RESEARCH REPORT

Reliability of Isometric Quadriceps Muscle Strength Testing in Young Subjects and Elderly Osteo-arthritic Subjects

J A Oldham
T E Howe

Key Words
Quadriceps, isometric, strength, reliability.

Summary
The purpose of this study was to examine the reliability of apparatus and the variability of subject performance in isometric strength testing of the quadriceps in young subjects and elderly osteo-arthritic subjects. Repeated applications of various loads were applied to a torque transducer within a specially designed chair for assessing isometric quadriceps function. The test-retest reliability of this procedure was high. Coefficients of variation were 0% for all test conditions, and the transducer acted linearly throughout the calibration range. Subject performance of maximum voluntary isometric contractions was measured six times over a 30-minute period and over three days of the same week. The reliability of these procedures was high for all groups. Repeated measures MANOVAs revealed no significant learning or training effects over one week ($p > 0.05$) and no fatiguing effects over a 30-minute period ($p > 0.05$). These results indicate that this test may be used for successive measures of strength in these subject groups.

Introduction

The quadriceps muscles are important in most weight-bearing functional activities involving the lower limb, and provide protection to the structures of the knee. However, these muscles are often preferentially affected by injury (Smillie, 1970; Young *et al*, 1982), and osteo-arthritis (OA) of the knee joint (Gibson *et al*, 1989). Impairment of quadriceps muscle performance may cause profound disability. For this reason, many patients are referred to physiotherapy departments for rehabilitation.

Rising costs of health care have promoted the Government to demand that the viability of new and existing interventions is evaluated (NAHAT, 1993). To achieve these ends, the extent of patients' impairments, and subsequent disability, should be quantified by appropriate objective measures. These measures should be specific to the muscles under investigation and proven to be both valid and reliable. Sequential measures may then be used by the clinician to assess the efficacy of interventions, formulate the choice of pertinent treatment modalities, and monitor disease processes.

Strength may be assessed by measuring the peak force to torque generated by the muscle. Force is 'that which tends to change an object's momentum' (Daintith and Deeson, 1980) and is measured at the point of application; and torque is 'a turning force' (Daintith and Deeson, 1980) and is measured at the axis of rotation.

Methods of Assessing Muscle Strength

Manual muscle testing (MMT), employing the Medical Research Council scale (MRC, 1986) and its equivalents, is traditionally used as a clinical measure of muscle performance. The assessor manually estimates the strength of a muscle and grades it on a predetermined scale. These scales were developed for use on patients recovering from peripheral nerve injuries; grade 0 indicating no contraction and grade 5 full function. However, strict adherence to standardised positioning of the patient for each muscle tested is required. The scales as a whole are insensitive and are not accurate enough for research projects. Indeed the muscle strength of most patients is covered by only one or two grades of MMT (Munsat, 1990).

The need for more objective and sensitive methods of assessment have led to the development of hand-held dynamometers, such as the Hammersmith myometer (Edwards and Hyde, 1977; Hyde *et al*, 1983), which measure isometric muscle strength. However, these have some drawbacks. Correct positioning of both subjects and testers is required to ensure that the action of the muscle under test is correctly opposed, and stabilisation of the dynamometer may prove difficult when the subject is stronger than the tester. For comparative measures it is essential that the joint and the dynamometer are placed in the same position for every test (Stokes, 1985; Delitto, 1990). Therefore, in view of these difficulties, hand-held dynamometers are rarely of use, and have been recommended only for children and adults with severe muscle weakness (Edwards and Hyde, 1977).

A more sensitive measure of muscle performance may be obtained from isokinetic dynamometers. These are quantitative myometers, using hydraulic or motor driven instruments to impose constant velocity movement through a given range at present velocities, commonly 30, 60, 90, 120° s^{-1} (Delitto, 1990). These devices estimate the muscular torque generated during testing. The reliability of these dynamometers has been investigated for the quadriceps muscles. The test-retest reliability coefficient for repeated applications was $r = 0.995$, $n = 70$ (Thorstensson, 1976). Although isokinetic testing has gained popularity in the last decade, Reilly (1985) proposes that these devices may interfere with the natural acceleration pattern used in explosive competitive actions and should be used with caution. However, the major disadvantage of isokinetic dynamometers for most physiotherapy departments is their prohibitive cost.

A cheaper and more practical alternative is to use a strain gauge device. Several such devices have been described for measuring isometric strength of the quadriceps (Hyde *et al*, 1983; Maxwell *et al*, 1984), and the adductor pollicis muscle (Edwards *et al*, 1977). The test-retest reliability of such a device has been investigated for the quadriceps muscles ($r = 0.91$, $n = 24$) with no significant difference between observers (Hyde *et al*, 1983). However, these results should be judged with caution as the use of correlation coefficients is an inappropriate measure of reliability (Bland and Altman, 1986).

Measures of isometric quadriceps strength using a strain gauge chair dynamometer have been shown to be repeatable on different days in normal males and females of different ages. The coefficient of variation in all groups of subjects was 8% (Young *et al*, 1985). Similar studies in boys with muscular dystrophy, however, showed day-to-day readings varied more than normal. This stresses the importance of testing the repeatability of measures in the population in which they are to be used (Stokes, 1985).

A further option is to use a torque transducer to estimate the torque produced at the anatomical joint during isometric contractions. Such devices include the Kin Com (Chattanooga) and the Biodex (Akron). Unlike hand-held dynamometers and strain gauges, measurements of torque are not affected by the point of application of the instrument (Delitto, 1990). With suitable stabilisation these transducers are capable of withstanding large amounts of torque and cannot be overpowered by the subject, thus eliminating the examiner's strength as a factor.

Sandardisation of Isometric Torque Assessment

Careful standardisation of assessment procedures is essential when sequential measurements of muscle performance are to be compared. When testing the quadriceps, the subject must be positioned to prevent vicarious movements occurring. This may be achieved by using a 'quadriceps chair' initially described by Tornvall (1963) and modified by others (Edwards *et al*, 1977; Hyde *et al*, 1983; Maxwell *et al*, 1984). The dimensions of the chair should be adjustable to position the subject's hip and knee angles at 90°. This ensures that muscle length is constant between tests. Any alteration in length may affect the force developed by the muscle and therefore give inaccurate comparisons of results (Astrand and Rodahl, 1988). When the knee is positioned at 90° the need to perform gravity correction calculations is eliminated. The provision of a pelvic strap prevents concomitant hip extension (Edwards *et al*, 1977).

To confirm that subjects are producing a maximum voluntary contraction, twitch interpolation techniques have been described (Merton, 1954; Belanger and McComas, 1981). This involves superimposing supermaximal 1 Hz twitches on to muscles undergoing progressively stronger voluntary contractions. Any motor units not fully activated during the contraction give a detectable twitch response, the amplitude of which decreases as the force of contraction increases. This extra force generated is independent of the percentage of the muscle stimulated (Chapman *et al*, 1984). If the contraction is maximal, the interpolated twitches are no longer detectable or represent a decrease in observed force.

The purpose of this study was to examine the test-restest reliability of the torque transducer (study one); the variability of performance during maximal voluntary isometric contractions (MVICs) of the quadriceps muscles in young subjects and elderly subjects with osteo-arthritis of the knee joint (studies 2 and 3).

Methods

Subjects

All subjects described below were purposively selected for participation in the studies described. A total of 12 young subjects (six male and six female, age range 19-35 years) and 13 elderly subjects (seven male and six female, age range 65-87 years) were recruited into the study investigating subject variability in strength within a 30-minute period (study 2). Twelve young subjects (six male and six female, age range 19-33 years) and 15 elderly subjects (seven male and eight female, age range 62-87 years) were recruited into the study investigating subject variability in strength over a one-week period (study 3). All the elderly osteo-arthritic subjects were recruited from a waiting list for total knee joint replacements. All were on stable drug therapy and did not have any history of neurological insult. The young healthy subjects had not suffered any recent trauma to the musculoskeletal system in the past 12 months. Ethical approval was provided by Salford Health Authority and informed consent was obtained from each subject before testing.

Apparatus

A modified version of the quadriceps chair described by Tornvall (1963) was used, with a seat which was adjustable for variations in thigh length.

A torque transducer (Sensotec QFFM-9/227-03 with a range of 0-3,000 lb) was attached to the frame of the chair on left- and right-sided mountings to facilitate measurement of either leg. The transducer output was connected to a high-speed pen recorder (Multitrace 2, Lectromed). A padded bar, against which the lower leg exerted force, was attached to the lever arm of the transducer, the position of which was adjustable to ensure patient comfort during testing. The padded bar was situated on the anterior border of the tibia above the medial malleolus. The torque transducer and the chart recorder were calibrated immediately before testing.

A stimulator (Type 3072, Digtimer) was used with two re-usable self-adhesive electrodes (5 in x 3 in, Chattanooga UK), to deliver the interpolated 1 Hz stimulus with a pulse width of 150 μs and maximum current of 100 mA to the quadriceps muscles. One electrode was placed on the superior lateral aspect of the thigh and the other approximately 3 cm above the base of the patella.

Procedure

Subjects were tested in the sitting position with their legs hanging vertically. The seat dimensions were adjusted to ensure that the hip and knee angles were at 90°, and the pelvis was fixed by a strap secured to the frame. The length of the lever arm of the torque transducer and the position of the backrest were kept constant for each subject. A warm-up period of two or three contractions with feedback familiarised subjects with the procedure. During this period the 1 Hz stimulus was delivered to their quadriceps at increasing intensities, to determine the intensity (volts) required to produce a supramaximal twitch on the torque trace (ie that point at which torque increased no further). This was followed by a rest period of five minutes to avoid fatigue.

Instructions to subjects were standardised. The interpolated 1 Hz stimulus was delivered to the subjects immediately before they began a contraction of their quadriceps. The trial ended after five seconds or earlier if the interpolated twitch was abolished. A rest period of five minutes was allowed between successive trials. Subjects remained seated in the same position on the 'torque chair' between recordings.

Study 1: Test-retest reliability of apparatus

Six repeated applications of various loads (1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60 kg) were applied at a distance of 0.5 m from the centre of the transducer, which was securely fixed in a socket attached to the chair. These loads gave a range of applied torque 4.9 to 294.3 Nm. Pen deflections from the zero position (unloaded transducer) were recorded, at two gain settings, with x 20 preamplification, for both left and right transducer arrangements.

Study 2: Subject variability performing six MVICs in a 30-minute period

To test the hypotheses that five minutes was enough time to recover between tests, subjects were asked to perform six maximum voluntary isometric contractions of their quadriceps over a 30-minute period. Five-minute rest intervals were allowed between successive contractions. The twitch interpolation technique was applied during each recording.

Study 3: Subject variability of performance of MVIC over three days in the same week

Subjects were asked to perform three maximum voluntary quadriceps contractions on three occasions in the same week. Five-minute intervals were allowed between each contraction. The twitch interpolation technique was applied during each recording. The highest reading of the three values produced on each day was used for analysis (cf Tornvall, 1963).

Data Analysis

Data obtained from subjects during this study were analysed by repeated measures MANOVAs using a statistical package for the social sciences (SPSS/PC+). The linearity of the torque transducer was examined by performing Pearson's correlation coefficient on the data obtained during calibration procedures. Coefficient of variation was calculated as the standard deviation/mean of subsequent recordings x 100%.

Results

Study 1

A significant position correlation (Pearson's $r = 0.999$, $p < 0.001$) was seen between the load applied and the amount of pen deflection produced, for all test conditions. Therefore the transducer acted linearly through the calibration range (4.9 - 294.3 Nm). This procedure was performed on six separate occasions over a six-month period. Coefficients of variation for all calibration settings (200 mV, 500 mV, and 1 V plus x 20 amplification, for both left and right sides) were 0% (ie results were identical on each occasion). Drift, calculated as the difference between pen deflections on the chart recorder, over a 30-minute period at a calibration measurement of 125 Nm was 0 Nm. Margins of error for reading the torque were ± 0.5 mm over total deflections up to 100 mm. This translated to errors of ± 1.5 and ± 3.5, and ± 0.6 Nm at gain settings of 200 mV, 500 mV and 1 V plus 20 x pre-amplifier, full scale deflection respectively.

Study 2

Results of repeated measures analysis of variance tests revealed non-significant ($p > 0.05$) within-subjects effects for all groups (table 1 and graphs).

Table 1: Results of repeated measures MANOVA on data obtained during MVICs of the quadriceps muscles, performed over a 30-minute period, from young subjects with osteo-arthritis (OA) of the knee joint

Subjects	*n*	*Mean % CV*	*Within-subject effects F values*	*p values*
Young women	6	6.4	$F_{25,5} = 1.87$	0.13
Young men	6	3.6	$F_{25,5} = 0.86$	0.52
Elderly women with OA	6	9.8	$F_{25,5} = 0.11$	0.99
Elderly men with OA	7	5.9	$F_{30,5} = 1.58$	0.20

Study 3

The maximum torque value from the three values obtained on each occasion was used in the analysis. Results of a repeated measures analysis of variance (Pillais multivariate test) on these values revealed no significant difference at the 0.05 level for any individual group, or when men and women were combined into young and elderly groups (table 2).

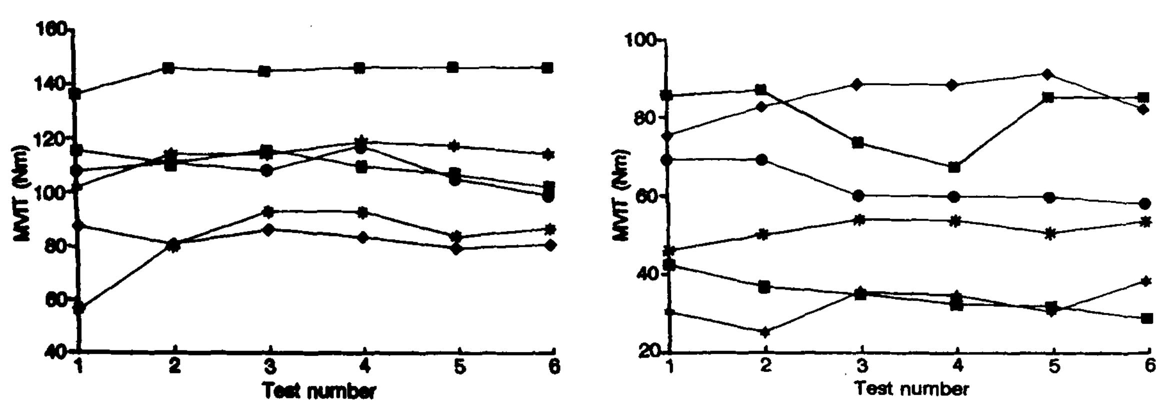
Table 2: Results of repeated measures MANOVA on data obtained during MVICs of the quadriceps muscles, performed on three occasions in the same week, from young subjects and elderly subjects with osteo-arthritis of the knee joint

Subjects	*n*	*F values*	*p values*
Young women	6	0.38	0.71
Young men	6	0.45	0.67
Young (women and men)	12	0.44	0.65
Elderly women with OA	8	1.83	0.24
Elderly men with OA	7	0.35	0.72
Elderly (women and men) with OA	15	1.58	0.24

Discussion

Measurements of muscle performance are used clinically to assess the outcome of therapeutic interventions in patients with injuries to or pathology of the knee joint. The appropriateness of these measures must be examined before they may be universally applied.

It is important to investigate the reliability of equipment, to determine the margin of error and limitations of the procedure, before examining subject variability. If the error margin is small compared to expected values or expected change in response to treatment the system may be used confidently. When testing patients, however, the readings obtained from their weakened muscles may be much smaller than those of normal subjects and the improvement following therapy smaller. Therefore, the measurement system should be sensitive to these small changes and produce reliable readings throughout the range of application. If this is not the case, observed improvements may be inappropriately attributed



Graphs showing the results of repeated measures of MVIT of the quadriceps, performed over a 30-minute period, on young female subjects (left) and elderly female subjects with osteo-arthritis of the knee joint respectively

to the therapeutic intervention (type I error) or an effect may be present but not observed (type II error) (Thomas and Nelson, 1990).

Study 1

The test-retest reliability of the torque transducer used in this study was high and acted linearaly throughout the calibration range. The maximum margin of error was ± 3.5 Nm recorded at the highest gain setting. Coefficient of variations (CVs) for all test conditions were 0%, through the calibration range (4.9 to 294.3 Nm). Torque values obtained from subjects in studies 2 and 3 ranged from 25.1 to 268 Nm and were within the calibration range of the transducer. Therefore measurements of torque using the apparatus in the present study were assumed to be accurate and reliable.

Study 2

To test the hypothesis that a five-minute rest interval was sufficient recovery time between tests, six repeated measures of MVIT, with five-minute rest intervals between each, were performed on the same day. The reliability of this procedure was high for all groups (table 1). Within-subject effects were not significant for any group ($p > 0.05$), confirming that there was no constant pattern over time, and indicating that there were no training or fatiguing effects present during the 30-minute testing period. The mean percentage coefficients of variation for young women, young men, elderly women and elderly men were 6.4, 3.6, 9.8 and 5.9 respectively. These results compare with those of Young *et al* (1985) of 8% using a strain gauge device on the quadriceps of young normal subjects, and with normal biological systems showing a 10-15% coefficient of variability (Stokes, 1985). Thus a natural variability of performance exists within human beings of between 10-15%. Therefore this test may be confidently used as a sequential measure to assess quadriceps strength in this population.

Study 3

No significant difference at the 0.05 level was seen for any group when a repeated measures MANOVA (Pillais multivariate test) was performed on data obtained on three days of the same week (table 2). This suggests that no significant training or learning effects were present over the three test days. This time period was used to ensure that subject variability did not alter significantly over time but was short enough to reduce the effects of life events, eg deterioration due to exacerbation of pathology, on muscle performance. The results of the present study indicate that the best of three torque values during isometric testing of the quadriceps is a reliable test and may be employed as a sequential measure, in the clinical setting.

Limitations of the Study

The principal limitation of this study is that only two populations were studied, ie young healthy subjects and elderly subjects with osteo-arthritis affecting the knee joint. Stokes (1985) stresses the importance of testing the repeatability of measures in the population in which they are to be used. Results from this study cannot be used to predict with confidence the effects of testing different populations and thus need to be extended to different populations in the future. Furthermore, the studies described incorporated a relatively small number of subjects purposively selected for participation in this study. A random selection of a much larger population would provide further conclusive evidence of the reliability of isometric quadriceps muscle strength testing.

Conclusions

To fulfil the requirements of health care professionals, to evaluate the efficacy of new and existing practice, it is essential that appropriate objective measurements are employed. It may be concluded from this study that the use of a torque transducer in conjunction with a 'quadriceps chair' to measure MVIT of the quadriceps is reliable, inexpensive and easy to use in a clinical setting. The best of three torque values obtained on the same occasion may be confidently used as sequential comparative measures of strength. These tests may be used to monitor a disease process, guide physical rehabilitation programmes, and help to design and evaluate the efficacy of innovative therapeutic interventions.

Authors

Dr J A Oldham PhD BSc RGN is reader in the School of Nursing Studies, University of Manchester.

Dr T E Howe PhD MSc MCSP is a research associate, University of Manchester School of Nursing Studies, Centre for Habilitation and Muscle Performance, Ladywell Hospital.

Address for Correspondence

Dr J A Oldham, Centre for Habilitation and Muscle Performance, Ladywell Hospital, Manchester M9.

Acknowledgements

The authors would like to thank Professor R C Tallis, Dr T Petterson, Mr Smith and the Medical Physics Department, Hope Hospital, Salford, for their contribution to this study. Dr Howe was funded by North West Regional Health Authority while undertaking this study.

References

Astrand, P O and Rodahl, K (eds) (1988). *Textbook of Work Physiology*, McGraw Hill, New York, 3rd edn, pages 40–42.

Belanger, A Y and McComas, A J (1981). 'Extent of motor unit activation during effort,' *Journal of Applied Physiology*, 51, 5, 1131–35.

Bland, J M and Altman, D G (1986). Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet*, **1**, 8476, 307–310.

Chapman, S J, Edwards, R H T, Grieg, C and Rutherford, O (1984). 'Practical application of the twitch interpolation technique for the study of voluntary contraction of the quadriceps muscle in man', *Journal of Physiology*, **353**, 3.

Daintith, J and Deeson, E (eds). (1980). *Key Facts Dictionary of Physics*, Intercontinental Book Productions Ltd, London.

Delitto, A (1990). 'Isokinetic dynamometry', *Muscle and Nerve*, **13** (suppl), S53–S57.

Edwards, R H T and Hyde, S (1977). 'Methods of measuring muscle strength and fatigue', *Physiotherapy*, **63**, 2, 51–55.

Edwards, R H T, Young, A, Hosking, G P and Jones, D A (1977). 'Human skeletal muscle function: Description of tests and normal values', *Clinical Science and Molecular Medicine*, **52**, 283–290.

Gibson, J N A, Morrison, W L, Scrimgeour, C M, Smith, K, Stoward, P J and Rennie, M J (1989). 'Effects of therapeutic percutaneous electrical stimulation of atrophic human quadriceps on muscle composition, protein synthesis and contractile properties', *European Journal of Clinical Investigation*, **19**, 206–212.

Hyde, S A, Godard, C M and Scott, O M (1983) 'The myometer: Development of a clinical tool', *Physiotherapy*, **69**, 12, 424–427.

Maxwell, J D, Hodt, P and Taylor, W H (1984). 'Portable chair for testing isometric muscle strength', *Lancet*, 8367, 18–19.

Medical Research Council (1986). Aids to the Examination of the Peripheral Nervous System, Bailliere Tindall, London.

Merton, P A (1954). 'Voluntary strength and fatigue', *Journal of Physiology*, **123**, 553–564.

Munsat, T L (1990). 'Clinical trials in neuromuscular disease', *Muscle and Nerve*, **13**, suppl, S3–S5.

National Association of Health Authorities and Trusts (NAHAT) (1993). *House of Commons Health Committee Interim Reports on NHS Trusts*, NAHAT Briefing No 38, January.

Reilly, T (ed) (1985). *Sports Fitness and Sports Injuries*, Faber, London, page 58.

Smillie, I S (1970). *Injuries of the Knee Joint*, E & S Livingstone, Edinburgh, page 1.

Stokes, M (1985). 'Reliability and repeatability of methods for measuring muscle in physiotherapy', *Physiotherapy Practice*, **1**, 71–76.

Thomas, J R and Nelson, J K (eds) (1980). *Research Methods in Physical Activity*, Human Kinetics Books, Champaign, 101–103.

Thorstensson, A (1976). 'Muscle strength, fibre types and enzyme activities in man', *Acta Physiologica Scandinavia*, suppl 443.

Tornvall, G (1963). 'Assessment of physical capabilities with special reference to the evaluation of maximal voluntary isometric muscle strength and maximal working capacity', *Acta Physiologica Scandinavica*, **58**, suppl 201.

Young, A, Hughes, I, Round, J M and Edwards, R H T (1982). 'The effect of knee injury on the number of muscle fibres in the human quadriceps femoris', *Clinical Science*, **62**, 277–234.

Young, A, Stokes, M and Crowe, M (1985). 'The size and strength of the quadriceps muscles of old and young men', *Clinical Physiology*, **5**, 145–154.