Eur J Appl Physiol (1994) 68:310–314

European Journal of Applied Physiology
and Occupational Physiology


# Reproducibility of computer measurement of maximal isometric strength and electromyography in sedentary middle-aged women

**Ari Heinonen**[1], **Harri Sievänen**[1], **Jukka Viitasalo**[2], **Matti Pasanen**[1], **Pekka Oja**[1], **Ilkka Vuori**[1]

[1] The UKK Institute for Health Promotion Research, P.O. Box 30, FIN-35501 Tampere, Finland
[2] The Research Institute for Olympic Sports, Jyväskylä, Finland

Accepted: 28 December 1993

**Abstract.** The objectives of this study were (i) to determine the reproducibility of computer measurements of isometric strength and related electromyography in several muscle groups in sedentary middle-aged women, (ii) to evaluate the effects of different digital signal averaging methods on the reproducibility, (iii) to determine the final test score to be preferred in terms of improved reproducibility of isometric strength measurements, and (iv) to evaluate potential advantages provided by the computer measurement. Fifteen subjects were measured three times within a 2-week period. The measurements consisted of recordings of maximal isometric strength and rate of force production during trunk extension and flexion, leg extension and dominant forearm flexion with simultaneous recordings of surface electromyography, except in the trunk flexors. The following four final test scores were determined for each trial: the maximum of the three scores, the mean of the two highest scores, the median of the three scores and the mean of the three scores. The scores for the strength measurement were generally more reproducible (coefficient of variation, CV, approximately 6% and intraclass correlation coefficient ICCC, approximately 0.90) than those of the other measurements (CV>10%, ICCC 0.13–0.97). There was no obvious preference for any type of final test score or for the width of the averaging window in the computer analysis. For isometric strength the reproducibility of the computer measurements was comparable to that of the voltmeter assessments. Computer analysis seems to be a versatile method for determining parameters of neuromuscular performance with reasonable reproducibility.

**Key words:** Reproducibility – Precision – Isometric strength – Electromyography – Middle-aged women

*Correspondence to:* A. Heinonen

## Introduction

The neuromuscular performance of different muscle groups may be assessed by simultaneous electromyographic (EMG) analysis and measurement of force production. The reproducibility of these methods has been evaluated in many studies. For example, the reproducibility of maximal isometric strength measurements in different limb and torso positions, expressed as the coefficient of variation (CV), varies from some 4 to 16%, whereas the day-to-day CV of the integrated EMG varies from 7 to 20% when the precision of the EMG measurement is determined for various types of isometric contractions in different muscle groups and in individual muscles (Daanen et al. 1990; DeVries 1968; Gollhofer et al. 1990; Komi and Bushkirk 1970; Patterson and Baxter 1988; Veiersted 1991; Viitasalo et al. 1980; Yang and Winter 1983; Zeh et al. 1986).

The age, gender, and assessment experience of subjects can affect the magnitude of the reproducibility (Baumgartner 1989). In previous studies the subjects have been young males (Daanen et al. 1990; Gollhofer et al. 1990; Patterson and Baxter 1988; Viitasalo et al. 1980). Therefore it is important to study the reproducibility of EMG and force measurements in other groups, for example, in sedentary middle-aged women – a relevant group in osteoporosis research (Bauer et al. 1993; Sinaki et al. 1986; Sinaki and Offord 1988) – who are not accustomed to maximal effort.

Maximal isometric strength and the related EMG are usually measured in a selected number of test trials. The highest strength value obtained in the test trial is generally used as the test score. However, in terms of the purpose of the test and the overall test reliability, this may not be the optimal choice (Baumgartner 1989). Furthermore, the potentially beneficial effects of digital signal processing and computer analysis on reproducibility have not been evaluated sufficiently.

The objectives of this study were (i) to determine the reproducibility of computer measurements of isometric strength and the related EMG in several muscle groups, (ii) to evaluate the effects of different digital signal averaging methods on the reproducibility, (iii) to determine the final test score to be preferred in terms of improved measurement reproducibility, and (iv) to evaluate the potential advantages provided by computer measurement in assessing the isometric muscle strength of sedentary middle-aged women.

## Methods

*Subjects.* Fifteen sedentary middle-aged and healthy women participated in the study. Informed consent was obtained from each subject. Their mean (SD) age was 53 (4) years, mean height 162 (6) cm and mean weight 71.9 (12.0) kg.

*Measurements.* The isometric strength of trunk extension and flexion was measured by a trunk dynamometer (Digitest, Muurame, Finland; measurement range 0–2.5 kN; sensitivity 2 $mV \cdot N^{-1}$; time constant $\tau$ approximately 0.1 s). The force transducer was adjusted horizontally at the level of the subject's inferior angle of the scapula. The subjects were tested in the standing position, the pelvis and knees being stabilized by the testing apparatus.

The isometric strength of bilateral leg extension was measured by a leg press force transducer dynamometer (Tamtron, Tampere, Finland; measurement range 0–25 kN; sensitivity 2 $mV \cdot N^{-1}$; time constant $\tau$ approximately 0.5 s). The subjects sat on a chair in an upright position keeping their hands on the chair arms, their hips in 45° flexion and their knees and ankles in 90° flexion. They then pressed maximally against force transducers located under their feet.

The isometric strength of forearm flexion was measured by an arm dynamometer (Digitest). The subjects stood with the upper arm in a vertical direction with their elbow flexed 90°, and stabilized by the testing apparatus. The wrist was in a neutral position. The untested arm was in a relaxed position at the side of the body. The force transducer (measurement range 0–2.5 kN; sensitivity 2 $mV \cdot N^{-1}$; time constant $\tau$ approximately 0.1 s) was mounted on the hand grip.

All of the force transducers proved to be linear ($r$ approximately 1) within the measurement range (10–90 kg for the trunk measurements, 20–355 kg for the leg measurements and 10–40 kg for the arm measurements). The transducers were calibrated monthly. The long-term stability over 1 year, expressed as the CV, was better than 1%. All of the values representing maximum isometric strength were manually recorded for further analysis and comparison with those produced by the computer analysis.

The surface EMG was recorded by disposable surface electrodes (E-10-VS, Medicotest, Olstykke, Denmark). The electrode sites were cleaned and decornified by rubbing the skin with sandpaper, the electrodes being placed longitudinally on the belly of the muscle as suggested by Gollhofer et al. (1990). The interelectrode distance was 18 mm. The electrode sites were marked with ink and the subjects kept the marks throughout the study so that the electrodes could be properly repositioned. Before the measurement, the quality of the EMG was checked for artefacts using an oscilloscope. The EMG was amplified by a five-channel differential amplifier (Myosystem 1008, Noraxon, Oulu, Finland), its input impedance was more than 1 MΩ, the gain 1000 and 3-dB bandwidth 20–350 Hz.

The measurement sessions were repeated three times within a 2-week period. The testing protocol and the measurement position were kept the same throughout the study and the measurements were conducted by the same tester (A.H.). The measurements consisted of recordings of the maximal isometric strength ($F_{w,\max}$) and rate of force production ($\dot{F}$) during trunk extension and flexion, leg extension and dominant forearm flexion, in this order, and simultaneous recordings of the surface EMG produced by the trunk extensors (erector spinae at vertebral level L1 and L3 bilaterally), leg extensors (rectus femoris, vastus lateralis and vastus medialis in the dominating limb) and forearm flexors (flexor carpi ulnaris and brachioradialis in the dominating limb). Before the recordings the subjects performed three warm-up contractions to become familiar with the testing apparatus and protocols. The isometric force production and EMG were recorded during three maximal trials. The subjects were asked to do the trial as intensively and fast as possible. An oscilloscope was used for visual feedback, and verbal encouragement was used to motivate the subjects to reach their maximum. All of the subjects maintained their habitual level of activity during the study and they were asked to avoid strenuous exercise prior to the measurement days.

All of the raw data representing force and EMG signals during each test trial were sampled digitally at 1 kHz by a DT2801 12-bit A/D converter (Data Translation, Marlborough, Mass., USA) installed in a Toshiba TS3200 microcomputer and stored for further processing.

*Digital signal processing.* Digital processing was implemented by NST software (Noraxon) as follows. First, all of the sampled signals were smoothed by a 101-impulse moving average low-pass filter (Challis and Kitney 1982). In addition, the EMG signal was digitally rectified prior to the filtering, the average of the whole EMG sequence being the zero level. The preprocessed EMG sequences from each muscle measured during the same trial were summed. From the preprocessed signal sequences ($F_i$ for force, $E_i$ for the summed EMG), four test scores representing maximal isometric strength ($F_{w,\max}$), maximal rate of force production ($\dot{F}_{d,\max}$) and integrated EMG [$IEMG_w$] were determined for each trial.

The $F_{w,\max}$ scores were determined according to the following formula:

$$F_{w,\max} = \frac{1}{w+1} \sum_{i=T-\frac{w}{2}}^{T+\frac{w}{2}} F_i$$

where $w$ is the width of the averaging window (100, 500, 1000, 2000 samples) and $T$ is the time point of the global maximum observed in the averaged $F_i$ sequence between the start [$i = S + (w/2) + 1$] and end ($i = E - w/2$) points (Fig. 1). $S$ is the starting point of the trial and $E$ is the end point.

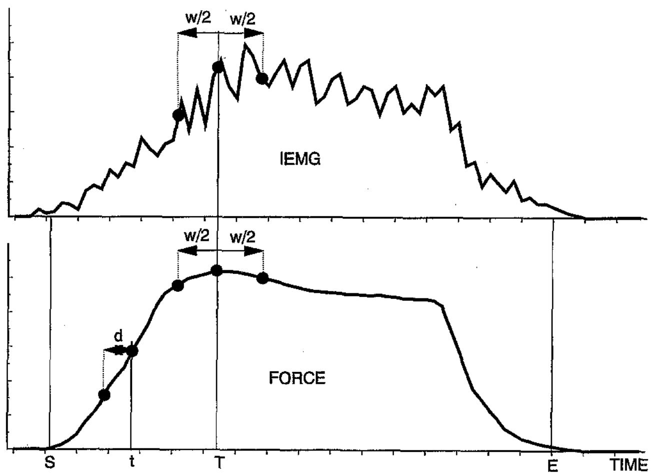


**Fig. 1.** Diagram of the determination of the parameters analysed in the study, where $w$ is the width of the averaging window (100, 500, 1000, 2000 samples), $T$ is the time point of the global maximum observed in the force production curve, $S$ is the starting point of the trial, $E$ is the end point, $d$ is the distance (5, 10, 20, 30 samples) between the points used for the difference calculation and $t$ is the time point of the global maximum in the rate of force production. The *vertical axis* represents the amplitude in arbitrary units

The $\dot{F}_{d,max}$ scores were determined according to the following formula:

$$\dot{F}_{d,max} = \frac{F_t - F_{t-d}}{d}$$

where $d$ is the distance (5, 10, 20, 30 samples) between the two samples used for the difference calculation and $t$ is the time point of the global maximum in the $\dot{F}_{d,max}$ sequence between the starting ($i=S+d+1$) and end ($i=E-d$) points (Fig. 1).

In a manner similar to the determination of the $F_{w,max}$ scores, the $IEMG_w$ scores were determined according to the following formula:

$$IEMG_w = \frac{1}{w+1} \sum_{i=T-\frac{w}{2}}^{T+\frac{w}{2}} E_i$$

where $w$ and $T$ are as defined for $F_{w,max}$ above.

As possibilities for the final test score we determined the following four values for each trial:
1. The maximum of the three scores (MAX)
2. The mean of the two highest scores (MEAN2)
3. The median of the three scores (MED)
4. The mean of the three scores (MEAN3)

*Statistical methods.* The mean and SD are given as descriptive statistics. As the measures of test reproducibility the CV [(SD/mean) × 100%], representing the relative magnitude (in per cent) of the score variability from trial to trial or day to day, and the intraclass correlation coefficient (ICCC), representing the error-free proportion of the inter-subject score variation, were determined. A one-way analysis of variance (ANOVA) was used to calculate the ICCC which considered the score variability between subjects and days, but not between trials. Evidently, measurement imprecision arises both from trial-to-trial and day-to-day variation. Thus to gain insight into the effects of variance on day-to-day behaviour in strength measurements and due to the electrode repositioning in EMG measurements, the trial-to-trial CV was determined. However, in this study, an actual test was always assumed to consist of three test trials, out of which the final test score was determined, and therefore the day-to-day reproducibility was considered the primary quantity, and the day-to-day CV and ICCC were determined accordingly. Only the test results of those subjects who successfully completed the test on all 3 days were included in the analyses. A two-way ANOVA was used to study the effects of averaging the windows and of the final test scores on reproducibility. The level of statistical significance was chosen to be $P<0.05$.

## Results

The maximal muscular performance characteristics of the subjects on the first measurement day are given in Table 1. The reproducibility of the non-computerized measurements, based on the voltmeter values for each type of final test score, is given in Table 2.

Figure 2 shows the reproducibility of the final test scores for $F_{w,max}$, $\dot{F}_{d,max}$ and the $IEMG_w$ scores. The trial-to-trial reproducibility (Table 3) was considerably better than the day-to-day reproducibility, except for $\dot{F}_{d,max}$, for which no difference was observed. In addition the recording of $F_{w,max}$ was generally more reproducible than the other measurements. Although individual CVs of any test score did not depend on the individual strength levels within each muscle group, the ICCC seemed to be associated with the absolute level

**Table 1.** Average maximal muscular performance characteristics of the subjects on the first measurement day

| | Maximal force (N) | | | Maximal rate of force production ($N \cdot s^{-1}$) $d=10$ ms | | | Integrated EMG ($mV \cdot s$) $w=1000$ ms | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Range | Mean | SD | Range | Mean | SD | Range |
| Trunk extension | 597 | 110 | 440– 770 | 993 | 1146 | 600–4400 | 115.1 | 71.4 | 56.0–281.8 |
| Trunk flexion | 453 | 64 | 360– 560 | 1707 | 721 | 700–3400 | | | |
| Leg extension | 1477 | 293 | 1060–2040 | 1667 | 592 | 1000–3000 | 104.7 | 50.1 | 44.3–243.0 |
| Forearm flexion | 185 | 32 | 133– 232 | 754 | 299 | 300–1300 | 220.4 | 131 | 69.5–552.5 |

*d*, Distance between the two samples used for the difference calculation; *w*, width of the averaging window

**Table 2.** The trial-to-trial and day-to-day coefficient of variation (CV) and intraclass correlation coefficient (ICCC) of the non-computerized measurements

| | Maximal strength | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Trunk extension ($n=15$) | | Trunk flexion ($n=15$) | | Leg extension ($n=15$) | | Forearm flexion ($n=15$) | |
| | CV | ICCC | CV | ICCC | CV | ICCC | CV | ICCC |
| Trial-to-trial | 2.8 | | 2.8 | | 2.2 | | 4 | |
| MAX | 5.1 | 0.88 | 5.9 | 0.82 | 5.4 | 0.93 | 9.7 | 0.63 |
| MEAN2 | 5.3 | 0.88 | 4.9 | 0.87 | 5.6 | 0.92 | 9.6 | 0.63 |
| MED | 5.6 | 0.88 | 4.2 | 0.89 | 5.8 | 0.92 | 9.6 | 0.62 |
| MEAN3 | 5.3 | 0.89 | 4.6 | 0.89 | 5.8 | 0.91 | 9.4 | 0,63 |

Each muscular exercise was carried out three times. MAX is the maximum of the three scores, MEAN2 the average of the highest two scores, MED the median of the three scores and MEAN3 the average of all three scores

**Table 3.** Trial-to-trial CV of the computer measurements based on different processing methods

| | Averaged window | | | |
|---|---|---|---|---|
| | $w=100$ ms<br>$d=5$ ms | $w=500$ ms<br>$d=10$ ms | $w=1000$ ms<br>$d=20$ ms | $w=2000$ ms<br>$d=30$ ms |
| | CV | CV | CV | CV |
| Trunk extension | | | | |
| Isometric strength ($n=12$) | 3.2 | 3.3 | 3.4 | 3.2 |
| Rate of force production ($n=12$) | 14.9 | 14.9 | 15.3 | 17.6 |
| Integrated EMG ($n=9$) | 7.7 | 8.1 | 14.3 | 5.0 |
| Trunk flexion | | | | |
| Isometric strength ($n=13$) | 2.8 | 2.6 | 2.9 | 2.9 |
| Rate of force production ($n=14$) | 19.0 | 17.5 | 18.3 | 20.0 |
| Leg extension | | | | |
| Isometric strength ($n=14$) | 2.4 | 2.3 | 2.4 | 2.5 |
| Rate of force production ($n=14$) | 8.4 | 12.0 | 11.7 | 12.9 |
| Integrated EMG ($n=14$) | 12.1 | 9.8 | 6.5 | 5.5 |
| Forearm flexion | | | | |
| Isometric strength ($n=9$) | 6.5 | 6.6 | 6.8 | 7.0 |
| Rate of force production ($n=9$) | 11.3 | 13.8 | 15.4 | 11.7 |
| Integrated EMG ($n=9$) | 15.7 | 17.0 | 22.8 | 14.7 |

$w$, Width of the averaging window for the maximal isometric strength and the integrated EMG; $d$, distance between the two samples used for the difference calculation for the maximal rate of force production

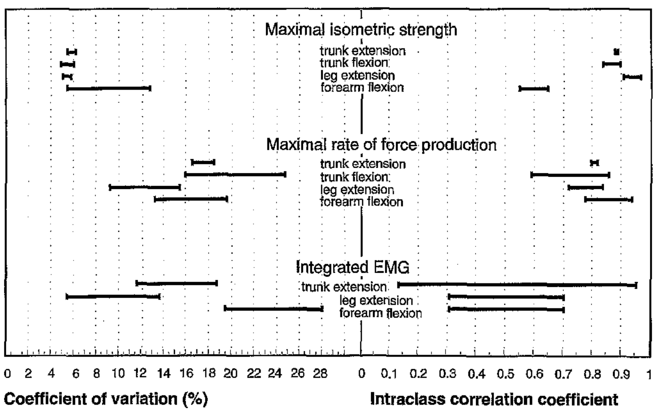


**Fig. 2.** The day-to-day reproducibility (coefficient of variation and intraclass correlation coefficient) of the computer measurements of maximal isometric strength, maximal rate of force production and integrated EMG. The *lines* represent the ranges from different final test scores and different processing method

of strength produced (Fig. 2). Especially the forearm flexion measurement was considerably less reliable than the other muscle group measurements, which involved more muscular strength.

There was no obvious preference for any type of final test score. However, the MEAN3 seemed to be more reproducible than the other scores for forearm flexion. No systematic beneficial effect of the increased width of the averaging window on the reproducibility of $F_{w,\max}$ or $\mathrm{IEMG}_w$ measurements was observed. In addition the reproducibility of $\dot{F}_{d,\max}$ was not systematically affected by changes in $d$.

The conventional measurement of $F_{w,\max}$ with a voltmeter was equally, or slightly more, reproducible than the corresponding computer measurement.

## Discussion

Measurements of maximal isometric strength and related EMG measurements are commonly used in intervention experiments in which changes in neuromuscular performance due to a specific treatment are of primary interest. This study showed that, even with the potential advantages of computer processing and analysis, the reproducibility of these measurements remains relatively poor (CV exceeding 10%), except in the determination of $F_{w,\max}$, for which the average reproducibility is moderate (CV approximately 6%). In the published literature, these reproducibility levels seem to be typical of maximal strength measurements (Viitasalo et al. 1980; Viljanen et al. 1991). Fortunately, maximal strength is usually the primary quantity assessed in muscular research, the other parameters providing supplementary information.

Test procedures and subject characteristics affect overall reliability (Baumgartner 1989). The former include the study and test arrangements, the measured object, subject fixation and measurement equipment and protocol. The latter comprises age, gender, physical and motor skills, motivation and readiness for the test. The arrangements and equipment for isometric strength testing have been properly validated and are reliable (Viitasalo et al. 1980; Viljanen et al. 1991). However the error variance due to subject characteristics remains large. The subjects in our study had been sedentary for several years and therefore probably lacked the physical skills and readiness needed for maximal performance tests. In addition, the effects of age on neuromuscular performance, for example $F_{w,\max}$, seem to be more pronounced after 50 years of age (Viitasalo et al. 1985; Viljanen et al. 1991) and many untrained people may not be able to fully activate the motor units within a muscle during maximal contractions. On the other hand, those who have trained can readily achieve a higher level of activation (Sale 1987). It is therefore likely that the test reproducibility obtained would also apply to younger, more active and better motivated subject groups. The study of Viljanen et al. (1989) supports these considerations. They found that the younger subjects showed better test reproducibility than older ones.

The selection of the final test score depends on objectives of the test (i.e., which parameter of neuromuscular performance the test score is supposed to indicate). We evaluated four test scores as good candidates for the final score. From the statistical point of view, the mean of the trial scores (here MEAN3) is more reliable than the best score (MAX) (Baumgartner 1989). The best score, on the other hand, is supposed to indicate the maximum performance of the subject at the time of the particular test. This is a reasonable assumption for athletes, but for ordinary, non-athletic persons the best score can yield a randomly deviant performance level. The observed difference between trial-to-trial and day-to-day CVs may demonstrate the considerable day-to-day variation in muscular performance. Obviously, the best indication of a subject's typical performance is the mean score. The modified mean (MEAN2) or median (MED) could act as an equally reproducible compromise between the other two scores.

In this study, it was impossible to improve the overall test reproducibility of the strength measurements by computer measurement and analysis. It is known that in many other physiological measurements (e.g., in electrocardiography) computer analysis provides significant advantages (Simoons et al. 1981). Apparently the lack of improvement was due to the greater subjective effect on performance test outcomes, and thus the reproducibility level achieved by standard methods could not be improved. On the other hand, the relevance of computer analysis in providing computational quantities, practically unaccessible by conventional methods, is without doubt. Although the CV of these quantities ($\dot{F}_{d,\max}$, $IEMG_w$) was poor (greater than 10%) in our study, their ICCC was mostly moderate (approximately 0.80), and therefore they may be potentially useful in intervention experiments. Further experimental studies are needed to determine which computational quantities would provide accurate and reliable information on the performance characteristic of interest.

In summary, we found no systematic preference, in terms of test reproducibility, for any width of averaging window, varying between 100 and 2000 ms, in the computer analysis of the isometric strength and EMG measurements, nor was the determination of the rate of force production affected. Furthermore, none of the final test scores evaluated in this study differed systematically from each other. We conclude that computer analysis offers a versatile determination of parameters representing neuromuscular performance at a reasonable level of reliability, but it provides no major advantages over conventional methods.

## References


Bauer DC, Browner WS, Cauley JA, Orwoll ES, Scott JC, Black DM, Tao JL, Cummings SR (1993) Factors associated with appendicular bone mass in older women. Ann Intern Med 118:657–665

Baumgartner TA (1989) Norm-referenced measurement: reliability. In: Safrit MJ, Wood TM (eds) Measurement concepts in physical education and exercise science. Human Kinetics Books, Champaign, Ill., pp 62–65

Challis RE, Kitney RI (1982) The design of digital filters for biomedical signal processing. 1. Basic concepts: the design of digital filters for biomedical signal processing. J Biomed Eng 4:267–278

Daanen HAM, Mazure M, Holewijn M, Van der Velde EA (1990) Reproducibility of the mean power frequency of the surface electromyogram. Eur J Appl Physiol 61:274–277

DeVries HA (1968) Efficiency of electrical activity as a physiological measure of the functional state of muscle tissue. Am J Phys Med 47:10–22

Gollhofer A, Horstmann GA, Schmidtbleicher D, Schönthal D (1990) Reproducibility of electromyographic patterns in stretch-shortening type contractions. Eur J Appl Physiol 60:7–14

Komi PV, Bushkirk ER (1970) Reproducibility of electromyographic measurements with inserted wire electrodes and surface electrodes. Electromyography 4:357–367

Patterson RP, Baxter T (1988) A multiple muscle strength testing protocol. Arch Phys Med Rehabil 69:366–368

Sale DG (1987) Influence of exercise and training on motor unit activation. Exerc Sport Sci Rev 15:95–151

Simoons ML, Hugenholtz PG, Ascoop CA, Distelbrink CA, DeLand PA, Vinke RVM (1981) Quantitation of exercise electrocardiography. Circulation 63:471–475

Sinaki M, Offord KP (1988) Effect on back muscle strength and bone mineral density of the spine. Arch Phys Med Rehabil 69:277–280

Sinaki M, McPhee MC, Hodgson SF, Merritt JM, Offord KP (1986) Relationship between bone mineral density of spine and strength of back extensor in healthy postmenopausal women. Mayo Clin Proc 61:116–122

Veiersted KB (1991) The reproducibility of test contractions for calibration of electromyographic measurements. Eur J Appl Physiol 62:91–98

Viitasalo JT, Saukkonen S, Komi PV (1980) Reproducibility of measurements of selected neuromuscular performance variables in man. Electromyogr Clin Neurophysiol 20:487–501

Viitasalo JT, Era P, Leskinen AL, Heikkinen E (1985) Muscular strength profiles and anthropometry in random samples of men aged 31–35, 51–55 and 71–75 years. Ergonomics 28:1563–1574

Viljanen, T, Viitasalo JT, Kujala U (1989) Isometric trunk extension and flexion strength of Finnish urban women and men aged 25–55 years. In: Telama R, Laakso L, Pieron M, Ruoppila I, Vihko V (eds) Physical education and life-long physical activity. Gummeruksen Kirjapaino, Jyväskylä, pp 550–559

Viljanen T, Viitasalo JT, Kujala UM (1991) Strength characteristics of a healthy urban adult population. Eur J Appl Physiol 63:43–47

Yang JF, Winter DA (1983) Electromyography reliability in maximal and submaximal isometric contractions. Arch Phys Med Rehabil 64:417–420

Zeh J, Hansson T, Bigos S, Spengler D, Battie M, Wortley M (1986) Isometric strength testing: recommendations based on a statistical analysis of the procedure. Spine 11:43–46