Reproducibility of an isokinetic eccentric muscle endurance task
Abstract

Despite the widespread use of isokinetic eccentric muscle endurance protocols, there appear to be no studies that have examined the absolute and relative reliability of such tests. The aim of this study was to examine the reliability of an isokinetic eccentric knee muscle endurance task. Fifteen healthy untrained males volunteered to take part in the study and written informed consent was obtained. The procedures received approval from the University Research Ethics Committee. Participants visited the laboratory on three separate occasions; two weeks prior to testing to familiarise them with the experimental procedures, and on two separate occasions, two weeks apart and at the same time of day. Isokinetic eccentric knee extension and flexion movements of the dominant limb were performed using a calibrated dynamometer (Biodex system 3). Range of motion during testing was set using voluntary maximal full extension (0 rad) to 1.57 rad of knee flexion and testing occurred at 1.56 rad.s⁻¹. Participants performed four maximal efforts to determine maximal peak torque (PT). A 120 s rest period was then given before completing the all-out endurance test of 50 continuous eccentric repetitions. Data were gravity corrected and windowed to only include constant velocity periods. Repeated measures ANOVAs were used to investigate differences in maximal peak extension and flexion torque and the maximal torque measurement recorded during the endurance test. Absolute and relative reliability of the torque fatigue index (TFI), work fatigue index (WFI) and total work (TW) were assessed through calculation of intra-class correlation coefficients (ICC), coefficient of variation (CV), repeatability coefficients (RC) and absolute 95% limits of agreement (LOA) using the methods described by Bland & Altman (1986). Differences between the strength test and the endurance trial for PT were found to be non-significant for both quadriceps (323 Vs 323 Nm) and hamstrings (183 Vs 178 Nm). The ICCs revealed significant (p < 0.05) positive moderate to strong correlations (r = 0.44 – 0.94) across repeated trials for all parameters except hamstring torque fatigue (p =
0.11) and hamstring work fatigue index (P = 0.08). Coefficient of variation values ranged from to % and were large for WFI and TFI but acceptable for TW for both the extensors and flexors respectively. Repeatability Coefficients and LOA indicated systematic bias in repeated trials for both TFI and WFI for extensors and flexors, ranging from -3% to -10%, with less fatigue evident in the 2nd test. There was also systematic bias for TW with more work performed during the 2nd test compared to the 1st for both extensors and flexors. The random error was large for all variables and there was greater random error in the hamstrings compared to the quadriceps. Along with moderate to strong ICC’s and large CV, these data suggest that there is small systematic bias in repeated eccentric muscle actions for both the quadriceps and hamstrings, however the random error is large despite habituation procedures. Therefore, data from an isokinetic eccentric muscle endurance task should be viewed with a degree of caution even when participants receive considerable habituation prior to assessment.

Keywords: Fatigue, Knee joint, Reliability, Limits of Agreement, Bias, Random Error
**Introduction**

Muscle endurance and fatigue are common in daily life and clinical settings, yet the mechanisms associated with muscle endurance and fatigue are far from understood. Muscle fatigue is difficult to define but it has been referred to as the failure to maintain the required or expected force (Westerdlad, Lee, Lannergren, & Allen, 1991). Fulco and colleagues (Fulco, Rock, Muza, Lammi, Cymerman, & Lewis, 2000) have suggested that progressive muscle fatigue during endurance exercise is typically unobservable, unquantifiable, and not systematically studied during conventional ergometry such as cycling or treadmill running. Nicholas, Gauthier, Michaut, & Davenne (2007) suggest that the rate of the development of fatigue is task and muscle action dependent. It has therefore been recommended that any assessment of muscular endurance and fatigue must measure dynamic muscle forces (Fulco et al., 2000; Vollestad, 1997) and repeated isokinetic muscle actions have become a popular approach when examining muscle endurance (Vollestad, 1997). To improve reproducibility, and to isolate the events associated with fatigue, there is an emphasis on measuring the progressively declining dynamic torque (Kaminski, Godfrey, Braith, & Stevens, 2000).

The majority of studies that have utilised isokinetic dynamometry to examine muscle endurance have focused on concentric (CON) rather than eccentric (ECC) muscle actions (Gleeson & Mercer, 1992; Kellis, Kellis, Gerodimos, & Manou, 1999; Larsson, Karlsson, Eriksson, & Gerdle, 2003). It is well recognised that CON and ECC muscle actions are under different neuromuscular regulation and therefore their response to an endurance task will be different (Nicolas *et al.*, 2007). Some studies have examined repeated eccentric actions of the knee extensors/flexors (Bilcheck 1994; Emery, Sitler, & Ryan, 1994; Kaminski *et al.*, 2000), elbow flexors (Gleeson, Eston, Marginson, & McHugh, 2003), dorsiflexors (Pasquet, Carpentier, Duchateau, & Hainaut, 2000) and shoulder rotators (Mullaney & McHugh, 2006),
but have predominantly used the exercise challenge to elicit delayed onset muscle soreness (DOMS). Most studies that have examined the effect of fatiguing repeated eccentric actions have demonstrated a degree of force preservation, which has been attributed to the activation of unrecruited fibres that compensate for fatigued fibres (Hortobagyi, Tracy, Hamilton, & Lambert, 1996). Despite the widespread use of such protocols there appear to be no studies that have examined the absolute and relative reliability of such exercise challenges. Relative reliability has been defined as the degree to which individuals maintain their position in a sample with repeated measurements and absolute reliability the degree to which repeated measurements vary for individuals (Atkinson & Nevill 1998).

One of the common difficulties in examining muscle endurance and fatigue from repeated isokinetic actions is the choice of outcome measure to use to demonstrate that fatigue has taken place. The most common outcome measure for muscle endurance is the fatigue index. Endurance is normally expressed as the percentage decline in peak torque from the first three or five repetitions to the last three or five, usually during a set of 50 repetitions. Data using this technique must be viewed with a degree of caution as a recent study has shown that adult participants do not produce a maximal effort at the beginning of an endurance trial and that not all motor units are recruited (Pincivero, Gear, Sterner, & Karunakara, 2000; Wretling & Henriksson-Larsen 1998). It has been shown that a 50 repetition task of concentric muscle actions elicits a two phase response, which includes a ‘fatigue phase’ and an ‘endurance phase’ (Larsson et al., 2003). Data from ECC actions are less readily available and it remains to be identified if there is a similar two-phase response.

It is well recognised in the literature that single maximal isokinetic concentric actions are more reliable than eccentric actions (Deighan, De Ste Croix, & Armstrong, 2003; De Ste
Croix, Armstrong, & Welsman, 2003), and that extensor movements appear to be more reliable than flexor movements (Deighan et al. 2003; De Ste Croix et al., 2003). Most studies have shown isokinetic concentric muscle endurance tests to be reliable (Burdett & Van Swearingen 1987; Pincivero, Lephart, & Karunakara, 1997; Pincivero, Gear, & Sterner, 2001), however, the majority of endurance studies using isokinetic dynamometers have focused upon decline in force production rather than decreases in contractile work. Kaminski et al., (2000) suggest that isokinetic work performed may be more representative of muscle function as it accounts for the force output throughout the range of motion. This is reinforced by Kannus, Cook, & Aosa, (1992) who have also suggested that work ratios are as valuable as peak torque measurements when assessing muscular fatigue. However, Burdett & Swearingen (1987) reported low reliability for a work fatigue index after 25 repeated concentric actions as did Pincivero and colleagues (1997, 2001) after a 30 repetition task. It has been suggested therefore that a linear model should be used to describe the rate of isokinetic fatigue, rather than a fatigue index (Pincivero et al., 2001). These conflicting data demonstrate the difficulty in comparing previous studies of isokinetic endurance as differing numbers of repetitions and differing methods to calculate the fatigue index have been used. Whether indices of peak torque or work performed are more appropriate when examining eccentric fatigue remains to be identified.

Most isokinetic endurance studies have used differing statistical analyses in describing reliability (Burdett & Van Swearingen 1987; Pincivero et al., 1997, 2001) but only one study, on children, has used the techniques described by Bland & Altman (1986) (De Ste Croix et al., 2003). The 95% limits of agreement (LOA) are widely considered by researchers as the most appropriate statistic to report when assessing the level of agreement between test-retest measurements (Atkinson & Nevill, 1998). To our knowledge no studies have examined the
reliability of an isokinetic eccentric muscle endurance task. Therefore the aims of the present study were to examine the absolute and relative reliability of a repeated isokinetic eccentric muscle endurance task of the knee extensors and flexors.

Methods
Participants

Fifteen healthy and physically active males, from a University population, volunteered to take part in the study, and written informed consent was obtained from all participants. The study procedures received approval from the University Research Ethics Committee. Participants visited the laboratory two weeks prior to testing to fully familiarise them with the laboratory environment and the experimental procedures. For experimental tests the participants visited the laboratory on two separate occasions, exactly two weeks apart and at the same time of day. Testing took place in a quite laboratory where the room temperature was maintained at between 19-21°C and 2 investigators were present at all sessions.

Anthropometry

Age was computed from date of birth and date of testing. Stature and body mass were measured according to the techniques described by Weiner & Lourie (1981). Stature was measured using a Stadiometer (Holtain, Crymych, Dyfed, UK) and body mass was measured using calibrated scales (Cranlea, Birmingham, UK).

Endurance Test

Isokinetic eccentric knee extension and flexion were performed using a calibrated dynamometer (Biodex system 3, Shirley, New York, USA). All testing sessions involved a standardised procedure, including a warm up of 2 min cycling on a Monark cycle ergometer.
814E (Varberg, Sweden) at 50W and 2 min of static stretching the hamstring and quadriceps muscles. Participants were placed in a seated position with the seatback tilted at 1.48 rad hip flexion from the anatomical position. The lever arm of the dynamometer was lined with the lateral epicondyle of the knee, and the force pad placed approximately 3 to 5 cm superior to the medial malleolus with the foot in a plantigrade position. Range of motion during testing was set using voluntary maximal full extension (0 rad) to 1.57 rad of knee flexion. Cushioning was set using a hard deceleration and therefore 1.40 rad constituting the tested range. All settings including seat height, seat length, dynamometer height and lever arm length were noted so as to be identical for the second test session. At the start of each test session the subject was asked to relax their leg so that passive determination of the effects of gravity on the limb and lever arm could be accounted for. The dominant limb, determined from kicking preference (Rizzardo, Wessel, & Bay, 1998), was examined, and testing occurred at a velocity of 1.56 rad.s⁻¹. Pilot work also showed that when subjects were fatigued in the reactive eccentric mode they could not maintain the required torque output throughout the range of motion, subsequently causing stalling of the lever arm. Therefore, passive eccentric mode was chosen so that the full range of movement would be completed for every action. Participants were instructed to resist the lever arm during extension and flexion as hard as possible, with extension undertaken first. The participants performed four continuous maximal efforts to determine maximal peak torque. A 120 s rest period was then given before the endurance test. The test consisted of 50 continuous eccentric repetitions. Participants were encouraged to give a maximal effort for each action by using both visual feedback from torque curves and strong verbal encouragement. The verbal encouragement was given by the same experimenter and was standardised by using key words such as ‘resist’ and ‘hard and fast as possible’. Immediately after the test participants completed a 2 min cool down on the cycle ergometer at 50W.
Statistical analyses

Isokinetic peak extension (PET) and peak flexion (PFT) torque were gravity corrected and windowed using the Biodex software to only include constant velocity periods. Maximal PET and PFT were determined as the highest values recorded. Total work for extensors and flexors for the complete 50 repetitions was recorded from the windowed torque curves. Average torque and average work for the first 3 and the last 3 repetitions were calculated, with the percentage difference between these values used to represent percentage decline in torque and work (i.e., fatigue index). Data were analysed using SPSS for Windows (SPSS Inc., Chicago, USA, V16.0). Repeated measures ANOVAs were used to investigate differences in maximal PET and PFT and the maximal torque measurement recorded during the endurance test.

Relative reliability was assessed through calculation of intra-class correlation coefficients (ICC) and absolute reliability through co-efficient of variation (CV) repeatability coefficients (RC) and 95% limits of agreement (LOA) using the methods described by Bland & Altman (1986) for absolute and percentage decline in PET, PFT, total work and percentage decline in work for extensors (WkET) and flexors (WkFT). The relationship between the differences and mean absolute values between test 1 and test 2 were examined using Pearson correlation coefficients. No significant relationships were observed for all outcome variables and therefore the data were deemed to be homoscadastic and did not require logarithmic transformation. Bland and Altman plots were also determined to provide a visual interpretation of the data and to determine reproducibility bias. In these plots the mean of tests 1 and 2 (x-axis) were plotted against the differences between test 1 and 2 (y-axis).
Results

Participant characteristics were: age 26 (±7) y, stature 1.83 (±0.07) m, bodymass 86 (±12) kg, (values are mean (SD)). Table 1 shows peak torque, torque fatigue index (TFI), work fatigue index (WFI) and total work for extensor and flexor movements and for both test and re-test. Maximal peak torque difference between the preliminary strength test and the endurance trial were found to be non-significant for both quadriceps (p > 0.05) and hamstrings (p > 0.05) respectively.

**************Table 1 Here**************

ICCs indicated significant (p < 0.05) positive moderate-to-strong correlations between test 1 and 2 for all parameters except hamstring torque fatigue (p = 0.11) and hamstring work fatigue index (p = 0.08), ranging from r = 0.44 – 0.94. The RCs and LoA indicated systematic bias for both TFI and WFI for extensors and flexors, ranging from -3 to -10, with less fatigue evident in the 2nd test. There was also systematic bias for TW with more work performed during the 2nd test compared to the 1st. RCs and LoA demonstrated greater systematic bias of the hamstrings for TFI and WFI compared to the quadriceps. However, there was more systematic bias for the quadriceps for TW compared to the hamstrings. LoA demonstrated large random error for all variables (see figure 1).

**************Table 2 Here**************

The 95% LoA are presented graphically for PTFI, WFI and TW for hamstrings and quadriceps respectively (figure 1). The central line represents the mean difference (i.e bias); top line represents +1.96 SD and bottom line -1.96 SD of differences.
Discussion

It is widely accepted that muscle fatigue during a muscle endurance challenge must include assessment of maximal voluntary contraction force (Vollestad, 1997). The main advantage of using isokinetic force production to investigate fatigue during an endurance challenge is that all levels of the muscle activation process are examined. Any decline in the force generating capability can be due to either central fatigue or to factors distal to the motorneurones (Vollestad, 1997). Previous studies have examined the reliability of isokinetic concentric fatigue tasks (Burdett & Van Swearingen 1987; De Ste Croix et al., 2003, Pincivero et al., 1997, 2001), however, the current study appears to be the first to examine the reliability of fatigue during an isokinetic eccentric endurance trial of the knee extensors and flexors.

One previous study, using an isokinetic concentric endurance protocol, identified that adult subjects do not perform maximal efforts at the start of an endurance protocol (Wretling & Henriksson-Larsen, 1998). This has important implications in the calculation of fatigue, especially if these actions are chosen as a reference for calculating a fatigue index. Data from the current study indicated no significant difference between maximal eccentric torque measured during a preliminary strength test and maximal eccentric torque achieved at the start of the endurance trial for both extensors and flexors. These data suggest that the maximal eccentric torque produced at the start of an endurance test are not significantly different to those produced during a preliminary strength test and that these initial actions can be validly used to calculate a fatigue index or reduction in torque output from linear slope regression (Pincivero et al., 2001). The lack of a difference between the preliminary strength test and the endurance trial outcomes may be due to the habituation of subjects to the eccentric movement prior to testing and the use of the passive eccentric mode on the Biodex dynamometer.
Motivation has been identified as important during strength testing and even more so for endurance tests where subjects are asked to give up to 50 repeated maximal efforts (Burdett & Van Swearingen, 1987). Thus the reliability of TW over 50 repetitions is a good indicator that a similar effort is given during repeated trials. The ICCs of 0.91 and 0.94 for hamstrings and quadriceps TW in the present study indicate that a similar effort was given during the repeated endurance task. These data are similar to ICCs reported previously for total work of the quadriceps (0.89 and 0.98) and hamstrings (0.95 and 0.91), albeit from repeated concentric actions (Burdett & Van Swearingen 1987, Kellis et al., 1999). This is further supported in the current study by the acceptable CV for both hamstring TW (5.7%) and quadriceps TW (6.2%). The use of the Biodex passive eccentric mode rather than the reactive eccentric mode may have aided in the ability to apply force throughout the entire range of motion during the fatigue task. During pilot work when subjects were fatigued they found it difficult to maintain force output throughout the range of motion causing ‘juddering’ of the lever arm in the more commonly used reactive mode, which would affect the calculation of work outcomes.

Both the work and torque fatigue index in the present study showed a decline in output during 50 maximal isokinetic eccentric actions. These data are in agreement with others who have demonstrated a 53% decrease in elbow flexors after 100 repeated eccentric actions (Linnamo, Bottas, & Komi, 2000) and repeated actions of the knee extensors (Brown, Child, Day, & Donnelly, 1997; MacIntyre, Reid, Lyster, Szasz, & McKenzie, 1996). However, others have shown that repeated eccentric muscle actions are extremely fatigue resistant despite high force production (Emery et al., 1994; Hortobagyi et al., 1996; Tesch, 1990). Work by Warren and colleagues (2000) indicated that a 2nd bout of 50 repeated eccentric actions demonstrated an increased activation of slow motor units and reduced activation of fast motor units. Data from
the current study does not support the suggestion that activation of unrecruited fibres or selective recruitment of slow twitch fibres compensates for fatigued fibres during a single bout of repeated eccentric actions. However, whether repeated performance of an eccentric fatigue task results in a ‘learned’ altered pattern of muscle fibre recruitment requires further investigation.

It has been suggested that an ICC of 0.88 or above and CV not exceeding ± 6% demonstrates good reliability in isokinetic strength testing (Gleeson & Mercer, 1996). The relative reliability for the WFI and TFI in the current study are lower than this recommended value for both the knee extensors and flexors. However, values presented are higher than previous values from adult subjects for the fatigue index of knee extensors (ICC = 0.26) from concentric actions (Pincivero et al., 2001). Data from a study on children has demonstrated much higher reliability based on ICC’s for both work and torque fatigue index during concentric knee extension and flexion movements (De Ste Croix et al., 2003). It is difficult to attribute a reason for the higher reliability in that study, but it may be due to lower initial torque values in children, especially from concentric actions. Larsson et al., (2003) also reported higher ICC’s from 100 repeated concentric actions ($r = 0.93-0.95$). It is difficult to compare our data with Larsson et al., (2003) as concentric and eccentric actions are known to be under separate neuromuscular regulation (Kellis & Baltzopoulos, 1998).

The ICC’s in the present study revealed that the WFI and TFI were more reliable for the extensors than the flexors ($r = 0.47$ v 0.66 and $r = 0.44$ v 0.58 respectively) which is in agreement with data presented for repeated concentric isokinetic actions (Burdett & Van Swearingen, 1987; De Ste Croix et al., 2003). Likewise the WFI appears to be more reliable than the TFI and this is in agreement with a previous study from repeated concentric actions
in children (De Ste Croix et al., 2003). Porter and colleagues (2002) state that examining the
confidence intervals of the mean differences acts as a guide for determining whether true
differences exist between repeated trials. Their data indicated that measurement of torque was
more reliable than that of work due to the greater confidence intervals associated with work
outcomes. This was attributed to the factors associated with the calculation of work, most
noticeably that participants do not complete the full range of motion as they fatigue. This has
important implications, since the calculation of work is a function of torque produced over the
range of motion (area under the torque curve). Results from the present study do not support
the findings of Porter et al., (2002); this may be because by using the passive eccentric mode
in the current study the full range of motion was completed for each muscle action. Care
should nevertheless be taken when making conclusions about the relative reliabilities of peak
torque measurements based on ICCs, as the differences in ICCs may not be mirrored in other
reliability statistics. This is evident in the current study, when data were analysed for bias and
random error using RC and LoA. Nevertheless, the high CV values in the current study
support the notion that eccentric extensor movements are more reliable that flexor movements
and that the WFI is more reliable than TFI.

As stated earlier, Bland & Altman (1986) demonstrated that the use of correlation coefficients
can be misleading in test-retest studies as they measure the strength of the relationship but not
how well the two variables agree with each other. The high CV values in the current study
for TFI and WFI for the extensors and flexors are much higher than the acceptable limits and
would suggest that calculation of a fatigue index may be inappropriate when examining
isokinetic muscle endurance and fatigue. Pincivero and colleagues (2001) suggest that
isokinetic dynamometry may not be the modality of choice for the determination of a fatigue
index. Kaminski et al. (2000) suggest that the traditional fatigue index could be replaced with
a ‘fatigue resistance index’ that effectively quantifies changes in muscle force production. It may be that the use of a traditional fatigue index in the current study may in part account for the high CV values observed. However, caution must be taken when interpreting CV values as they correspond to an arbitrary analytical goal and importantly the use of CV assumes heteroscedasticity (Atkinson & Nevill, 1998) but we have demonstrated that our data is homoscedastic.

The technique of Bland & Altman (1986) has, to our knowledge, not been applied to data from an isokinetic eccentric fatigue protocol. Systematic bias represents the tendency for a measurement to be different in either a positive or negative direction between assessments (Atkinson & Nevill, 1998). The relatively small systematic bias in the current study demonstrated a reduction in WFI and TFI for both the hamstrings and quadriceps from test 1 to test 2 (e.g., 25% to 18% for extensor TFI). There was also a tendency for total work to increase for both hamstrings and quadriceps between test occasions. This small bias was probably reduced due to good familiarisation but could be further reduced with additional practice. However, in most practical and clinical settings extensive habituation is not always possible. The systematic bias in the present study, demonstrating a lower fatigue index for both torque and work at test 2, may possibly be attributed to protective effects that may exist after initial bouts of eccentric exercise. Brown et al., (1997) reported a prophylactic effect following a single bout of relatively few eccentric actions, reporting no significant increase in serum creatine kinase in a 2nd bout (3 weeks later). Likewise, Warren et al., (2000) reported a reduction in the fatigue of the anterior crural muscles during a 2nd bout of 50 repeated eccentric actions 7 days after the initial test (18% Vs 29% respectively). They attributed this reduction in fatigue to an increased reliance on slower motor units based on similar total electromyography recorded neural activation (root mean square) but a concomitant decrease
in the median frequency. It is possible therefore that the small systematic bias in the current study may be a product of altered patterns in motor unit recruitment to preferably recruit slow units during lengthening muscle actions to resist fatigue in subsequent bouts. Alternatively it may be due to damage of fast twitch motor units during the initial bout. Further study is needed using multiple repeated tests to examine if systematic bias and possible learning effect is reduced with additional practice/habituation.

Random error is the chance fluctuation in a measurement and is caused by inherent biological and mechanical variation (Iga et al., 2006). The error component of the LoA indicates the level of random variation with a probability of 95% (Atkinson & Nevill, 1998). The random error for all variables in the current study was large and therefore care should be taken when drawing conclusions from data from repeated eccentric muscle endurance tasks. This judgement is made on the application of the measurement error to the real application of the measurement tool (Atkinson & Nevill, 1998); in this case, whether the variability associated with the measurement is small enough to allow the detection of expected changes through eccentric endurance training. To our knowledge there are no studies that have examined the changes in eccentric endurance following a period of eccentric endurance training. Therefore it is not possible to ascribe analytical goals to describe whether the limits of agreement are narrow enough for the test to be of practical use. Others have used analytical goals to report acceptable random error from single eccentric muscle actions of the hamstrings, albeit in young boys (Iga et al., 2006). The large random error reported in the current study may also have implications for sample size in future investigations that intend to examine isokinetic eccentric endurance changes. It is difficult to identify reasons for the large random error in the current study but it may be associated with the relatively small sample size (Bland & Altman suggest n > 40) and the unaccustomed nature of eccentric testing.
To our knowledge, no previous studies have examined the bias and random error of repeated isokinetic eccentric tasks and therefore we have no data with which to directly compare our findings. Despite moderate ICC’s and relatively small systematic bias, we would suggest that the high CVs and large random error associated with a repeated eccentric task precludes the use of such tests as a reproducible indicator of isokinetic eccentric muscle endurance unless extensive habituation procedures and numerous repeat trials are in place.
References


Kellis, E.; Baltzopoulos, V., 1998 Muscle activation differences between eccentric and concentric isokinetic exercise. Medicine & Science in Sports & Exercise 30 1616-1623


