Clinical and Experimental Pharmacology and Physiology (2002) **29**, 39–47

MUSCLE STRENGTH AND AGEING: METHODOLOGICAL ASPECTS OF ISOKINETIC DYNAMOMETRY AND ANDROGEN ADMINISTRATION

Lam P Ly and David J Handelsman

Department of Andrology, Concord Hospital, ANZAC Research Institute, University of Sydney, Sydney, New South Wales, Australia

SUMMARY

1. To evaluate interventions aiming to increase muscle strength in older men, it is necessary to use an objective and reproducible method to measure strength. The most reliable method to evaluate muscle strength is isokinetic peak torque (PT) measured by a dynamometer; however, raw PT varies with differences in body size and an optimal scaling for body physique to control confounding effects of body size is necessary to make the most valid comparisons.

2. The present study was designed to estimate the effects of age (part A) and androgen administration (part B) on muscle strength and to estimate reproducibility and evaluate various scaling methods in order to optimize comparisons of isokinetic PT measurements.

3. A single isokinetic exercise protocol was used to compare the muscle strength of 31 healthy men of two age groups (< 40 and > 60 years; part A) and change in strength due to administration of dihydrotestosterone (DHT; 70 mg/day) or placebo gel for 3 months in 35 healthy older (≥ 60 years) men (part B).

4. Muscle strength was assessed by a total of 16 PT measurements using a Cybex NORM dynamometer (Cybex, Ronkonkoma, NY, USA). Age-related differences in muscle strength were estimated by using PT evaluated as raw data or scaled by normalizing methods, including simple ratio (PT/weight), allometric PT (PT/weight$^{0.67}$) and adjustment of PT by weight, height, body mass index and body surface area (BSA). The goodness-of-fit for various scaling methods was compared using the Akaike Information Criterion (AIC) as an objective measure of model-based entropy reduction.

5. The effects of DHT administration according to different scaling methods were estimated by eta-squared measure of effect size in treatment models. In part A, older men were weaker than younger men in five knee PT, consistently by all eight analysis models but not in shoulder PT. In part B, DHT treatment resulted in an increase one knee PT (dominant knee flexion at 120°/s) with the difference consistent in all seven models.

6. The scaling model using BSA proved superior to other comparison models throughout both parts of the present study according to entropy minimization criteria (AIC) for goodness-of-fit of the model or eta-squares for treatment effect size.

7. We conclude that differences in muscle strength due to age or androgen administration in older men are restricted to a minority of lower limb contractions and that use of BSA scaling for PT values is considered the best scaling method for muscle strength comparisons in either cross-sectional or longitudinal studies.

Key words: age, dihydrotestosterone androgen, isokinetic dynamometry, muscle strength.

INTRODUCTION

Age is associated with a marked reduction in muscle mass and strength for which therapeutic strategies to enhance quality of life in older men have been proposed to reduce frailty, falls and fractures.[1] In order to evaluate androgen effects on muscle strength, it is essential to have sensitive, specific and reproducible measures to distinguish drug effects reliably from background 'noise'. Muscle strength is a complex measure influenced by factors such as gender, age, motivation, muscular training and body size, which contribute to normal variability and make it difficult to quantify reproducibly and objectively.[2,3] Objective, standardized measurement of muscle strength has most effectively been based on isokinetic dynamometry.[4–6]

An isokinetic dynamometer estimates peak torque (PT) during maximal strength exertion measured at a fixed angular speed with the limb and adjacent joints in a fixed configuration. The precision of a dynamometer in providing reproducible PT estimates depends on the application of resistive forces by the dynamometer while the subject generates force rotating the lever arm to maintain the preset constant angular velocity. Torque is recorded throughout the range of motion, with gravity correction being performed by the accompanying software. Reliability of isokinetic PT measurements has been reported, but the measurements are difficult to compare due to the limited range of test conditions.[7] Most dynamometric investigations have studied knee movements, with very few studying contractions of major upper limb joints. This is problematic because it is anticipated that androgen effects on muscles may be greater on the upper, rather than the lower, limb girdle.[8] Therefore, in order to quantify most sensitively the effects of androgen supplementation on muscular strength in older men in a prospective therapeutic study, it was necessary to refine methods of muscle strength and data analysis in a cross-sectional study.

Correspondence: Professor DJ Handelsman, ANZAC Research Institute, Sydney, NSW 2139, Australia. Email: djh@med.usyd.edu.au

Received 12 September 2000; revision 1 August 2001; accepted 4 September 2001.

When making comparisons of PT, an analysis restricted to raw data may not account fully for differences in body size and musculature; for example, differences in limb length may influence torque through physical effects on the limb acting as a fulcrum. Such intermediary variables may be taken into account by scaling methods for subject differences in body size.[2,3] Such scaling methods help refine analyses to partition out confounding effects of body size and allow more powerful tests of hypotheses being evaluated.

Potential scaling methods for PT include direct ratio, power function ratio and various linear adjustments. In the direct ratio method, the physiological variable of interest is expressed relatively to body mass (e.g. PT/weight). This makes the simplistic assumption that PT and weight vary in a constant linear proportion.[9] Other authors have proposed a curvilinear exponential function.[10,11] According to the method of allometry, the dependent variable should be expressed as a ratio of some power function of weight, such as PT/weight^b, with b being a variable scaling exponent. The most frequently proposed value for the scaling exponent has been 0.67,[12,13] which has been supported by empirical studies,[10,14,15] although 0.75 was also suggested.[16] A disadvantage of such scaling methods is that they ignore height. Height is relevant to muscular strength because the longer a limb, the greater the force moment a fixed muscular force can generate.[17,18] The present study extends the evaluation of scaling models by introducing linear regression models to incorporate height into analysis by using body mass index (BMI), a weight-for-height measure usually used to reflect adiposity or non-lean mass, and body surface area (BSA), a weight × height measure often used to indicate overall body size, as well as covariates to enhance the sensitivity of analyses of muscular strength. Empirical findings from both a cross-sectional study of age effects and a longitudinal study of androgen effects in a randomized double-blind, placebo-controlled clinical trial are used to demonstrate the usefulness of these scaling methods.

METHODS

Designs

The present study was conducted in two parts. Part A was a cross-sectional cohort study comparing the muscular strength of men in two different age groups (young, 20–40 years; older, > 60 years). Part B was a prospective double-blind clinical trial with men > 60 years randomized to daily treatment for 3 months with transdermal dihydrotestosterone (DHT) or placebo gel. None of the men participated in both studies. All participants were requested not to vary their diet and patterns of physical activities during the study. All were given an information sheet and provided a signed consent form prior to entry into the study. The studies were approved by the Central Sydney Area Health Service Ethics Committee (Royal Prince Alfred Hospital – RPAH zone) within the guidelines of the National Health and Medical Research Council of Australia (NH&MRC) for human experimentation.

Subjects

Recruitment of healthy, ambulant men with no history of neurological or joint pathology and unrestricted knee and shoulder movement was undertaken by advertisement. None had prior experience with isokinetic exercise. For part A (cross-sectional study), men were assigned to form two age groups with young subjects being aged 20–40 years and the older subjects being > 60 years of age. For part B (longitudinal study), men > 60 years of age were eligible if they had plasma testosterone levels ≤ 15 nmol/L on two separate occasions but were excluded if they had health conditions that interfered with safe participation, muscle testing, topical drug delivery or used inappropriate medication. The therapeutic intervention in the longitudinal study was a 3 month treatment of 70 mg hydroalcoholic dermal gel containing DHT (Andractim; Laboratoires Besins-Iscovesco, Paris, France) or placebo applied on the skin of their chest daily after a morning shower.

Procedures

In part A, men underwent isokinetic dynamometry on two visits 1 day apart. In part B, men underwent isokinetic dynamometry and body composition measurements before and after 3 months of treatment. Height was measured to the nearest 0.5 cm using a standard stadiometer and weight was measured to the nearest 0.1 kg with subjects lightly dressed.

Body composition

Lean mass was estimated from bio-impedance readings according to Lukaski's formula for men.[19] Using bioelectrical impedance analysis (BIA) readings (all in Ω) for resistance (R), reactance (X_c) and impedance (Z), with height (H; cm) and weight (W; kg) and S = 1 for male gender:[20]

$$\text{Fat-free mass (FFM)} = 0.734(H^2/R) + 0.116(W) + 0.096\ X_c + 0.878(S) - 4.03$$

Bioelectrical impedance was measured with a SEAC Model BIM 3.0 Bio-Impedance meter (Inderlec Australia, Baulkham Hills, NSW, Australia). Lean mass was also calculated from bodyweight by subtraction of fat mass that was estimated from Siri's equations[21] using body density[22] and the sum of four skinfold thickness (SUM = biceps + triceps + subscapular + suprailiac skinfold thickness) according to the formulas:

$$\text{Body density} = 1.1715 - 0.0779 \times \log \text{SUM}$$
$$\%\ \text{Body fat} = (4.95/\text{body density} - 4.50) \times 100$$

Skinfold thicknesses were assessed at biceps, triceps, subscapular and suprailiac positions at standard sites on the right side of the body[23] with a Harpenden Skinfold Caliper (British Indicators, Bedfordshire, UK).

Muscle testing

The maximal muscle strength during a contraction was measured as PT using a Cybex® NORM™ dynamometer (Cybex, Division of Lumex, Ronkonkoma, NY, USA) interfaced with an IBM-compatible computer for data reduction and analysis. The dynamometer measures torque as the product of the force (monitored on a transducer orientated perpendicularly to the limb segment) and the distance between the cuff attachment on the limb (in case of knee movements) or the hand-held handle (shoulder movements) and the joint centre. Measurements were made during an isokinetic contraction at a constant preset angular velocity throughout the full range of motion. The PT was defined as the highest isokinetic torque produced during five repetitions of the contraction. There were altogether 16 contractions assessed, specified by limb (knee and shoulder), side of the limb (dominant and non-dominant), movement (extension and flexion) and angular speed (90 and 120°/s). These angular speeds were selected based on our observation that a lower speed (e.g. 60°/s) was too labourious for older men to perform and speeds higher than 120°/s were more effort dependent and had lower reproducibility.[24] The dynamometer was weight and speed calibrated using its internal system before each test session and a position calibration was performed before the testing of each subject. Before each test, men were positioned as required with individual adjustment for maximal comfort while ensuring alignment of the dynamometer axis to the rotation axis of the tested joint. The setup data were recorded and reproduced in following sessions by built-in software of the dynamometer.

Knee

Subjects were seated with their back against the chair at a reclining angle of 85°. The body was stabilized by a shoulder belt, by both hands grasping the handlebars at the sides of the chair firmly and the contralateral limb by a stabilizer. The limb of the knee being tested was strapped at the proximal

thigh, approximately 2 cm proximal to malleolus so that full ankle dorsiflexion was still possible. The position set-up was completed with the alignment of the axes of rotation of the knee (by a line passing through the femoral epicondyles) and of the dynamometer.

Shoulder

The chair was set on a horizontal position. The body was stabilized by a belt around the torso, the feet were kept in a non-weight-bearing position by being placed in a footrest and the hand of the non-testing shoulder firmly held the handle bar. The hand of the testing shoulder held the handgrip of the shoulder lever arm. During testing, the elbow was kept straight and the subject was reminded not to move his shoulder upwards.

It was explained to particpants that the purposes of the set-up were to position the alignment of the axes of the testing joint and the dynamometer and to maintain the alignment during tests, so as to provide conditions for the most reliable performance of muscle strength testing.

Standard test protocol

The dominant sides were defined by which limbs were used to throw or kick a ball. All details of the set-up for each individual, including range of motion, dynamometer and chair, were recorded on the interfaced computer and reproduced for that person on further tests. All joints tested were evaluated for their isokinetic strength through their full range of motion. Gravity correction was calculated by the Cybex internal computer for the effect of limb weight on torque production after details were obtained during the set-up. No overshoot was observed in graphics of any tests.

Familiarization

A familiarization exercise was required before the isokinetic test on the first visit. This aimed to demonstrate the nature of the isokinetic test to ensure the subject could perform the exercise comfortably. A set of five repetitions, continuous concentric–concentric extension–flexion cycles, at 90°/s was used. Test subjects were encouraged to push the limb as fast and as hard as they could until the limb reached the end of its range of motion, then to pull it back in the same manner to the other end. Throughout the exercise, the tester provided regular verbal encouragement, aiming to obtain maximal effort by the subject. The familiarization was considered complete when subjects felt confident in performing maximal isokinetic contractions and this was confirmed by graphs of consistent performance on the computer screen.

Tests

A brief description of an isokinetic exercise assessment is provided in the Table 1.

Joint order was obtained in the first visit by: (i) random selection of joint (knee or shoulder), then of side (dominant or non-dominant); and (ii) fixed order of velocities: 90 then 120°/s. Dynamometry was conducted at the same time of day in subsequent visits. Subjects were encouraged to perform the exercise with maximum force before commencing a test session and the verbal feedback to perform at maximal level during the test was kept at the same level for all subjects. A single operator (LPL) conducted all the isokinetic tests. The practice repetitions, five repetitions at each speed and increasing order of speeds, were designed to ensure high reliability.[6,25]

Data analysis

Data

In part A, comparison of muscle strength of the two age groups was analysed as mean raw PT for each of the 16 contractions performed on the first days. In part B, changes in PT (= after – before treatment) were used as the response variable due to the DHT or placebo treatment.

Scaling

Data were analysed in raw values or after scaling. Scaling methods were simple ratio (= PT/weight), allometry (= PT/weight$^{2/3}$) and adjustment by body size indices (part A) or change in body composition (part B).

Statistical tests

Tests used for part A were: (i) *t*-test for raw, simple ratio, allometric PT; and (ii) ANCOVA for raw PT adjusted by BSA, BMI, weight, height and weight and height (eight models). To estimate the changes in muscle strength due to treatment, *t*-tests were used for changes in raw and allometric PT and ANCOVA for changes in raw PT adjusted by changes in BSA, in directly measured lean mass, in FFM estimated from skinfold, in BMI and in weight (seven models).

Relevant formulae are:

$$\text{Relative PT (Nm/kg)} = \text{PT/W, with weight, W, in kg}$$

$$\text{Allometric scaled PT (Nm/kg}^{0.67}\text{)} = \text{raw PT/W}^{0.67}$$

$$\ln \text{BSA (m}^2\text{)} = -3.751 + 0.422 \ln H + 0.515 \ln W,^{26} \text{ with height, H, in cm and weight, W, in kg}$$

$$\text{BMI} = W/H^2, \text{ with height, H, in m and weight, W, in kg}$$

Selection of the best fit model (part A)

Estimators for PT were compared between the various models according to the quality of model fit adjusted for differing numbers of variables according to the Akaike Information Criterion (AIC).[27] The AIC is a general measure of entropy reduction due to the model that allows comparisons between models using different numbers of variables. The AIC is calculated as:

$$\text{AIC} = \chi^2 - 2\text{d.f.}$$

where χ^2 and d.f. were calculated from ANCOVA or cross-tabulation depending on the tests. The AIC values closer to zero indicate better fit and greater parsimony (i.e. more efficient reduction in entropy).

Evaluation of muscle strength effect size (part B)

The proportion of the variance in the response variable that can be attributed to the independent variable was expressed using the eta-squared test as an unbiased estimator of effect size in order to compare various models.[25,28] In this study, eta-squares estimate the association between the effects of DHT on PT and the two experimental groups.

Reproducibility of measurements

Inter-day reliability of PT was estimated by calculation of intraclass correlation coefficients (ICC) and reproducibility by coefficients of variation (CV; expressed as a percentage) for each contraction. The ICC were computed by the ANOVA method in accordance with methods described by Shrout and Fleiss.[29] The ICC(2,1) used in the present study estimates the inter-rater reliability in measuring results of two sets of tests in which raters are assumed to be randomly sampled; hence,

Table 1 Protocol of isokinetic tests

	Joint test	Peak torques collected
Warm-up	Five repetitions of extension–flexion (three sub-maximum and two maximum), then rest 60 s	0
Test at 90°/s	Five repetitions of extension–flexion, then rest 30 s	One extension + one flexion
Test at 120 °/s	Five repetitions of extension–flexion	One extension + one flexion
Total		4

Four joints were tested, and the joint order decided on day 1 was followed afterwards. An average of 12 min was required to set up the dynamometer for the next joint test; there was another 5 min warm-up before the dynamometry test not included in this table.

for the present study, this ICC should be interpreted as a measure of reliability, which can be generalized for similar raters. Values of the ICC test statistic range between 0 and 1 and values closer to 1 indicate greater variability.

The related formula was:

$$ICC(2,1) = (BMS - EMS)/(BMS + (k-1)EMS + k(JMS - EMS)/n)$$

where BMS is the mean squares for subjects, EMS is the subject by day error mean squares, JMS is the between-days mean squares, k is the number of occasions (k = 2) and n is the number of subjects (n = 16 young and n = 15 older men).

The CV is defined as:

$$CV\ (\%) = 100 \times SD/mean$$

where SD (standard deviation) is the square root of the EMS above and the mean was the grand mean in the two-way ANOVA for computing ICC.

A significant difference was defined if $\alpha = 0.05$. Data are expressed as the mean±SEM.

RESULTS

Subjects

In part A, older men were similar to their young counterparts in terms of height, weight and BSA, but had higher BMI (Table 2). In part B, both groups of older men were similar, except for weight and BMI, despite the randomization. The older men of both studies (Table 2) were comparable on all characteristics.

Comparison of PT

Baseline muscle strength (raw PT) for both studies is given in Table 3.

In part A, among 16 PT of knee and shoulder, strength about the knee joint was significantly less for older men compared with young men for all four extension and one flexion (non-dominant; 90°/s)

Table 2 Subject characteristics

	Part A			Part B	
Characteristics	DHT (n = 17)	Young (n = 16)	Older (n = 15)	All (n = 35)	Placebo (n = 18)
Age (years)	28.5 ± 1.6	69.9 ± 1.6	69.3 ± 1.1	68.0 ± 1.5	70.6 ± 1.5
Height (cm)	175.2 ± 1.5	173.6 ± 1.6	172.7 ± 1.0	172.2 ± 1.5	173.1 ± 1.5
Weight (kg)	71.6 ± 3.1	80.4 ± 3.2	82.1 ± 2.0	85.8 ± 2.9	78.3 ± 3.0*
BSA (m^2)	1.87 ± 0.04	1.98 ± 0.05	1.99 ± 0.03	2.04 ± 0.04	1.95 ± 0.04
BMI (kg/m^2)	23.2 ± 0.8	26.6 ± 0.8	27.5 ± 0.6	28.9 ± 0.8	26.0 ± 0.8*
Limb dominance					
Right hand	16	13	32	17	15
Right leg	16	14	33	18	15

Data expressed as the mean ± SEM. $P < 0.05$ for comparison between young and older men in part A; *$P < 0.05$ for comparison between treatment groups dihydrotestosterone (DHT) and placebo in part B. There was no difference in limb dominance among the groups by Chi-squared tests ($\alpha = 0.05$).

BSA, body surface area; BMI, body mass index.

Table 3 Baselines of peak torques (in N·m) and changes in peak torques due to dihydrotestosterone treatment

	Part A		Part B		
				Difference from baseline at the end of 3 month treatment	
Isokinetic peak torque	Young men (n = 16)	Older men (n = 15)	Baseline (n = 35)	DHT (n = 17)	Placebo (n = 18)
Knee extension					
Dominant, 90°/s	141.6 ± 5.9	113.2 ± 6.1†	109.0 ± 4.0	0.1 ± 2.8	0.8 ± 2.9
Dominant, 120°/s	127.2 ± 5.1	96.0 ± 5.2†	93.8 ± 3.4	3.6 ± 2.5	1.2 ± 2.6
Non-dominant, 90°/s	141.8 ± 6.8	113.3 ± 7.0†	106.9 ± 4.6	2.4 ± 2.8	3.1 ± 1.8
Non-dominant, 120°/s	127.9 ± 6.0	97.5 ± 6.2†	91.9 ± 4.1	4.3 ± 1.7	1.6 ± 2.7
Knee flexion					
Dominant, 90°/s	92.0 ± 4.3	82.5 ± 4.6	82.4 ± 2.9	6.7 ± 2.6	3.5 ± 1.7
Dominant, 120°/s	79.6 ± 4.1	74.9 ± 4.2	74.3 ± 2.8	6.0 ± 2.2§	0.2 ± 1.5
Non-dominant, 90°/s	90.2 ± 4.0	73.5 ± 4.2†	81.6 ± 2.7	4.3 ± 2.2	5.1 ± 1.6
Non-dominant, 120°/s	78.3 ± 4.0	68.9 ± 4.1	73.8 ± 2.7	2.7 ± 1.8	1.2 ± 1.7
Shoulder extension					
Dominant, 90°/s	65.1 ± 3.5	69.1 ± 3.6	78.2 ± 2.4*	1.3 ± 1.5	–0.2 ± 1.5
Dominant, 120°/s	59.6 ± 3.4	62.7 ± 3.5	69.3 ± 2.3	0.1 ± 2.1	–1.9 ± 1.8
Non-dominant, 90°/s	64.7 ± 3.8	68.5 ± 3.9	74.8 ± 2.6	4.7 ± 2.2	2.3 ± 1.7
Non-dominant, 120°/s	57.1 ± 3.8	61.1 ± 3.9	65.5 ± 2.6	2.1 ± 2.4	0.9 ± 1.1
Shoulder flexion					
Dominant, 90°/s	41.0 ± 2.7	39.3 ± 2.8	40.0 ± 1.8	1.8 ± 1.6	2.3 ± 1.3
Dominant, 120°/s	36.9 ± 2.8	33.9 ± 2.9	34.5 ± 1.9	1.2 ± 1.7	0.7 ± 2.0
Non-dominant, 90°/s	40.1 ± 2.4	37.3 ± 2.5	39.2 ± 1.7	2.5 ± 1.5	1.4 ± 1.8
Non-dominant, 120°/s	34.4 ± 2.4	32.7 ± 2.5	33.3 ± 1.6	0.1 ± 1.5	1.1 ± 1.7

Data are expressed as the mean ± SEM. Baseline data from part B have been grouped together. †$P < 0.05$ for comparisons between young and older men in part A; *$P < 0.05$ for comparisons between 15 older men of part A and 35 men of part B (*); §$P < 0.05$ for comparisons between treatment groups in part B.

DHT, dihydrotestosterone.

PT; for the remainder, there were no significant differences according to age.

In part B, for 15 of 16 PT, changes in raw PT were not different between the two groups of ageing men due to 3 month DHT treatment. An increase in muscle strength was found only in the PT of the isokinetic contraction of dominant knee flexion at 120°/s.

The results and patterns of statistical significance using raw PT and seven other PT-based estimators of muscle strength for knee flexion and extension using the data from the two age groups in part A are depicted in Fig. 1. Due to differences in scale, neither direct ratio nor allometric scaled data could be depicted on the upper part of the figure, although their statistical pattern of significance is included.

Selection of the best-fitted model

For all eight PT estimators, goodness-of-fit (as measured by AIC) and effect size (percentage difference) for the comparison between age groups in the cross-sectional study are given in Table 4 using the five PT that demonstrated significant between-group differences. All five PT show that BSA-associated PT provided the best model for the data. These findings were corroborated by an ANCOVA regression model in which the highest adjusted R^2 and lowest residual mean square were found for BSA as a covariate compared with models using BMI, weight, height as well as weight and height together as covariates (data not shown).

Reproducibility coefficients of PT

Older men had higher reproducibility than did younger men according to consistently higher intraclass correlation coefficients and lower CV (Table 5). Using the criterion of ICC ≥ 0.90 and CV < 8.0% as indicating high reliability and low variability, respectively, the older men had 10 and 11 contractions demonstrating high reproducibility compared with three and six for their young counterparts.

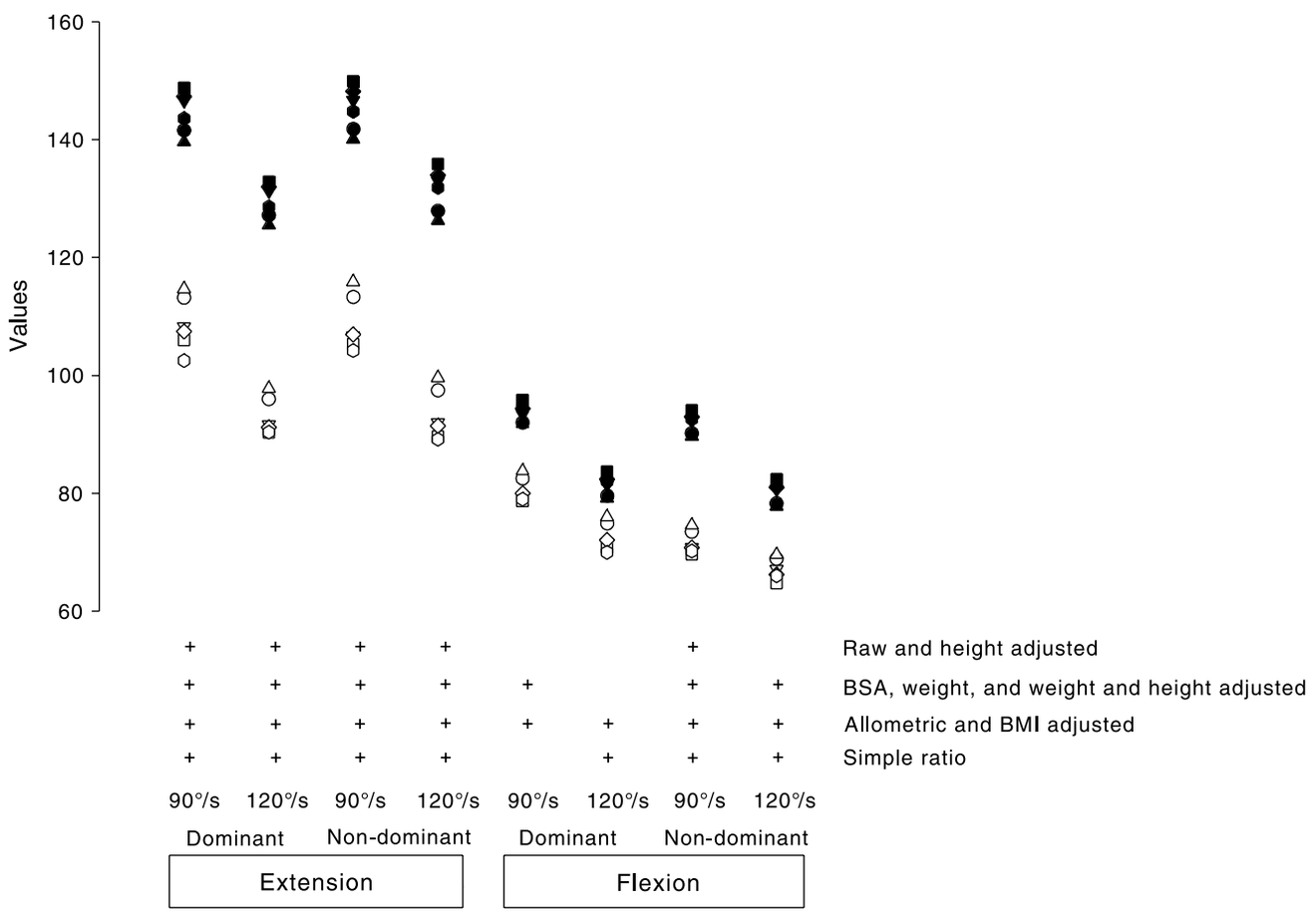

Fig. 1 Comparison of raw peak torque (PT) and seven other PT-based estimators of strength of knee movements. The upper part of the figure shows plots of raw or adjusted PT values (*y* axis) of the young (filled symbols) and the older (open symbols) over eight knee movements (*x* axis). (●), raw PT, young; (○), raw PT, older; (▼),body surface area (BSA)-adjusted PT, young; (▽), BSA-adjusted PT, older; (■), body mass index (BMI)-adjusted, young; (□), BMI-adjusted PT, older; (◆), weight-adjusted PT, young; (◇), weight-adjusted PT, older; (▲), height-adjusted PT, young; (△), height-adjusted PT, older; (⬣), weight- and height-adjusted PT, young; (⬡), weight- and height-adjusted PT, older. The lower part of the figure shows the outline of the knee PT in which older men are weaker: The five PT that have four plus (+) symbols in the column are consistently different between age groups through all eight models. Older men have most knee PT weaker than the young men. Note that: (i) the plus symbol (+) in the lower part of the figure indicates that the difference in PT of a particular isokinetic contraction (specified by test conditions in the bottom lines) between older and young men is significant ($P < 0.05$) in the analysis of a particular model (on the right side of the figure); and (ii) figures for shoulder movements are not displayed because of no group difference.

Table 4 Measures of goodness-of-fit by Akaike Information Criterion and per cent difference between young and older men in eight peak torque-based estimators of the five consistently different peak torques (part A)

Peak torques with consistent difference	Raw (Nm)	Simple ratio (Nm/kg)	Allometric (Nm/kg$^{0.67}$)	BSA adjusted (Nm)	BMI adjusted (Nm)	Weight adjusted (Nm)	Height adjusted (Nm)	Weight and height adjusted (Nm)
AIC values								
Knee extension								
Dominant, 90°/s	–27.01	–29	–29	–10.41	–18.18	–11.84	–21.53	–20.42
Dominant, 120°/s	–23.00	–27	–29	–12.34	–21.43	–13.98	–18.84	–24.78
Non-dominant, 90°/s	–29.01	–29	–29	–9.05	–17.51	–10.10	–19.94	–22.62
Non-dominant, 120°/s	–20.01	–29	–29	–8.23	–16.35	–9.04	–21.35	–21.50
Knee flexion								
Non-dominant, 90°/s	–21.67	–29	–29	–20.65	–22.16	–20.72	–26.00	–26.26
Per cent difference								
Knee extension								
Dominant, 90°/s	25	42	36	35	40	37	22	40
Dominant, 120°/s	32	50	44	43	47	45	28	42
Non-dominant, 90°/s	25	43	36	38	42	39	21	39
Non-dominant, 120°/s	31	49	42	45	52	46	27	48
Knee flexion								
Non-dominant, 90°/s	23	40	34	30	35	31	20	32

This table is based on the peak torques (PT) of five isokinetic contractions that consistently differ between groups through eight models.
An Akaike Information Criterion (AIC) value closer to zero indicates a better goodness-of-fit of the model.
The per cent difference = (older PT – young PT)/older PT.

Table 5 Intraclass correlation coefficients and coefficients of variation for 1 day inter-trial reproducibility of Cybex dynamometer (part A)

Isokinetic peak torque	ICC(2,1) Young	ICC(2,1) Older	CV% Young	CV% Older
Knee extension				
Dominant, 90°/s	0.83	0.95	9.63	3.86
Dominant, 120°/s	0.90	0.88	6.82	7.65
Non-dominant, 90°/s	0.89	0.94	7.58	5.29
Non-dominant, 120°/s	0.87	0.92	9.05	7.20
Knee flexion				
Dominant, 90°/s	0.64	0.82	12.99	8.61
Dominant, 120°/s	0.79	0.91	9.82	7.03
Non-dominant, 90°/s	0.87	0.84	6.27	8.19
Non-dominant, 120°/s	0.78	0.91	11.03	6.57
Shoulder extension				
Dominant, 90°/s	0.93	0.91	5.91	6.11
Dominant, 120°/s	0.90	0.87	7.88	8.56
Non-dominant, 90°/s	0.80	0.93	10.72	6.45
Non-dominant, 120°/s	0.81	0.93	11.93	6.40
Shoulder flexion				
Dominant, 90°/s	0.80	0.83	7.79	11.12
Dominant, 120°/s	0.77	0.83	11.36	14.28
Non-dominant, 90°/s	0.62	0.92	12.40	6.60
Non-dominant, 120°/s	0.84	0.95	9.76	6.85
Descriptive statistics				
Mean	0.82	0.90	9.43	7.55
SD	0.09	0.05	2.22	2.40
SEM	0.02	0.01	0.60	0.55
Minimum	0.62	0.82	5.91	3.86
Maximum	0.93	0.95	12.99	14.28

Reproducibility is higher if the intraclass correlation coefficient (ICC) is closer to 1 or the coefficient of variation (CV%) is closer to 0.

Using the reproducibility parameters, power and sample size estimates for raw and BSA-adjusted PT (angular speed 90°/s) could be established (Table 6) to detect changes of 10, 20, 30 and 40% in PT ($\alpha = 0.05$ and $\beta = 0.20$) for eight extensions and flexions of the knee and shoulder of either sides at a speed of 90°/s. Sample sizes are calculated for fixed type I and II errors and the smallest number of subjects required for knee flexor PT is predicted in Table 6. This prediction is vindicated in the longitudinal study (part B).

Table 6 Sample size to detect a change in strength (part B)

Peak torque at 90°/s (used value = BSA-adjusted PT of aged group)	Sample size for an increase in value (two–tailed $\alpha = 0.05$ and power = 0.80) 10%	20%	30%	40%
Knee extension				
Dominant	82	21	10	6
Non-dominant	130	33	15	9
Knee flexion				
Dominant	63	16	7	4
Non-dominant	64	16	8	4
Shoulder extension				
Dominant	70	18	8	5
Non-dominant	106	27	12	7
Shoulder flexion				
Dominant	120	30	14	8
Non-dominant	91	23	11	6

The numbers in the table indicate the sample size necessary to detect a change in body surface area (BSA)-adjusted peak torque (PT) of 10, 20 or 30% with $\alpha = 0.05$ and $\beta = 0.20$.

Note that knee flexion appears the best variable for longitudinal studies of muscle strength due to its minimal requirement for sample size to detect changes.

Evaluation of scaling methods for PT estimators of drug effects in part B

Seven PT estimators of treatment effects (differences between 3 months of DHT or placebo treatment) were compared. In the intervention therapy part, among the 16 contractions, DHT treatment

Table 7 Effect of sizes of groups in seven models of comparison of the consistently different dominant knee flexor at 120°/s (part B)

Response variable (test)	Covariate	Group difference Mean (±SEM) difference value	*F* ratio (*P*)†	Eta-squared*
Raw PT, Nm, (*t*-test)		5.8 ± 2.7	4.78 (.04)	0.13
Allometric PT, Nm/kg$^{2/3}$, (*t*-test)		0.34 ± 0.14	6.20 (.02)	0.16
Raw PT, Nm, (ANCOVA)	Changes in BSA	8.3 ± 3.0	4.00 (.05)	0.20
Raw PT, Nm, (ANCOVA)	Changes in lean mass (BIA)	5.9 ± 2.7	5.79 (.02)	0.13
Raw PT, Nm, (ANCOVA)	Changes in FFM (Siri)	3.2 ± 2.8	7.50 (.01)	0.04
Raw PT, Nm, (ANCOVA)	Changes in BMI	7.4 ± 2.6	7.18 (.01)	0.19
Raw PT, Nm, (ANCOVA)	Changes in weight	7.4 ± 2.6	7.00 (.01)	0.18

Response variable: difference in group changes due to treatment (changes in dihydrotestosterone group – changes in placebo group) calculated from the peak torque (PT) of dominant knee flexion at 120°/s.

*A higher eta-square indicates a better effect size due to groups (i.e. a stronger relationship between groups and response variable).

†*F* ratio = Between-groups mean square/within-groups mean square.

$P < 0.05$ indicates that the difference is significant.

BSA, body surface area; BIA, bioelectrical impedance analysis; FFM, fat-free mass; BMI, body mass index.

produced differences only in strength of dominant knee flexion at 120°/s (Table 7). The ANCOVA model using BSA as a covariate provided highest treatment effect size, as indicated by eta-squared, closely followed by the PT adjusted for BMI or weight as covariates.

DISCUSSION

The present study evaluated a wide range of contractions, their reproducibility and optimal PT-based estimators to obtain a comprehensive view of the most sensitive measure of muscular strength by isokinetic dynamometry in older men suitable for therapeutic intervention studies. Previous studies have examined fewer (less than six) contractions.[4,5,9,24,30–33] Among the 16 joint contractions studied, knee flexion proved the most sensitive to both age and androgen effects, whereas BSA-adjusted PT proved the most consistent PT-based estimator to model the experimental data for both age- and treatment-related changes. Our observations that knee flexion strength was the most consistently sensitive to the effects of age and of androgen effects indicates that it is the most useful variable to detect therapeutic interventional effects in future studies of muscular strength if any less comprehensive set of joint contractions is to be used.

The generally high reproducibility of isokinetic dynamometry, particularly among older men, supports the validity of this testing regimen for therapeutic studies of the effects of drugs and/or exercise in older men. Whether the superior reproducibility among older men is due to more consistent motivation remains unclear, but it is advantageous for muscle strength studies in this age group. Using the demanding criteria of an ICC ≥ 0.90 or a CV ≤ 8.0% for high reproducibility for a biological measurement,[34] the Cybex NORM dynamometer provides reliable measurement of muscle strength. The day-to-day, within-subject reproducibility estimates we have measured allow development of power and sample size estimates for future studies. These shed light on previous studies in explaining why ones with small sample size fail to show beneficial effects when they may have been missed,[35] as well as estimating the effect size when larger studies are negative.[36]

The present study finds that healthy ambulant older men have reduced strength in knee contraction, notably in flexion, but that shoulder strength is little affected by age or androgen treatment. Our findings of 25–32% decreases with age in knee strength are consistent with previous studies reporting 22–32% reductions in older compared with younger men.[37–39] Variations in knee strength with age and androgen therapy may reflect changes in muscle bulk with age[32,40] or androgens,[36,41,42] with or without additional changes in intrinsic muscular function. The reduced muscular strength in older men may be due to age-related changes in factors either intrinsic to neuromuscular units, such as changes in muscle structure and innervation, and/or extrinsic factors, such as motivation, habitual physical activity and circulating hormones, including androgens. Age-related changes in neuromuscular structure include reduced muscle protein synthesis[43] and muscle mass,[40,44] as well as changes in muscle fibre composition[38,45] and denervation.[39,44,46]

Among extrinsic factors that influence muscle structure and function, the age-associated decline in circulating hormones, notably testosterone concentration, may be important, as suggested by correlations in observational studies.[40,47] Previous studies of androgen administration to older men are limited and provide conflicting results with regard to improvement in muscle strength. The long controversy over whether androgen administration improves muscular strength in eugonadal men[48] has been resolved by a definitive study showing that supraphysiological testosterone doses increase muscle mass and strength in healthy, eugonadal young men.[49] Notably, in that study, muscle strength increased in squat lifting but not bench-press lifting,[49] consistent with our findings of greater improvement in muscles of lower compared with upper limb strength; similar findings have been reported following sublingual testosterone administration in hypogonadal men.[50] In older men, controlled studies of androgen administration studies have shown consistent increases in lean (i.e. muscle) mass,[36,51,52] whereas muscle strength is reported as unchanged[36,52] or increased.[51] Two of these three studies have used isokinetic dynamometry.[36,51] The limited increase in muscle strength due to DHT treatment in the longitudinal part of the present study (part B) may be due to the fact that the participants were relatively healthy, rather than debilitated, elderly men with little scope for improvement or increase in their routine daily activities during treatment. The minimal improvement in muscle strength despite increased muscle mass suggests that changes in muscle function and efficiency with age and androgens need to be further evaluated.

The present study is the first systematic study of shoulder strength changes with age and androgens. Our finding of no difference between age groups or androgen treatment suggests that shoulder strength, as measured by isokinetic dynamometry, is not a sensitive measure of age or androgen effects on muscle. It is possible that this reflects the long-term changes in elective exercise load for a non-weight-bearing limb compared with the little-varying habitual weight-bearing load on lower limbs in ambulant men. This inability to find changes in shoulder strength may indicate that such changes are too small (or absent) to be detected in studies of this size. If so, any such changes are likely to have minimal clinical or therapeutic significance. Whether there are age-related changes in muscle composition, such as fibre proportions or functions in shoulder girdle muscles, is unclear. Alternatively, older men may compensate for reductions in muscle power by complex changes in muscle activation patterns. Gross changes are unlikely due to the extensive standardization of limb position and movement in the dynamometry testing; however, more subtle changes in sequential activation would need to be evaluated by electromyography.

Important caveats for studies of muscular strength include the methodological issues of fatigue and effort dependence of strength measures. There was no report or observation of fatigue or soreness for any participant in the present studies. This is consistent with the modest physical demands of the testing, which involved less than 45 min, during which time 12 short (< 15 s) exercise sets were separated by rest periods (30–60 s between speeds, 10 min between joints). To overcome the effort dependence, participants were regularly encouraged and appeared to provide a maximal effort during exercise. The standardized patterns of testing were also designed to minimize systematic bias according to scheduling. Nevertheless, isokinetic dynamometry is an effort-dependent test and variations in motivation and effort may introduce variability, so that larger sample size is necessary to overcome methodological variance and evaluate therapeutic effects.

Scaling of physiological measurement for individuals of different body size has mostly applied to studies of maximal oxygen uptakes,[10,11,14,15,53,54] which reflect muscular aerobic function, whereas only a few have been applied to muscle strength measurements.[3,30] The study of Davies *et al*.[30] examined three methods of scaling (ratio standard, allometry and ANCOVA) for knee contraction in older men and women and found differences between whole body mass and lean mass adjustments. To reconcile these observations, it has been suggested that scaling according to body mass for weight-bearing and lean mass for non-weight-bearing exercise[10,30] is desirable, whereas for variables that vary with body size, allometric scaling is preferable.[10] The present study contributes the suggestion that BSA adjustment of raw PT be used as an estimator of muscle strength, which is consistent with a previous finding of maximal oxygen consumption during treadmill testing.[54] This seems rational because body height has an important influence on raw muscular strength, including its influence on long bone length.[17,18] We found BSA-adjusted PT to be the most consistent estimator of muscle strength for age and androgen effects.

In conclusion, the present study demonstrates the reproducibility of isokinetic dynamometry for testing muscle strength in older men and its suitability for evaluating therapeutic effects. The most satisfactory estimator of muscle strength was BSA-adjusted PT, which provided the best detection of age and androgen effects in older men. These approaches provide the basis for further analyses of age-related changes in muscle function, as well as evaluation of therapeutic studies to enhance muscle strength in ageing men.

ACKNOWLEDGEMENTS

The authors thank the staff of the Department of Andrology, Concord Hospital, for their valued and skilful assistance in conducting this study, as well as the participants for their time, exertion and co-operation. This study was supported by the NH&MRC.

REFERENCES

1. Tenover JS. Androgen replacement therapy to reverse and/or prevent age-associated sarcopenia in men. *Baillieres Clin. Endocrinol. Metab.* 1998; **12**: 419–25.
2. Hortobagyi T, Katch FI, Katch VL, LaChance PF, Behnke AR. Relationships of body size, segmental dimensions, and ponderal equivalents to muscular strength in high-strength and low-strength subjects. *Int. J. Sports Med.* 1990; **11**: 349–56.
3. Weir JP, Housh TJ, Johnson GO, Housh DJ, Ebersole KT. Allometric scaling of isokinetic peak torque: The Nebraska Wrestling Study. *Eur. J. Appl. Physiol. Occup. Physiol.* 1999; **80**: 240–8.
4. Murray MP, Gardner GM, Mollinger LA, Sepic SB. Strength of isometric and isokinetic contractions: Knee muscles of men aged 20–86. *Phys. Ther.* 1980; **60**: 412–19.
5. Kramer JF. Reliability of knee extensor and flexor torques during continuous concentric–eccentric cycles. *Arch. Phys. Med. Rehab.* 1990; **71**: 460–4.
6. Wilhite M, Cohen E, Wilhite S. Reliability of concentric and eccentric measurements of quadriceps performance using the Kin–Com dynamometer. *J. Orthop. Sports Phys. Ther.* 1992; **15**: 175–82.
7. Keating JL, Matyas TA. The influence of subject and test design on dynamometric measurements of extremity muscles. *Phys. Ther.* 1996; **76**: 866–89.
8. Gooren LJG, Polderman KH. Safety aspects of androgen therapy. In: Nieschlag E, Behre HM (eds). *Testosterone: Action Deficiency Substitution.* Springer-Verlag, Berlin. 1990; 182–203.
9. Aniansson A, Grimby G, Rundgren A. Isometric and isokinetic quadriceps muscle strength in 70-year-old men and women. *Scand. J. Rehab. Med.* 1980; **12**: 161–8.
10. Nevill AM, Ramsbottom R, Williams C. Scaling physiological measurements for individuals of different body size. *Eur. J. Appl. Physiol. Occup. Physiol.* 1992; **65**: 110–17.
11. Welsman JR, Armstron N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO_2 for differences in body size. *Med. Sci. Sports Exer.* 1996; **28**: 259–65.
12. Sholl D. The quantitative investigation on the vertebrate brain and the applicability of allometric formulae to its study. *Proc. R. Soc. B* 1948; **135**: 243–58.
13. Astrand P, Rodahl K. *Textbook of Work Physiology*, 3rd edn. McGraw-Hill, New York. 1986.
14. Nevill AM, Holder RL. Scaling, normalizing, and per ratio standards: An allometric modeling approach. *J. Appl. Physiol.* 1995; **79**: 1027–31.
15. Heil DP. Body mass scaling of peak oxygen uptake in 20- to 79-yr-old adults. *Med. Sci. Sports Exer.* 1997; **29**: 1602–8.
16. Kleiber M. Body size and metabolic rate. *Physiol. Rev.* 1947; **27**: 511–41.
17. Madsen OR, Lauridsen UB, Hartkopp A, Sorensen OH. Muscle strength and soft tissue composition as measured by dual energy X-ray absorptiometry in women aged 18–87 years (published erratum appears in *Eur. J. Appl. Physiol. Occup. Physiol.* 1997; **76**: 195). *Eur. J. Appl. Physiol. Occup. Physiol.* 1997; **75**: 239–45.
18. Gross MT, McGrain P, Demillo N, Plyler L. Relationship between multiple predictor variables and normal knee torque production. *Phys. Ther.* 1989; **69**: 54–62.

19. Lukaski HC, Bolonchuk WW. Theory and validation of the tetrapolar bioelectrical impedance method to assess human body composition. In: Ellis KJ, Yasumura S, Morgan WD (eds). *Proceedings of the International Symposium on In Vivo Body Composition Studies.* Institute of Medical Physics, London. 1987; 410–14.
20. Lukaski HC. Methods for the assessment of human body composition: Traditional and new. *Am. J. Clin. Nutr.* 1987; **46**: 537–56.
21. Siri W. *Body Composition from Fluid Spaces and Density Analysis of Methods*. National Academy of Sciences, Washington, DC. 1961.
22. Durnin JV, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: Measurements on 481 men and women aged from 16 to 72 years. *Br. J. Nutr.* 1974; **32**: 77–97.
23. Lohman TG, Roche AF, Martorell R. *Anthropometric Standardization Reference Manual.* Human Kinetics Books, Champaign, IL. 1988.
24. Montgomery L, Douglass L, Deuster P. Reliability of an isokinetic test of muscle strength and endurance. *J. Orthop. Sports Phys. Ther.* 1989; **10**: 315–22.
25. Wessel J, Gray G, Luongo F, Isherwood L. Reliability of work measurements recorded during concentric and eccentric contractions of the knee extensors in healthy subjects. *Physiother. Can.* 1989; **41**: 250–3.
26. Bailey BJ, Briars GL. Estimating the surface area of the human body. *Stat. Med.* 1996; **15**: 1325–32.
27. Akaike H. Factor analysis and AIC. *Psychometrika* 1987; **52**: 317–32.
28. SPSS Inc. GLM general factor analysis. In: *SPSS Advanced Statistics 7.5*. SPSS Inc., Chicago, IL. 1997; 1–12.
29. Shrout P, Fleiss J. Intraclass correlation: Uses in assessing rater reliability. *Psychol. Bull.* 1979; **86**: 420–8.
30. Davies MJ, Dalsky GP. Normalizing strength for body size differences in older adults. *Med. Sci. Sports Exer.* 1997; **29**: 713–17.
31. Tredinnick T, Duncan P. Reliability of measurements of concentric and eccentric isokinetic loading. *Phys. Ther.* 1988; **68**: 656–9.
32. Frontera WR, Hughes VA, Lutz KJ, Evans WJ. A cross-sectional study of muscle strength and mass in 45- to 78-yr-old men and women. *J. Appl. Physiol.* 1991; **71**: 644–50.
33. Kramer JF, Vaz MD, Hakansson D. Effect of activation force on knee extensor torques. *Med. Sci. Sports Exer.* 1991; **23**: 231–7.
34. Birmingham TB, Kramer JF, Speechley M, Chesworth BM, MacDermid J. Measurement variability and sincerity of effort: Clinical utility of isokinetic strength coefficient of variation scores. *Ergonomics* 1998; **41**: 853–63.
35. Clague JE, Wu FC, Horan MA. Difficulties in measuring the effect of testosterone replacement therapy on muscle function in older men. *Int. J. Androl.* 1999; **22**: 261–5.
36. Snyder PJ, Peachey H, Hannoush P *et al*. Effect of testosterone treatment on body composition and muscle strength in men over 65 years of age. *J. Clin. Endocrinol. Metab.* 1999; **84**: 2647–53.
37. Overend TJ, Cunningham DA, Kramer JF, Lefcoe MS, Paterson DH. Knee extensor and knee flexor strength: Cross-sectional area ratios in young and elderly men. *J. Gerontol.* 1992; **47**: M204–10.
38. Larsson L, Grimby G, Karlsson J. Muscle strength and speed of movement in relation to age and muscle morphology. *J. Appl. Physiol.* 1979; **46**: 451–6.
39. Stalberg E, Borges O, Ericsson M *et al*. The quadriceps femoris muscle in 20–70-year-old subjects. Relationship between knee extension torque, electrophysiological parameters, and muscle fiber characteristics. *Muscle Nerve* 1989; **12**: 382–9.
40. Frontera WR, Hughes VA, Fielding RA, flatarone MA, Evans WJ, Roubenoff R. Aging of skeletal muscle: A 12-yr longitudinal study. *J. Appl. Physiol.* 2000; **88**: 1321–6.
41. Bhasin S, Storer TW, Berman N *et al*. Testosterone replacement increases fat-free mass and muscle size in hypogonadal men. *J. Clin. Endocrinol. Metab.* 1997; **82**: 407–13.
42. Sih R, Morley JE, Kaiser FE, Perry 3rd HM, Patrick P, Ross C. Testosterone replacement in older hypogonadal men: A 12-month randomized controlled trial. *J. Clin. Endocrinol. Metab.* 1997; **82**: 1661–7.
43. Sheffield-Moore M, Urban RJ, Wolf SE *et al*. Short-term oxandrolone administration stimulates net muscle protein synthesis in young men. *J. Clin. Endocrinol. Metab.* 1999; **84**: 2705–11.
44. Lexell J, Taylor CC, Sjostrom M. What is the cause of the ageing atrophy? Total number, size and proportion of different fiber types studied in whole vastus lateralis muscle from 15- to 83-year-old men. *J. Neurol. Sci.* 1988; **84**: 275–94.
45. Lexell J. Human aging. Muscle mass, and fiber type composition. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* 1995; **50**: 11–16.
46. Larsson L, Ansved T. Effects of ageing on the motor unit. *Prog. Neurobiol.* 1995; **45**: 397–458.
47. Denti L, Pasolini G, Sanfelici L *et al*. Aging-related decline of gonadal function in healthy men: Correlation with body composition and lipoproteins. *J. Am. Geriatr. Soc.* 2000; **48**: 51–8.
48. Elashoff JD, Jacknow AD, Shain SG, Braunstein GD. Effects of anabolic-androgenic steroids on muscular strength. *Ann. Intern. Med.* 1991; **115**: 387–93.
49. Bhasin S, Storer TW, Berman N *et al*. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N. Engl. J. Med.* 1996; **335**: 1–7.
50. Wang C, Eyre DR, Clark R *et al*. Sublingual testosterone replacement improves muscle mass and strength, decreases bone resorption, and increases bone formation markers in hypogonadal men: A clinical research center study. *J. Clin. Endocrinol. Metab.* 1996; **81**: 3654–62.
51. Urban RJ, Bodenburg YH, Gilkison C *et al*. Testosterone administration to elderly men increases skeletal muscle strength and protein synthesis. *Am. J. Physiol.* 1995; **269**: E820–6.
52. Tenover JS. Effects of testosterone supplementation in the aging male. *J. Clin. Endocrinol. Metab.* 1992; **75**: 1092–8.
53. Winter EM. Scaling: Partitioning out differences in size. *Pediatr. Exer. Sci.* 1992; **4**: 296–301.
54. Rogers DM, Olson BL, Wilmore JH. Scaling for the VO_2-to-body size relationship among children and adults. *J. Appl. Physiol.* 1995; **79**: 958–67.