SPINE Volume 28, Number 13, pp 1455–1460
©2003, Lippincott Williams & Wilkins, Inc.

# Reliability of Computed Tomography Measurements of Paraspinal Muscle Cross-Sectional Area and Density in Patients With Chronic Low Back Pain

Anne Keller, MD,* Ragnhild Gunderson, MD,† Olav Reikerås, MD, PhD,‡ and Jens I. Brox, MD, PhD‡

**Study Design.** A reliability study was conducted.

**Objective.** To estimate measurement errors related to equipment and the observer in computed tomography measurements of cross-sectional area and density of paraspinal muscles. Interobserver reliability was not investigated in the current study.

**Summary of Background Data.** Computer tomography (CT) had been used to measure the cross-sectional area and degeneration of the back muscles in patients with low back pain.

**Methods.** This study included 31 patients, mean age 47 years, with chronic low back pain. The measurements comprised cross-sectional area ($cm^2$) and density (Hounsfield units [HU]) of the paraspinal muscles at Th12–L1, L3–L4, and L4–L5. To measure the reliability of the equipment and the observer (total reliability), two independent CT scans were performed for each patient. The radiologist traced the cross-sectional area twice within 2 weeks for measurement of the intraobserver reliability.

**Results.** There were no significant differences in the assessments between the first and second CT scans, or between the radiologist's two measurements of the identical slices. The critical difference for the total reliability ranged from 11.3 to 22.8 for the density and from 10.0 to 16.0 for the cross-sectional area. For the cross-sectional area, the measurement error associated with the observer was higher than for the equipment. For the density, the measurement error related to the equipment was higher. The main measurement error was associated with the radiologist for the cross-sectional area and with the CT scanner for the density.

**Conclusions.** The reliability of the CT scan for measuring the cross-sectional area and density of the back muscles is acceptable. The authors do not know definitely whether their results can be generalized because the interobserver and intermachine reliabilities were not investigated. [Key words: cross-sectional area, density, muscle strength, paraspinal muscle, reliability] **Spine 2003;28:1455–1460**

From the *Department of Physical Medicine and Rehabilitation, Ullevaal University Hospital, Oslo, and the Departments of †Radiology and ‡Orthopaedics, Rikshospitalet University Hospital, Oslo, Norway.
Acknowledgment date: August 28, 2002. First revision date: October 14, 2002. Second revision date: December 2, 2002. Acceptance date: December 10, 2002.
The submitted manuscript does not contain information about medical devices or drugs.
No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this article.
Address correspondence and reprint requests to Anne Keller, MD, Department of Physical Medicine and Rehabilitation, Ullevaal University Hospital, N-0407 Oslo, Norway; E-mail: Anne.Keller@ioks.uio.no

The cross-sectional area and density of the paraspinal muscles are affected by several variables such as age, physical condition, diet, weight, and back pain. With increasing age, there is a reduction in muscle volume.[1–3] Current evidence suggests that the paraspinal muscles are smaller in patients with chronic low back pain[4–8] than in healthy individuals of similar age. Computed tomography (CT) and magnetic resonance imaging (MRI) have been used to measure the cross-sectional area and degeneration of muscles for evaluation of patients with muscular dystrophy.[9,10] These methods also have been applied to assess the effect of exercise on the muscles of the extremities[11–13] and the back.[14–24]

Errors in such measurements are associated with either the observer (the radiologist) or the equipment. The radiologist traces the borders of the back muscles, then uses a computer program to determine the muscle area and density. In patients with a large degree of muscular atrophy and fatty infiltration, the muscle borders are irregular and more difficult to trace precisely. Furthermore, errors connected with the equipment include the position of the subject, the CT scanner, and the technical staff. Measurement error related to the position of the subject is caused by variation of the angle in the sagittal position and the lumbar lordosis. The current study aimed to estimate the errors related to the equipment and to the observer in CT measurements of cross-sectional area and density of paraspinal muscles.

## Methods

**Patients.** The series consisted of 18 men and 13 women. Their mean age was 45 years (range, 21–64 years) and their mean body mass index was 23.3 kg/$m^2$ (range, 18.8–28.9 kg/$m^2$). All the patients were referred to the Department of Orthopedics, Rikshospitalet University Hospital, for diagnostic evaluation and treatment of chronic low back pain. The treatment included back surgery, exercise, or both. Three patients had previously undergone back surgery. Patients with myopathy, muscular dystrophy, inflammatory diseases, spinal deformity, idiopathic scoliosis, previous vertebral fractures, or congenital malformations were excluded from the study. The median duration of the symptoms in the patients was 5 years. The Ethics Committee for Medical Research in Health Region I of Norway approved the study.

**Computed Tomography.** A computer tomography scanner (Toshiba XPEED, Tokyo, Japan) was used to measure muscle area and density. The scanner was calibrated according to air, which is a standard medium, before every examination. The

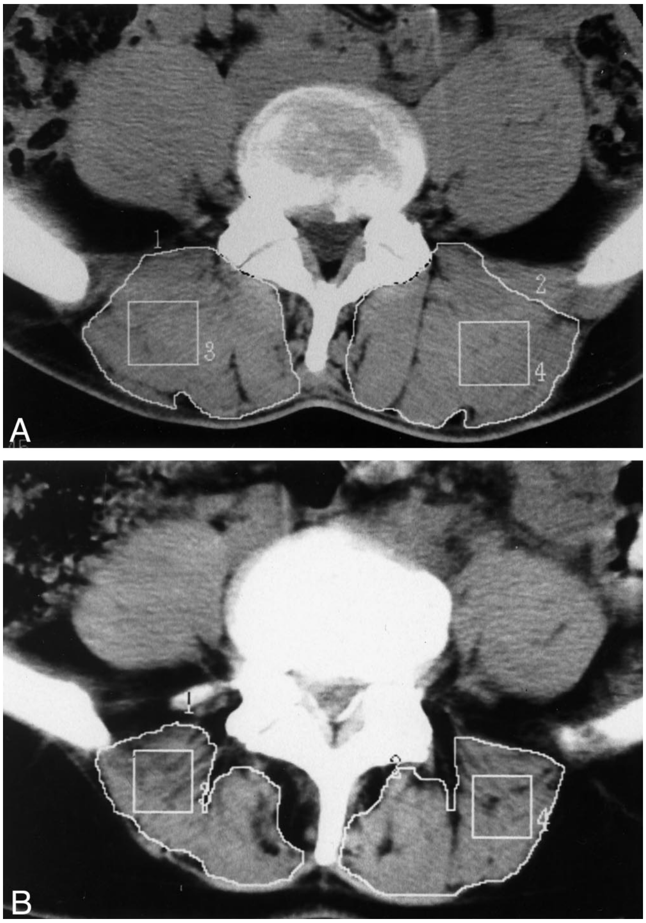


Figure 1. Computed tomography scans at L4–L5 of two patients with low back pain. The upper scan is from a 47-year-old man, and the lower scan is from a 48-year-old woman. The cross-sectional area was calculated within the borders of the back muscles. The muscle density was calculated both within the borders of the back muscles and within the square in a homogeneous part of the muscle.

patients were placed in a standard prone position with a small foam pad under the knees.

Computed tomography scans of muscle area and density were performed two times for each patient. After the first procedure, the patient went off the CT scanner and walked around approximately 5 minutes before the second procedure was performed. Each time, three contiguous 5-mm-thick slices parallel to the disc plane were obtained for each of the disc spaces at Th12–L1, L3–L4, and L4–L5. In the three patients who previously had undergone lumbar instrumental fusion, the slice was obtained at the level above the fixate segment because of the artifacts of the osteosynthesis material, implying loss of three scans at L4–L5.

The measurements comprised both the erector spinae (the longissimus thoracis and iliocostalis lumborum) and the multifidus muscles. An experienced radiologist (R.G.) traced the border of the back muscle to obtain the cross-sectional area, which was measured in square centimeters. The density was measured in Hounsfield units (HU) and evaluated for the whole cross-sectional area in the region of interest. Moreover, density also was measured within a square that the radiologist placed within a homogeneous part of the muscle (Figure 1). The first time, the radiologist determined a homogeneous part of the muscle by judgment. The second time, the judgment was independent of the first because the radiologist did not know the result of the first scan.

For evaluation of observer variance, 44 random slices were selected from eight patients at the three levels. Within a 2-week period, the radiologist made two assessments of the cross-sectional muscle area and the density of these slices. The second judgment was independent of the first. The radiologist had not seen the first CT scanning, and did not know the value of the cross-sectional area or the density from the first scan.

**Statistical Analyses.** The total variance of two consecutive measurements is given as

$$\text{variance}_{\text{total}} = \text{variance}_{\text{observer}} + \text{variance}_{\text{equipment}}$$

and

$$\text{variance}_{\text{equipment}} = \text{variance}_{\text{position}} + \text{variance}_{\text{CT-scanner}} + \text{variance}_{\text{staff}}.$$

The study was designed for direct estimation of $\text{variance}_{\text{total}}$ and $\text{variance}_{\text{observer}}$. Accordingly, it follows that

$$\text{variance}_{\text{equipment}} = \text{variance}_{\text{total}} - \text{variance}_{\text{observer}}.$$

All the groups of paired measurements were examined for statistically significant difference because an analysis of reliability is irrelevant if the first and second measurements differ significantly. The reliability was calculated as the coefficient of variation (CV) in percentage. For each variable in each subject, the CV was calculated as

$$CV = 100\% \cdot \frac{\sqrt{0.5 \cdot d^2}}{\overline{x_{pair}}}$$

and the pooled CV for a group of observations were calculated as described by Gluer *et al*[25]:

$$CV_{pool} = 100\% \cdot \frac{\sqrt{(\Sigma d^2)/2n}}{\overline{x_{pool}}}$$

where n is the number of the subject, d is the difference between the two measurements obtained from one subject, $\overline{x_{pair}}$ is mean of the two measurements, and $\overline{x_{pool}}$ is the mean of all the measurements.

A common term for measurement error is critical difference (CD). If the first of two observations is $X_1$ for a particular measurement and the CD at 0.05 significance level for this method has the value Y, then there is 95% chance that the second observation, $X_2$, lies within the interval from $X_1 - Y \cdot X_1$ to interval from $X_1 + Y \cdot X_1$. In other words, the expected difference between the two measurements would be less than the CD for 95% of the pairs of observations.[26–29] The critical difference at a 0.05 level of significance is defined in percentage terms and can be calculated as

$$CD = CV \cdot Z_{0.05} \cdot \sqrt{2}$$

where $Z_{0.05}$ is the standard deviate for a *P* value of 0.05.

The total measurement error was calculated for each side of the three levels, L4–L5, L3–L4, and Th12-L1, and expressed as CVt and CDt. The measurement error of the observer was calculated at each of the three levels and expressed as CVo and CDo. The measurement error of the equipment was calculated as explained in the preceding equations and expressed as CVe and CDe. The Wilcoxon sign-rank test was performed for paired comparisons. All *P* values less than 0.05 were considered significant. Plots of the diference between tests 1 and 2 against the mean of tests 1 and 2 were constructed according to

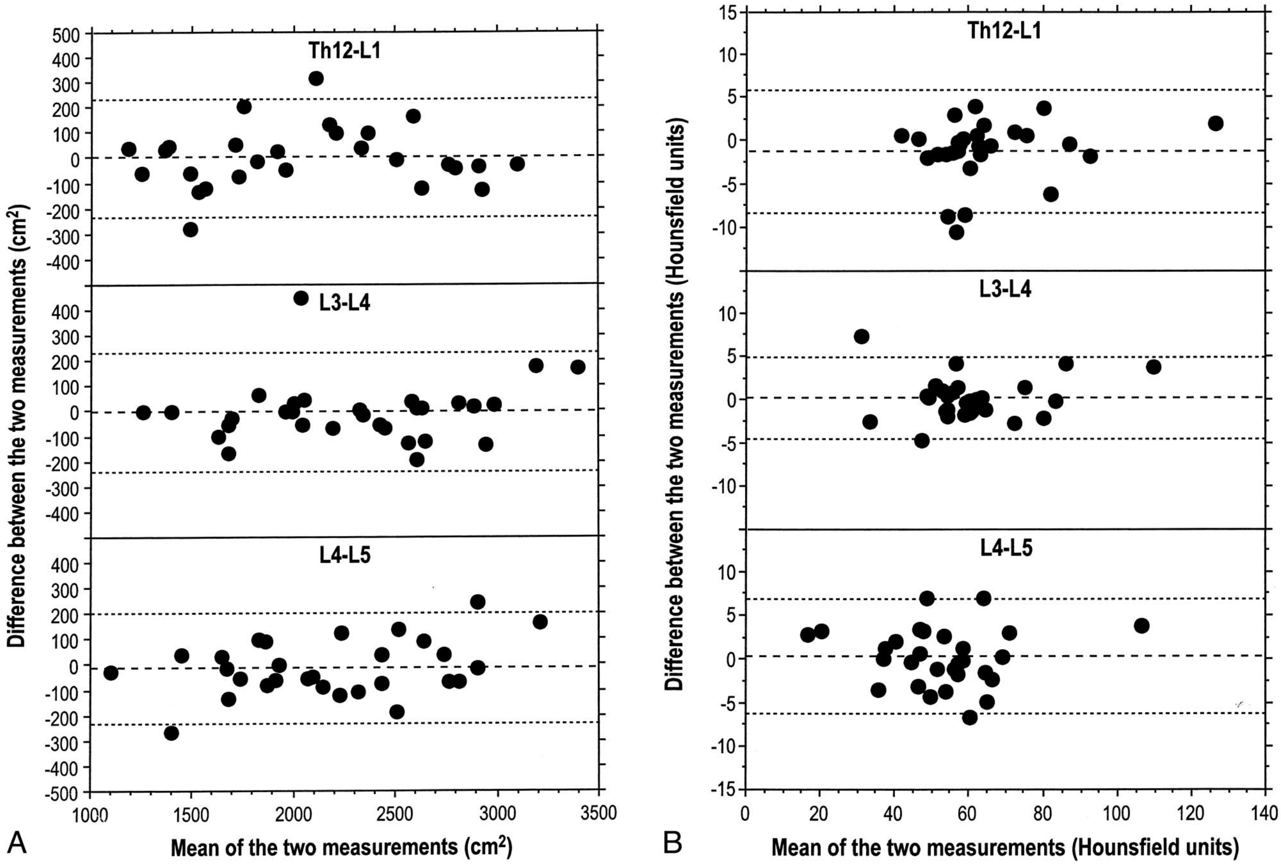


Figure 2. The means of the first and second scans plotted against the difference of the first and second scans for cross-sectional area (A) and total muscle density (B). On each plot, the central horizontal line (– – –) represents the mean of the differences between paired measurements, and the flanking lines (...) represent the 95% limits of agreement. Each point represents the mean of the left and the right side of the back muscles.

Bland's recommendations.[26,30] Software package StatView 5.0.1 was used for the statistical analyses.

## ■ Results

There were no significant differences between the assessments from the first and second CT scans or between the radiologist's two measurements of the identical slices. The measurement error of cross-sectional area and total muscle density are visualized in Figure 2 as plots of the mean versus the difference between the paired measurements.

The results from the first measurement of the cross-sectional area and density are presented in Table 1. The density of the representative part of the muscle was significantly higher than the density of the total cross-sectional area of the muscle ($P < 0.001$), except for L4–L5 (Table 1), but the measurement error for the two methods of density measurement were of the same magnitude (Table 2). Measurement error of the total variance was lower for the cross-sectional area (CDt range, 10.0–16.0) than for the total density and the density of a

**Table 1. Median (Quartiles) for the First Measurement of the Cross-Sectional Area (cm²), Total Density, and Density of a Representative Part of the Lumbar Paraspinal Muscles (Hounsfield units, HU)**

| | | | Density | |
|---|---|---|---|---|
| Spinal Level | Side | Cross-Sectional Area (cm²) | Total Muscle (HU) | Part of the Muscle (HU) |
| Th12–L1 | Right | 19.8 (14.7–25.6) | 58.0 (51.7–64.3) | 63.8* (56.4–71.7) |
| | Left | 20.8 (15.0–24.8) | 55.9 (49.5–62.7) | 59.0# (54.2–66.1) |
| L3–L4 | Right | 22.4 (19.9–26.2) | 53.7 (48.1–61.9) | 57.0* (53.7–64.3) |
| | Left | 24.2 (19.5–27.1) | 54.2 (47.5–62.0) | 59.3* (52.5–63.3) |
| L4–L5 | Right | 20.5 (17.5–25.8) | 49.9 (42.6–55.8) | 51.8 (42.1–58.0) |
| | Left | 21.2 (19.0–28.1) | 51.3 (46.1–59.1) | 52.8 (46.9–62.5) |

# Significant larger than the density of the total muscle, $P < 0.001$.
* Significant larger than the density of the total muscle, $P < 0.0001$.

**Table 2. Coefficients of Variance (CVt) and Critical Differences (CDt) Related the Total Variance of Two Consecutive Measurements at Different Spinal Levels**

| | | Cross-Sectional Area | | Density | | | |
|---|---|---|---|---|---|---|---|
| | | | | Total Muscle | | Part of the Muscle | |
| Spinal Level | Side | CVt (%) | CDt (%) | CVt (%) | CDt (%) | CVt (%) | CDt (%) |
| Th12–L1 | Right | 4.0 | 10.1 | 4.0 | 11.3 | 4.6 | 12.8 |
| | Left | 5.3 | 14.8 | 6.9 | 19.1 | 6.5 | 18.0 |
| L3–L4 | Right | 3.6 | 10.0 | 5.0 | 14.4 | 6.2 | 17.2 |
| | Left | 5.8 | 16.0 | 5.0 | 14.4 | 5.2 | 14.3 |
| L4–L5 | Right | 4.4 | 12.2 | 7.1 | 19.8 | 8.3 | 22.8 |
| | Left | 4.3 | 11.8 | 6.7 | 18.7 | 7.2 | 19.9 |

part of the muscle (CDt range, 11.3–22.8, Table 2). For the density, the measurement error of the total variance was highest at L4–L5, but the measurement errors for the cross-sectional area did not differ across spinal levels (Table 2).

When the measurement error for the radiologist was compared with the measurement error related to the equipment, the error for the cross-sectional area was higher for the radiologist (CDo range, 5.5–16.0) than for the equipment (CDe range, 3.9–8.0), as opposed to density, in which the measurement error was higher for the equipment (CDe range, 8.3–15.8) than for the observer (CDo range, 1.7–4.2) (Table 3).

## ■ Discussion

In the current study, the mean densities of the total cross-sectional area and a representative part of the muscle were measured. The last method was used because tracing the muscle's border could mistakenly include bone or fat, which would give a misleading high or low density. However, no difference in the reliability of these methods was found. The density based on the total cross-sectional area represents a more correct estimate of the density, and is assessed automatically by the computer, so the authors recommend this method.

Although several authors have used CT and MRI for measurement of muscle cross-sectional area,[14–24] the measurement error of the equipment and the radiologist have been investigated in only a few studies. Maughan *et al*[17,31] and Häggmark *et al*[32] evaluated the measurement error for the cross-sectional area of the thigh muscles in young individuals. The CV for the observer ranged from 1.2 to 2.4, as compared with 2.0 to 5.8 in the current study. The regular circumferences of thigh muscles in young persons compared with the back muscles in middle-age persons could account for this difference.

In a previous study, the authors determined the measurement error for a back muscle strength test and the Åstrand bicycle test. The CD was approximately 30% for the back muscle strength test, whereas the CD for the Åstrand test was only about 20%.[33] In the current study, the CDs ranged from 8% to 20% for the cross-sectional area and density. An obvious explanation for these differences is different demands for the patients. The back muscle strength test requires much more motivation of the patients with chronic low back pain than the Åstrand bicycle test and an examination by CT-scanning. The Åstrand test is considered as a highly reliable test,[34–36] and as seen in this context, the reliability for cross-sectional area and density is acceptable.

After separation of the total variance into the variances for the radiologist and the equipment, the findings showed that the measurement error related to the radiologist was high for the cross-sectional area and lower for density, as compared with the error related to the equipment. It is not surprising that the main measurement error for the cross-sectional area is related to the radiologist, who traces the muscle borders. Thus, it seems that the main methodologic error for density is

**Table 3. Coefficients of Variance and Critical Differences Associated With the Equipment (CVe and CDe) and the Observer (CVo and CDo) at Different Spinal Levels**

| | | Equipment-Related | | Observer-Related | |
|---|---|---|---|---|---|
| Spinal Level | Type of Measurement | CVe (%) | CDe (%) | CVo (%) | CDo (%) |
| Th12–L1 | Cross-sectional area | 1.4 | 3.9 | 5.8 | 16.0 |
| | Density | 5.7 | 15.8 | 0.6 | 1.7 |
| L3–L4 | Cross-sectional area | 2.0 | 5.5 | 2.0 | 5.5 |
| | Density | 3.0 | 8.3 | 1.4 | 3.9 |
| L4–L5 | Cross-sectional area | 2.9 | 8.0 | 2.5 | 7.0 |
| | Density | 3.1 | 8.6 | 1.5 | 4.2 |

related to the equipment, and the error of the cross-sectional area to the radiologist.

Degeneration of the back muscles is more marked at L4–L5 than in the cranial direction,[5,15] and higher measurement error at L4–L5 would be expected for both the cross-sectional area and density. However, at Th12–L1 the measurement error for the cross-sectional area was higher than at the two lower levels (Table 3). Because no obvious explanation has been found for these results, the authors interpret their observations as an example of the measurement error made by the radiologist. The measurement error for density demonstrated a tendency toward an increasingly higher error at the lower spinal levels. Because degeneration of the discs dominates at these levels,[7] they are of particular clinical interest. Therefore, the authors recommend CT scanning at L4–L5 for measuring the cross-sectional area of the paraspinal muscles.

Intraclass correlation (ICC), another statistical method for evaluating the reliability,[37] is calculated by the equation

$$\text{ICC} = (\text{BMS} - \text{WMS})/\text{BMS}$$

where BMS is the between-mean square, and WMS is the within-mean square variation. From this equation it is obvious that ICC increases as BMS increases. Therefore, in the case of a large BMS, the estimate of reliability expressed as ICC may be misleadingly high.[33,38] Because the coefficient of variation and critical difference depend only on the WMS, the authors chose to express the reliability as CD.

A weakness of the current study is that only the intraobserver reliability was studied. Because measurement errors related to different observers and different equipment were not tested, it is not known whether the use of different observers and scanners would lead to a similar degree of reliability. However, the patients were placed in a standard prone position in the scanner, and the CT scanner was calibrated before every examination, which is considered regular standard procedure for every type of CT scanner and will therefore decrease the difference between them. Several variables may contribute to the reliability of CT scans of the back muscles. Interobserver reliability usually is lower than intraobserver reliability because two observers rarely have identical observations. Therefore, the authors believe that differences between radiologists and the use of different CT scanners will increase measurement error. Interobserver error should be assessed in future studies.

### Conclusion

The reliability of CT scanning for measuring the cross-sectional area and density of the back muscles was acceptable for the equipment and the observer. In general, the measurement error related to the equipment was higher than that associated with the radiologist. However, the measurement error of the cross-sectional area depends mainly on the observer, whereas the measurement error of the density depends on the equipment. Evaluating density as the mean value of the total cross-sectional area is recommended. The authors do not know definitely whether their results can be relied upon generally because the interobserver and intermachine reliabilities were not investigated.

## Key Points

- A reliability study was conducted to evaluate the measurement errors related to the equipment and the observer in CT measurements of cross-sectional area and density of back muscles.
- The reliability of cross-sectional area and density measurements of the paraspinal muscles was acceptable.
- The measurement error for the cross-sectional area depends mainly on the observer.
- The measurement error for the density depends mainly on the equipment.

### Acknowledgments

Medical statistician Petter Mowinckel is kindly acknowledged for fruitful discussions regarding the statistical analyses.

## References

1. Grimby G, Saltin B. The ageing muscle. *Clin Physiol* 1983;3:209–18.
2. Lexell J. Human aging, muscle mass, and fiber type composition. *J Gerontol A Biol Sci Med Sci* 1995;50:11–6.
3. 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.
4. Cooper RG, St Clair Forbes W, Jayson MI. Radiographic demonstration of paraspinal muscle wasting in patients with chronic low back pain. *Br J Rheumatol* 1992;31:389–94.
5. Danneels LA, Vanderstraeten GG, Cambier DC, et al. CT imaging of trunk muscles in chronic low back pain patients and healthy control subjects. *Eur Spine J* 2000;9:266–72.
6. Mayer TG, Vanharanta H, Gatchel RJ, et al. Comparison of CT scan muscle measurements and isokinetic trunk strength in postoperative patients. *Spine* 1989;14:33–6.
7. Parkkola R. *MRI of the Lumbar Discs and Paraspinal Muscles: A Study of Patients With Low Back Pain and Healthy Controls* [doctoral thesis]. University of Turku, 1993.
8. Sihvonen T, Herno A, Paljarvi L, et al. Local denervation atrophy of paraspinal muscles in postoperative failed back syndrome. *Spine* 1993;18:575–81.
9. Grindrod S, Tofts P, Edwards R. Investigation of human skeletal muscle structure and composition by x-ray computerised tomography. *Eur J Clin Invest* 1983;13:465–8.
10. Jones DA, Round JM, Edwards RH, et al. Size and composition of the calf and quadriceps muscles in Duchenne muscular dystrophy: a tomographic and histochemical study. *J Neurol Sci* 1983;60:307–22.
11. Akima H, Kubo K, Imai M, et al. Inactivity and muscle: effect of resistance training during bed rest on muscle size in the lower limb. *Acta Physiol Scand* 2001;172:269–78.
12. McCartney N, Hicks AL, Martin J, et al. Long-term resistance training in the elderly: effects on dynamic strength, exercise capacity, muscle, and bone. *J Gerontol A Biol Sci Med Sci* 1995;50:B97–104.
13. Sipila S, Suominen H. Effects of strength and endurance training on thigh and leg muscle mass and composition in elderly women. *J Appl Physiol* 1995;78: 334–40.
14. Akima H, Kuno S, Takahashi H, et al. The use of magnetic resonance images to investigate the influence of recruitment on the relationship between torque and cross-sectional area in human muscle. *Eur J Appl Physiol* 2000;83: 475–80.

15. Alaranta H, Tallroth K, Soukka A, et al. Fat content of lumbar extensor muscles and low back disability: a radiographic and clinical comparison. *J Spinal Disord* 1993;6:137–40.
16. Gibbons LE, Latikka P, Videman T, et al. The association of trunk muscle cross-sectional area and magnetic resonance image parameters with isokinetic and psychophysical lifting strength and static back muscle endurance in men. *J Spinal Disord* 1997;10:398–403.
17. Maughan RJ, Watson JS, Weir J. Strength and cross-sectional area of human skeletal muscle. *J Physiol* (Lond) 1983;338:37–49.
18. Maughan RJ, Watson JS, Weir J. Muscle strength and cross-sectional area in man: a comparison of strength-trained and untrained subjects. *Br J Sports Med* 1984;18:149–57.
19. Maughan RJ, Watson JS, Weir J. The relative proportions of fat, muscle, and bone in the normal human forearm as determined by computed tomography. *Clin Sci* (Colch) 1984;66:683–9.
20. Mooney V, Gulick J, Perlman M, et al. Relationships between myoelectric activity, strength, and MRI of lumbar extensor muscles in back pain patients and normal subjects. *J Spinal Disord* 1997;10:348–56.
21. Parkkola R, Rytokoski U, Kormano M. Magnetic resonance imaging of the discs and trunk muscles in patients with chronic low back pain and healthy control subjects. *Spine* 1993;18:830–6.
22. Poelstra KA, Eijkelkamp MF, Veldhuizen AG. The geometry of the human paraspinal muscles with the aid of three-dimensional computed tomography scans and 3-Space Isotrak. *Spine* 2000;25:2176–9.
23. Reid JG, Costigan PA. Geometry of adult rectus abdominis and erector spinae muscles. *J Orthop Sports Phys Ther* 1985;6:278–80.
24. Salminen JJ, Oksanen A, Maki P, et al. Leisure time physical activity in the young: correlation with low back pain, spinal mobility, and trunk muscle strength in 15-year-old school children. *Int J Sports Med* 1993;14:406–10.
25. Gluer CC, Blake G, Lu Y, et al. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* 1995;5:262–70.
26. Bland J. *An Introduction to Medical Statistics*. 2nd ed. New York: Oxford University Press, 1995:396.
27. Bland JM, Altman DG. Statistics notes: measurement error. *BMJ* 1996; 313:744.
28. Madsen OR. Trunk extensor and flexor strength measured by the Cybex 6000 dynamometer: assessment of short-term and long-term reproducibility of several strength variables. *Spine* 1996;21:2770–6.
29. Vasikaran SD, McCloskey EV, Kahn S, et al. Intraindividual variation in fasting urinary calcium- and hydroxyproline-creatinine ratios measured in metabolic bone clinic patients as both outpatients and inpatients. *Ann Clin Biochem* 1994;31:272–6.
30. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
31. Maughan RJ, Watson JS, Weir J. Relationships between muscle strength and muscle cross-sectional area in male sprinters and endurance runners. *Eur J Appl Physiol Occup Physiol* 1983;50:309–18.
32. Haggmark T, Jansson E, Svane B. Cross-sectional area of the thigh muscle in man measured by computed tomography. *Scand J Clin Lab Invest* 1978;38: 355–60.
33. Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering–Sorensen test, and Astrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. *Spine* 2001;26:771–7.
34. Åstrand PO, Rodahl K. *Textbook of Work Physiology: Physiological Base of Exercise*. New York: McGraw-Hill, 1986.
35. Cox NJ, Hendriks JC, Binkhorst RA, et al. Reproducibility of incremental maximal cycle ergometer tests in patients with mild to moderate obstructive lung diseases. *Lung* 1989;167:129–33.
36. Roclevitch RA. *Cycle Ergometery Accuracy Validation Study Report*. In: Brooks: Department of the Air Force, Headquarters Human System Systems Center, 1998:102–14.
37. Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
38. Altman D, Bland J. Measurement in medicine: the analyses of method comparison studies. *Statistician* 1983;32:307–17.