generation, absorption, and transfer during pedaling. *Journal of Biomechanics, 29*, 81-90.

Giakas, G., & Baltzopoulos, V. (1997). Time and frequency domain analysis of ground reaction forces during walking: An investigation of variability and symmetry. *Gait & Posture, 5*, 189-197.

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

Hopkins, W. G., Schabort, E. J., & Hawley, J. A. (2001). Reliability of power in physical performance tests. *Sports Medicine, 31*, 211-234.

Hug, F. (2011). Can muscle coordination be precisely studied by surface electromyography? *Journal of Electromyography and Kinesiology, 21*, 1-12.

Hug, F., & Dorel, S. (2009). Electromyographic analysis of pedaling: A review. *Journal of Electromyography and Kinesiology, 19*, 182-198.

Jacobs, R., & van Ingen Schenau, G. J. (1992). Intermuscular coordination in a sprint push-off. *Journal of Biomechanics, 25*, 953-965.

Jobson, S. A., Hopker, J., Arkesteijn, M., & Passfield, L. (2013). Inter-and intra-session reliability of muscle activity patterns during cycling. *Journal of Electromyography and Kinesiology, 23*, 230-237.

Kadaba, M., Ramakrishnan, H., Wootten, M., Gainey, J., Gorton, G., & Cochran, G. (1989). Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *Journal of Orthopaedic Research, 7*, 849-860.

Kinoshita, H., Bates, B., & DeVita, P. (1985). Intertrial variability for selected running gait parameters. *Biomechanics IX-A* (pp. 499-502). Champaign, IL: Human Kinetics Publishers.

Konrad, P. (2005). *The ABC of EMG.* (No.1.4). Scottsdale: Noraxon INC.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155-163.

Laplaud, D., Hug, F., & Grélot, L. (2006). Reproducibility of eight lower limb muscles activity level in the course of an incremental pedaling exercise. *Journal of Electromyography and Kinesiology, 16*, 158-166.

Li, J., Lu, T., Lin, C., Kuo, M., Hsu, H., & Shen, W. (2017). Soft tissue artefacts of skin markers on the lower limb during cycling: Effects of joint angles and pedal resistance. *Journal of Biomechanics, 62*, 27-38.

Martin, J. C., Diedrich, D., & Coyle, E. F. (2000). Time course of learning to produce maximum cycling power. *International Journal of Sports Medicine, 21*, 485-487.

Martin, J. C., & Brown, N. A. T. (2009). Joint-specific power production and fatigue during maximal cycling. *Journal of Biomechanics*, *42*, 474-479.

Martin, J. C., Davidson, C., & Pardyjak, E. (2007). Understanding sprint-cycling performance: The integration of muscle power, resistance, and modeling. *International Journal of Sports Physiology and Performance, 2*, 5-21.

Martin, J. C., & Nichols, J. (2018). Simulated work loops predict maximal human cycling power. *The Journal of Experimental Biology, 221,* doi: 10.1242/jeb.180109

McDaniel, J., Behjani, N. S. E., S.J., Brown, N. A. T., & Martin, J. C. (2014). Joint-specific power-pedaling rate relationships during maximal cycling. *Journal of Applied Biomechanics, 30*, 423-430.

McGinley, J. L., Baker, R., Wolfe, R., & Morris, M. E. (2009). The reliability of three-dimensional kinematic gait measurements: A systematic review. *Gait & Posture, 29*, 360-369.

Mendez-Villanueva, A., Bishop, D., & Hamer, P. (2007). Reproducibility of a 6-s maximal cycling sprint test. *Journal of Science and Medicine in Sport, 10,* 323-326.

Milner, C. E., Westlake, C. G., & Tate, J. J. (2011). Test–retest reliability of knee biomechanics during stop jump landings. *Journal of Biomechanics, 44*, 1814-1816.

Myers, C. A., Laz, P. J., Shelburne, K. B., & Davidson, B. S. (2015). A probabilistic approach to quantify the impact of uncertainty propagation in musculoskeletal simulations. *Annals of Biomedical Engineering, 43*, 1098-1111.

Neptune, R. R., & Kautz, S. A. (2001). Muscle activation and deactivation dynamics: The governing properties in fast cyclical human movement performance? *Exercise and Sport Sciences Reviews, 29*, 76-81.

Neptune, R. R., Kautz, S. A., & Hull, M. L. (1997). The effect of pedaling rate on coordination in cycling. *Journal of Biomechanics, 30*, 1051-1058.

Pataky, T. C. (2010). Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *Journal of Biomechanics, 43*, 1976-1982.

Pini, A., Markström, J. L., & Schelin, L. (2019). Test–retest reliability measures for curve data: An overview with recommendations and supplementary code. *Sports Biomechanics,* doi: 10.1080/14763141.2019.1655089

Raasch, C. C., Zajac, F. E., Ma, B., & Levine, W. S. (1997). Muscle coordination of maximum-speed pedaling. *Journal of Biomechanics, 30*, 595-602.

Ryan, M. M., & Gregor, R. J. (1992). EMG profiles of lower extremity muscles during cycling at constant workload and cadence. *Journal of Electromyography and Kinesiology, 2*, 69-80.

Umberger, B. R., & Martin, P. E. (2001). Testing the planar assumption during ergometer cycling. *Journal of Applied Biomechanics, 17*, 55-62.

van Ingen Schenau, G. J., Boots, P. J. M., De Groot, G., Snackers, R. J., & Van Woensel, W. W. L. M. (1992). The constrained control of force and position in multi-joint movements. *Neuroscience, 46*, 197-207.

van Ingen Schenau, G. J., Van Woensel, W. W. L. M., Boots, P. J. M., Snackers, R. W., & De Groot, G. (1990). Determination and interpretation of mechanical power in human movement: Application to ergometer cycling. *European Journal of Applied Physiology and Occupational Physiology, 61*, 11-19.

Wakeling, J. M., Blake, O. M., & Chan, H. K. (2010). Muscle coordination is key to the power output and mechanical efficiency of limb movements. *The Journal of Experimental Biology, 213*, 487-492.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research, 19*, 231-240.

Winter, D. A. (2009). Biomechanics and motor control of human movement. (4th ed., pp. 70-73). Hoboken, NJ: Wiley.

Wren, T. A., Do, K. P., Rethlefsen, S. A., & Healy, B. (2006). Cross-correlation as a method for comparing dynamic electromyography signals during gait. *Journal of Biomechanics, 39*, 2714-2718.

Yavuzer, G., Öken, Ö, Elhan, A., & Stam, H. J. (2008). Repeatability of lower limb three-dimensional kinematics in patients with stroke. *Gait & Posture, 27*, 31-35.

Table 1: Between-sessions reliability for kinematic and kinetic variables, * indicates significant difference between sessions (*p* < 0.05), ICC(3,k) = Between-sessions intraclass correlation with lower (LB) and upper (UB) bound confidence intervals, SEM = standard error of measurement, MDD = minimal detectable difference

Variable	Units	Mean(SD)	Mean difference	*p*	ICC	95%	95%	SEM	MDD

		Session 1	Session 2			(3,*k*)	LB	UB		
Power (average for left crank)	W	445.3 ± 95.7	438.8 ± 111.5	-6.5	0.429	0.979	0.938	0.993	4.3	12
Pedalling rate	rpm	134.8 ± 1.3	134.7 ± 1.4	-0.2	0.021*	0.986	0.935	0.996	0.0	0.1
Max effective crank force	N	593.3 ±126.2	579.0 ± 130.9	-14.4	0.072	0.986	0.952	0.996	3.2	9
Max ineffective crank force	N	603.5 ± 172.1	605.3 ± 165.4	1.8	0.944	0.923	0.756	0.975	25.9	72
Min ineffective crank force	N	-192.7 ± 65.2	-207.3 ± 82.3	-14.7	0.136	0.937	0.805	0.980	8.7	24
Max instantaneous crank power	W	1387.2 ± 309.2	1348.4 ± 316.5	-38.7	0.043*	0.986	0.946	0.996	7.7	21
Peak ankle plantarflexion angle	°	141.7 ± 11.3	142.3 ± 11.5	0.6	0.446	0.983	0.948	0.994	0.4	1.1
Peak ankle dorsiflexion angle	°	113.1 ± 5.0	113.8 ± 5.8	0.7	0.281	0.955	0.863	0.985	0.5	1.3
Peak knee extension angle	°	142.7 ± 6.4	143.5 ± 5.7	0.8	0.489	0.864	0.580	0.956	1.6	4.4
Peak knee flexion angle	°	70.0 ± 3.6	70.2 ± 3.4	0.2	0.715	0.857	0.550	0.954	1.0	2.6
Peak hip extension angle	°	68.1 ± 5.0	68.4 ± 4.6	0.3	0.720	0.893	0.665	0.966	1.0	2.8
Peak hip flexion angle	°	26.1 ± 4.3	25.6 ± 4.2	-0.5	0.447	0.916	0.746	0.973	0.7	1.9
Peak ankle plantarflexion angular velocity	°/s	236.6 ± 65.7	247.1 ± 65.0	10.4	0.441	0.839	0.509	0.948	19.7	55
Peak ankle dorsiflexion angular velocity	°/s	-262.0 ± 91.2	-268.5 ± 107.2	-6.6	0.561	0.957	0.868	0.986	8.6	24
Peak knee extension angular velocity	°/s	472.8 ± 43.2	479.1 ± 33.8	6.3	0.434	0.838	0.504	0.948	11.8	33
Peak knee flexion angular velocity	°/s	-507.5 ± 57.6	-513.3 ± 43.6	-5.8	0.635	0.772	0.279	0.927	21.4	59
Peak hip extension angular velocity	°/s	265.6 ± 29.1	273.8 ± 21.9	8.2	0.141	0.814	0.447	0.939	8.5	24
Peak hip flexion angular velocity	°/s	-277.6 ± 30.7	-273.4 ± 35.1	4.2	0.390	0.924	0.769	0.975	4.9	14
Peak ankle plantarflexion moment	N.m	78.6 ± 18.6	81.4 ± 20.2	2.8	0.372	0.910	0.729	0.971	3.4	9
Peak ankle dorsiflexion moment	N.m	-14.0 ± 7.0	-12.3 ± 6.0	1.8	0.049*	0.928	0.743	0.978	0.8	2
Peak knee extension moment	N.m	90.0 ± 34.5	82.9 ± 33.5	-7.1	0.028*	0.965	0.852	0.990	2.0	6
Peak knee flexion moment	N.m	-50.7 ± 20.9	-57.7 ± 15.0	-7.0	0.151	0.697	0.127	0.900	9.4	26
Peak hip extension moment	N.m	132.3 ± 30.7	140.4 ± 32.8	8.1	0.086	0.919	0.737	0.974	4.6	13
Peak hip flexion moment	N.m	-47.7 ± 26.1	-41.3 ± 17.0	6.5	0.115	0.870	0.600	0.958	5.1	14
Maximum ankle power	W	259.6 ± 111.7	258.5 ± 107.8	-1.1	0.937	0.951	0.846	0.984	10.9	30
Maximum knee power	W	659.6 ± 321.7	620.4 ± 253.6	-39.2	0.160	0.968	0.901	0.990	17.6	49
Maximum hip power	W	519.8 ± 186.3	578.1 ± 153.0	58.3	0.104	0.826	0.474	0.944	52.1	144

Figure captions

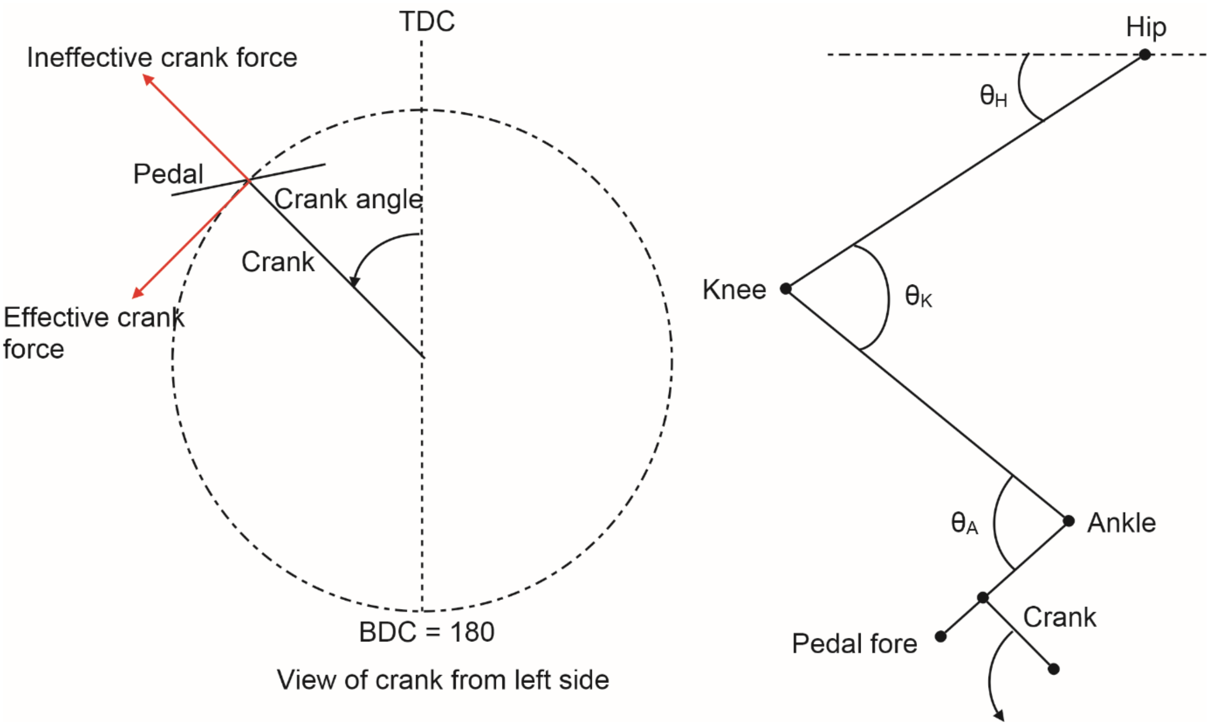

Figure 1: Joint angle and crank forces convention. TDC = top dead centre, BDC = bottom dead centre, θ_H = hip angle, θ_K = knee angle, θ_A = ankle angle

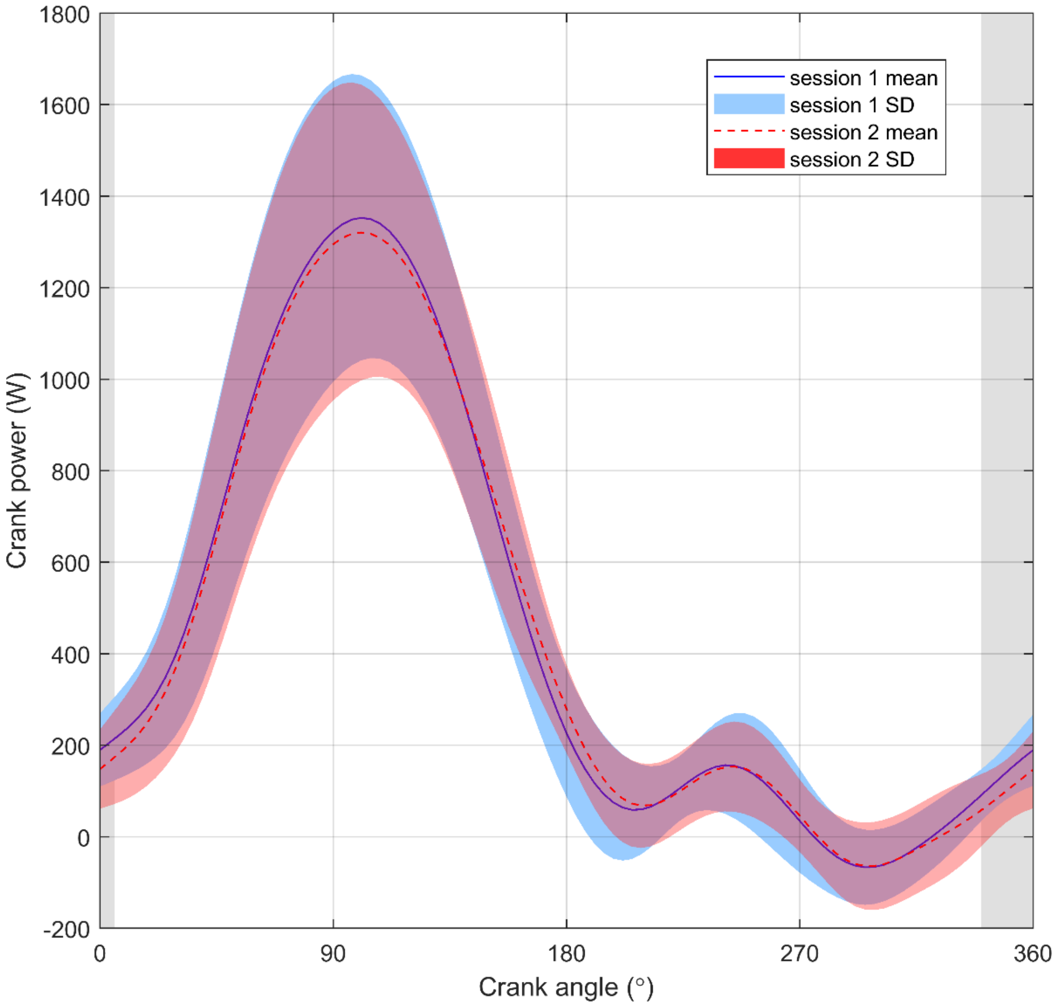

Figure 2: Crank power: group means for session one and two. Areas of the graph shaded grey where the Statistical parametric mapping (SPM) is significant.

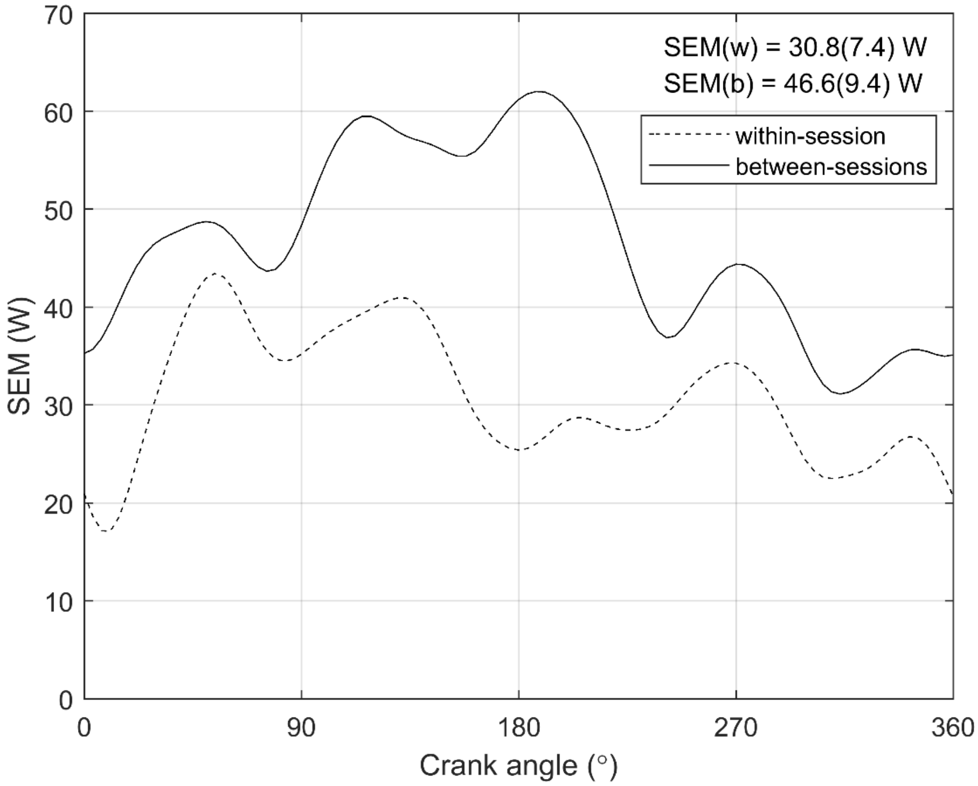

Figure 3: Crank power: standard error of measurement (SEM) within- and between-session. Mean and standard deviation of SEM within-session (w) and between-sessions (b) over complete crank cycle.

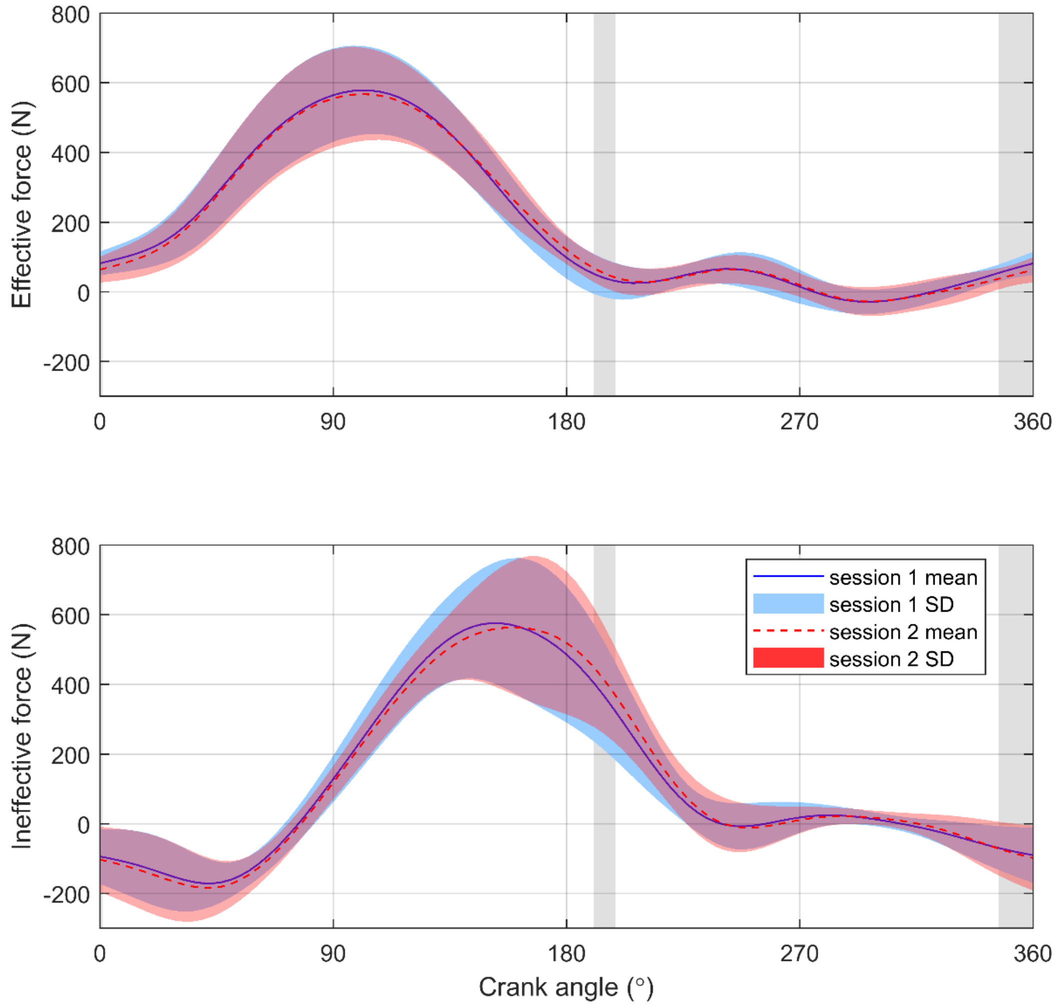

Figure 4: Crank forces: group means for session one and two. Areas of the graph shaded grey where the Statistical parametric mapping (SPM) is significant.

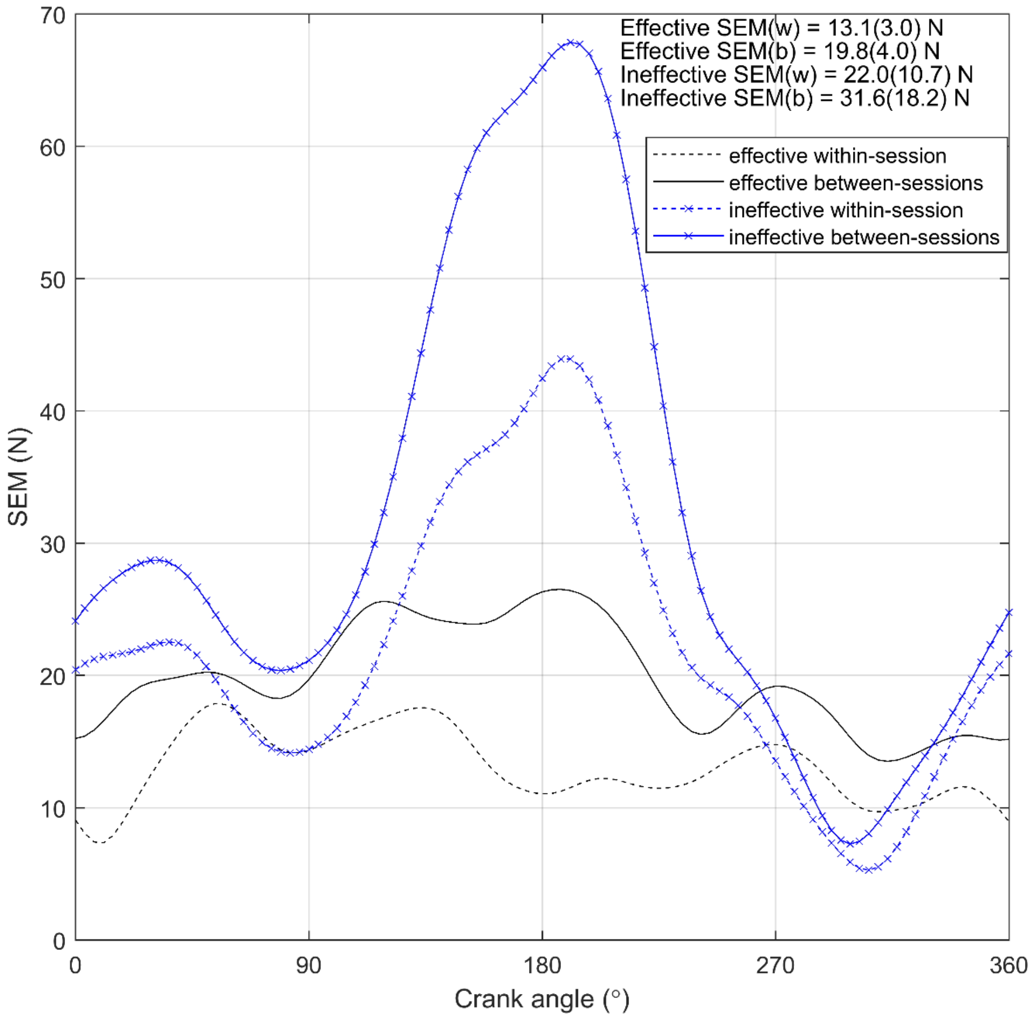

Figure 5: Crank forces: standard error of measurement (SEM) within- and between-session. Mean and standard deviation of SEM within-session (w) and between-sessions (b) over complete crank cycle.

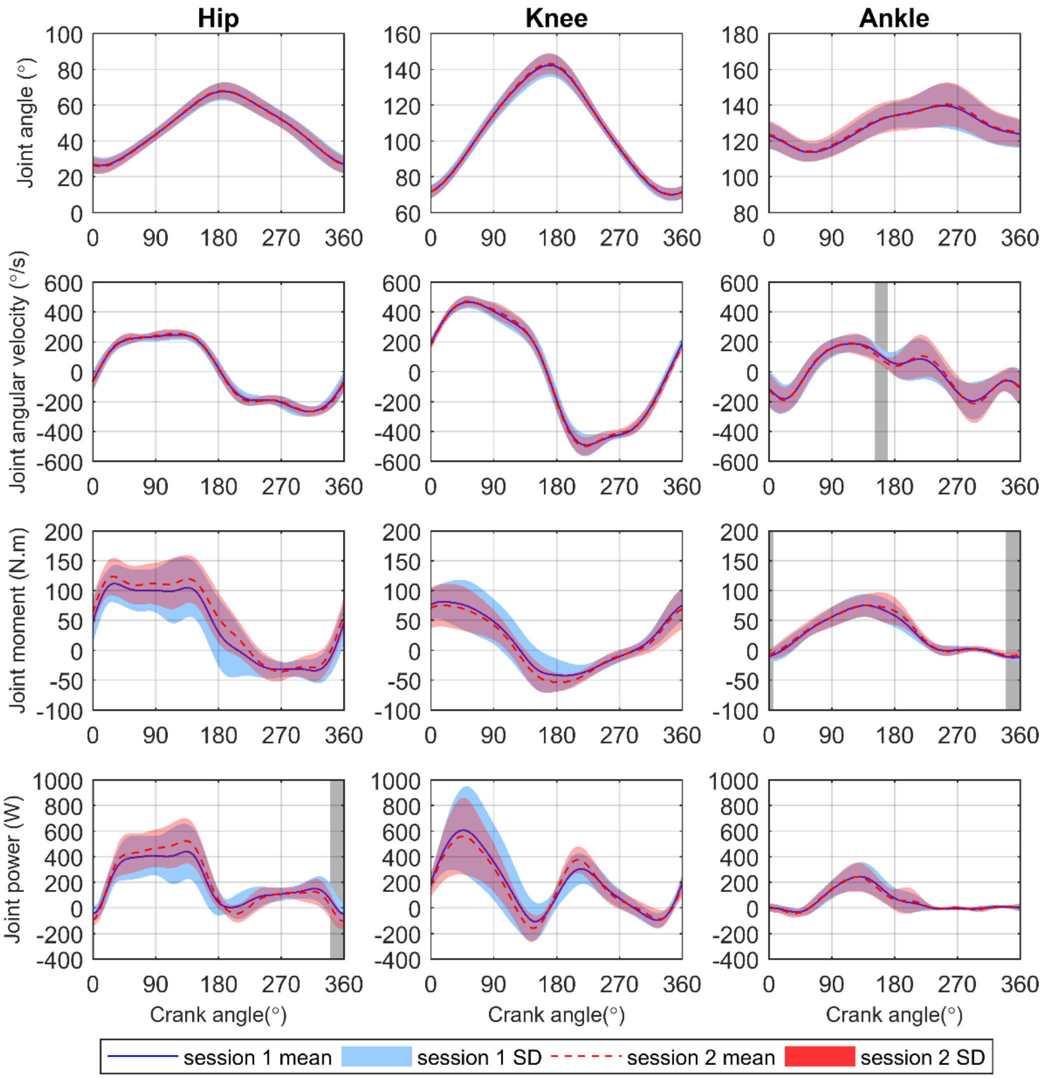

Figure 6: Joint angles, angular velocities, moments and powers: group means for session one and two. Areas of the graph shaded grey where the Statistical parametric mapping (SPM) is significant.

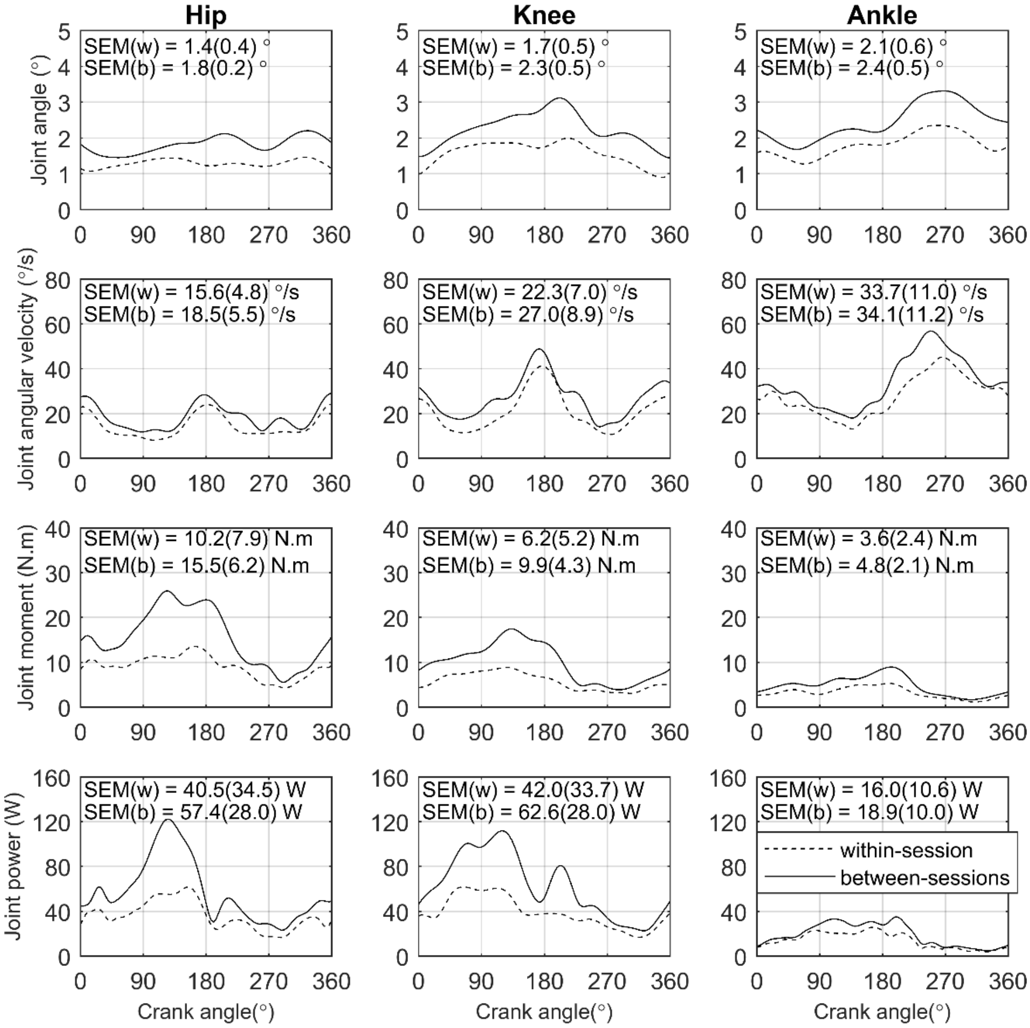

Figure 7: Joint angles, angular velocities, moments and powers: standard error of measurement (SEM) within- and between-session. Mean and standard deviation of SEM within-session (w) and between-sessions (b) over complete crank cycle.

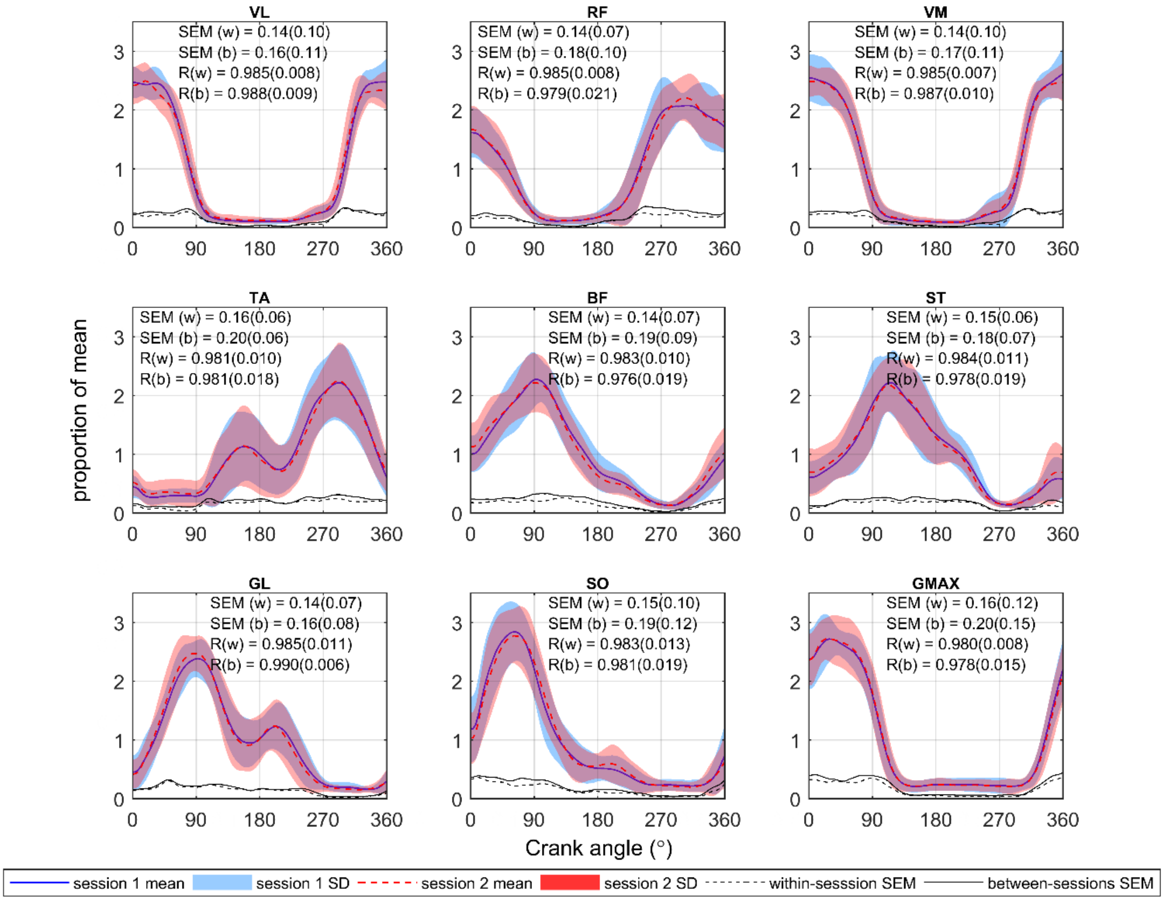

Figure 8: EMG linear envelopes (normalised to mean value in signal) for each muscle: group means for session one and two and standard error of measurement (SEM) within- and between-session. VL = vastus lateralis, RF = rectus femoris, VM = vastus medialis, TA = tibialis anterior, BF=biceps femoris, ST= semitendinosus, GL = gastrocnemius lateralis, SO = soleus, GMAX = gluteus maximus. Mean and standard deviation of SEM within-session (w) and between-sessions (b) over complete crank cycle.