Volume of ankle dorsiflexors and plantar flexors determined with stereological techniques

Paula Gadeberg, Henning Andersen, and Johannes Jakobsen. Volume of ankle dorsiflexors and plantar flexors determined with stereological techniques. J. Appl. Physiol. 86(5): 1670–1675, 1999.—The validity of the methods used for determination of muscle mass has not been evaluated previously. We determined muscle mass by estimating muscle volume with assumption-free stereological techniques applied to magnetic resonance imaging (MRI) in 18 healthy untrained subjects (6 women, 12 men) aged 41 yr (29–64 yr; median, range). Muscle mass was also estimated by measuring leg circumference and cross-sectional muscle areas (CSA) from MRIs at three predetermined levels. Power (peak torque [PT]) of the ankle dorsiflexors and plantar flexors was estimated by using isokinetic dynamometry. Dorsiflexor volume \( r^2 = 0.76, P < 5 \times 10^{-6} \) and CSA \( r^2 = 0.73, P < 5 \times 10^{-5} \) were related to PT, whereas circumference was not \( r^2 = 0.17, n.s. \). Correspondingly, a relationship to plantar PT was established for plantar flexor volume \( r^2 = 0.69, P < 5 \times 10^{-5} \) and CSA \( r^2 = 0.46, P < 5 \times 10^{-3} \) but not leg circumference \( r^2 = 0.15, n.s. \). SDs of the residuals were smaller for the relationship between dorsiflexor PT and volume than between PT and CSA (0.42 vs. 0.45) for plantar flexors (1.5 vs. 2.0). By using the Cavalieri method, six MRI sections and 15 min of point counting are sufficient to obtain a valid estimate of the volume of the muscles of the lower leg.

It is generally accepted that a close relationship exists between the size of the muscle and its ability to generate force. However, there are conflicting results concerning this relationship, probably because of differences in the techniques applied for estimating muscle size. Ikai and Fukunaga (16) found that arm strength was proportional to the cross-sectional area (CSA) of the flexors of the upper arm measured by ultrasound scanning (UL) regardless of age and gender. Maughan et al. (20, 21) found a positive relationship between strength and CSA measured along the leg; however, actual values were not given, and correlations between muscle volume and strength were not calculated. In a study by Roberts et al. (27), volume of all muscles was assessed and the coefficient of error related to sampling was described, but the relationship to muscle function was not studied.

In the present study, three techniques for evaluation of muscle mass of the ankle dorsiflexors and plantar flexors and their relationship to power were compared: 1) leg circumference, 2) CSA at predetermined levels obtained from MRI, and 3) volume assessed by stereological techniques applied to MRI. To study the relationship between muscle function and the various estimates of muscle size, maximal muscle power was assessed by applying isokinetic dynamometry.

Materials and Methods

Subjects. Eighteen sedentary and untrained volunteers (6 women, 12 men) participated in the study. The men, aged 37 yr (29–64 yr; median, range), had a body weight of 75 kg (64–98 kg) and a height of 180 cm (170–194 cm). The women, aged 44 yr (31–56 yr), had a body weight of 60 kg (50–77 kg) and a height of 165 cm (162–170 cm). The length of the lower leg, defined as the distance from the distal end of the lateral malleolus to the lateral articular cleft between the femur and tibia condyles, was 45 cm (42–50 cm) for the men and 40 cm (37–42 cm) for the women. None suffered from neuromuscular diseases or from any other medical or psychiatric disorders. All subjects gave informed consent to the study approved by the local ethics committee.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
MRI. The nondominant leg was evaluated at MRI with a 1.5-Tesla superconducting magnet (Gyrosan, Phillips, Eindhoven, The Netherlands). Scanning was performed with the subject in the supine position on a couch placed inside the gantry. Five-millimeter-thick MRI sections were obtained with a conventional T1-weighted Spin-Echo technique (echo time = 20 ms, repetition time = 550 ms). A 256 × 256 matrix and two excitations were used. MRI images were converted to bitmap files for analysis on a personal computer (Chameleon, Olympus, Ballerup, Denmark). Image resolution allowed unambiguous separation of the various tissues and of individual muscle groups. Muscular anatomy was defined in accordance with standard anatomy textbooks (7). Tibialis anterior, extensor digitorum longus, extensor hallucis longus, and peroneus tertius muscles were defined as the ankle dorsiflexors. Soleus, plantaris, medial and lateral gastrocnemius, flexor hallucis longus, flexor digitorum longus, tibialis posterior, and peroneus longus and peroneus brevis muscles were defined as the ankle plantar flexors. The identity of MRI images was blinded to the observers.

Cross-sectional muscle areas and external circumference at predetermined levels. Cross-sectional MRI scans were performed at levels 20, 50, and 80% of the distance from the lateral malleolus to the articular cleft between the tibial and femoral condyles. These levels were designated as distal, mid-, and proximal level, respectively, of the lower leg. External circumference was measured at the same levels.

Muscle volume estimation. For muscle volume estimation, the lower leg was intersected by a series of transverse and parallel MRI sections. Because the field of view of the MRI scanning was smaller than the lower leg, the subjects were moved, by computer-controlled movement of the couch, once during each imaging performance. Each examination consisted of two scanning sessions of eight sections, excluding two that were overlapping. Thus 14 systematic sections with a random start were performed. The first and last sections were placed outside the lower leg. Twelve sections were used for analysis. At each level, the CSA of the muscle groups of interest was estimated by a single observer by using stereological techniques with separation of contractile muscle from any tissues of interest in the serial slices throughout the leg, according to the Cavalieri principle (9). The empirical resampling method is reliable but time consuming. It is obtained by resampling systematically all possible data sets, e.g., all four sets consisting of every fourth section from the total data set. The standard deviation (SD) divided by the average of the four estimates gives the CE for that particular data set. The average of the estimates for the possible samples coincides exactly with the estimate obtained by using all sections, reflecting the unbiasedness of the Cavalieri estimator. Because the cross sections along the leg are neither independent nor identically distributed, the usual way of calculating the error of the estimates does not apply. Instead, a new error-prediction formula proposed by Matheron, CE (2/P) = (VarSRS (2a) + Noise)2/P, enables a prediction of the error of the Cavalieri estimator from a single systematic sample of n sections chosen at random from the whole set of sections. Matheron’s prediction formula for calculating CE (9, 10) is a reasonable prediction for the majority of objects. VarSRS (2a) is the variance (Var) of the sum of the areas under SRS for a given direction of sectioning, which is dependent on the random position of the first section. Noise is the variance due to the number of points counted over the object of interest, which indicates how much the estimate would change if the transparent test grid had been positioned differently. 2P is the actual number of points counted. Estimation of CE (2/P) is illustrated with a set of data in Table 1. A factor related to object shape for the calculation of the noise effect can be interpolated from the normogram by Gundersen and Jensen (9).

The purpose is to keep the variation introduced by noise, sampling, counting, and measuring procedures (CE) lower than the biological variation (CVbiol) of our material, where CV is coefficient of variation. The total variation (CVtot) can be expressed as $CV_{tot} = CV_{biol}^2 + CE^2$. For all patients, $CV_{tot}$ is the SD between patients divided by the mean of means (SD/Mean). CE is determined as described above, and $CV_{bid}$ is then calculated from the formula. If CE among patients is higher than $CV_{bid}$, the variation introduced by the method is too high and the method should be improved.

Isokinetic muscle testing. The maximal isokinetic muscle power (PT) of the ankle dorsiflexors and plantar flexors was
Table 1. Expected coefficient of error of the estimated volume

<table>
<thead>
<tr>
<th>i (n = 8)</th>
<th>a_i</th>
<th>a_i x a_i</th>
<th>a_i x a_{i+1}</th>
<th>a_i x a_{i+2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>900</td>
<td>1,350</td>
<td>1,680</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>2,025</td>
<td>2,520</td>
<td>4,060</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>3,136</td>
<td>3,808</td>
<td>1,376</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>4,624</td>
<td>2,108</td>
<td>580</td>
</tr>
<tr>
<td>5</td>
<td>31</td>
<td>961</td>
<td>310</td>
<td>155</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>100</td>
<td>50</td>
<td>230</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>25</td>
<td>115</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>529</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

ΣP = 268
A = 12,300
B = 10,261
C = 7,541

evaluated with an isokinetic dynamometer (LidoActive Multi-joint II, Loredan Biomedical, West Sacramento, CA). The dynamometer was calibrated in accordance with recommendations from Loredan. The computer software package Lidoact 5.3D was used for data collection.

The non-dominant leg was tested. The dominant leg was the leg preferred when the subject kicked a ball. Before the tests, subjects received instructions about the procedures and were asked to perform a warm-up of at least five submaximal repetitions with increasing power to become familiarized with the instrumentation. Before each test, a passive-movement sequence provided by the Lidoact software was used to weigh the limb throughout the defined range of motion. Force measurements were automatically corrected for the limb weight. The subjects were instructed to push and pull “as hard and fast as possible” at every trial through the full range of motion available. To ensure standardized instructions, the verbal instructions of one of the examiners were tape recorded and used during all examinations. Every test included eight reciprocal trials with maximal effort. A 10-s rest period was interposed between every trial. Data were accepted if the CV for torque values throughout the movement of the eight repetitions did not exceed 10%. This was done to exclude results obtained from subjects who did not show maximal effort. If the CV exceeded 10%, the person was retested once. In case the CV exceeded 10% in the second test, data were discarded if no outlier torque curve could be identified. Subjects were in a sitting position, with 70 and 80° flexion at the knee and hip joint, respectively, as measured with a hand-held goniometer. The anatomic axis, defined as a line perpendicular to a point just distal to the midpoint of the lateral malleolus, was aligned with the axis of the dynamometer. The foot was placed on the foot plate and secured by two straps placed over the dorsum. In addition, straps were placed around the pelvis and trunk, and a thigh cuff was placed just above the knee for further stabilization. The full range of motion was ±24° from the neutral position in plantar and dorsal direction, the velocity being 60°/s.

Statistical analyses. Comparisons of intrinsic muscle power between dorsiflexors and plantar flexors, as well as between men and women, were performed by applying unpaired t-tests. Relationships between the different measurements of muscle mass and PT were evaluated with linear regression analysis by using a 5% limit of statistical significance. The closeness of the relationships between power and the various measurements of muscle mass was evaluated by calculation of the SD of the residuals. Furthermore, multiple regression analysis was performed, including CSA and volume as the explanatory variables and muscle power as the dependent variable.

RESULTS

The relationship between precision of the estimates of muscle volume and number of sections used is illustrated in Fig. 2. When all possible data sets from the total number of sections are resampled, the empirical CE is lower for a given data set than the CE calculated by using Matheron’s prediction formula (Fig. 2). If all sections are used, i.e., no sampling is performed, the CE of the estimate is 1.5%. Sampling every second section in which muscle tissues are present provides an estimate of muscle volume with a CE of 6%. Sampling of every fourth, sixth, or eighth section provides estimates with CEs that vary between 20 and 30% (Fig. 2). CVbig between patients is 17%.

The sum of the external leg circumference at the three predetermined levels was 88 ± 6 cm. The CSA at the same levels was 27 ± 6 cm² for the dorsiflexors and 115 ± 25 cm² for the plantar flexors. Total volume of the dorsiflexors and plantar flexors was 363 ± 97 and 1,607 ± 396 cm³, respectively. The PT of dorsiflexion was 32 ± 8 Nm and that of plantar flexion was 117 ± 26 Nm. Intrinsic muscle power, defined as PT per unit muscle volume, was higher for the dorsiflexors than for the plantar flexors, 88 ± 12 and 73 ± 10 Nm/dm³, respectively (P < 5 x 10⁻⁴). In a comparison of men and women, no difference was observed for intrinsic muscle power of the dorsiflexors [86 ± 14 vs. 91 ± 6 Nm/dm³].
not significant (NS)] and plantar flexors (74 ± 10 vs. 72 ± 9 Nm/dm³, NS).

By univariate analysis, significant correlations were found between dorsiflexion PT and volume ($r^2 = 0.76$, $P < 5 \times 10^{-6}$) (Fig. 3) as well as CSA ($r^2 = 0.73$, $P < 5 \times 10^{-3}$), whereas circumference was not significantly related to PT ($r^2 = 0.17$, NS).

Correspondingly, plantar flexion PT was related to volume ($r^2 = 0.69$, $P < 5 \times 10^{-3}$) (Fig. 3) and CSA ($r^2 = 0.46$, $P < 5 \times 10^{-3}$). In contrast, leg circumference was not related to PT ($r^2 = 0.15$, NS).

The SD of the residuals for the relationship between volume and dorsiflexion PT was 0.42, smaller than that between CSA and dorsiflexion PT (1.5 vs. 2). Eventually, the SD of the residuals for the relationship between plantar flexion PT and volume was smaller than that between CSA and PT (1.5 vs. 2). Correspondingly, plantar flexion PT was related to volume ($r^2 = 0.69$, $P < 5 \times 10^{-3}$) (Fig. 3) as well as CSA ($r^2 = 0.46$, $P < 5 \times 10^{-3}$). In contrast, leg circumference was not related to PT ($r^2 = 0.15$, NS).

DISCUSSION

Leg muscle size has been assessed by a number of methods. Measurements of skinfolds and leg circumference are inexpensive and noninvasive and do not require sophisticated equipment (11). They are inaccurate, however, because all muscles and nonmuscular tissues that make no contribution to strength are included. The procedure leads to an overestimation of CSA of −20–25% (15). Ultrasound has been frequently used (6, 12, 13), although it provides images with low spatial resolution, impeding separation of neighboring muscles. The technique is also suboptimal, because identical cross-sectional planes in all subjects are difficult to define. CT has enabled acquisition of cross-sectional images with higher resolution and, consequently, has been used in numerous studies for determination of muscle CSA in humans (4, 19, 22, 23). However, the power to differentiate between tissues is not optimal, and, furthermore, CT is hazardous to the subjects because it introduces ionizing radiation. The introduction of MRI has enabled a more reliable separation of tissues and individual muscles (5, 8, 25), without introducing any known risk to the subjects. Stereology applied to MRI allows noninvasive and unbiased in vivo measurements of the amount of contractile muscle tissue (27).

The regular form of the leg muscles is ideal for systematic sampling (9), thereby requiring small sample sizes only. For a stereological method to be efficient, CE has to be lower than CVbiol. The smallest number of sections needed to fulfill this criterion is the optimal sample. CE for a particular data sample can be predicted by using Matheron’s prediction formula or the empirical resampling method described in Coefficient of error (CE). Matheron’s formula is not unbiased, because it is based on a mathematical model related to object shape. It is applicable to most objects and requires the use of at least three sections. The empirical procedure is simple to perform. However, its reliability decreases as the size of the samples approaches the size of the data set itself, because the available sample replications decrease. The efficiency of the Cavalieri estimator is emphasized by the fact that, when both Matheron’s formula and the empirical procedure for calculating CE are used, the error of the Cavalieri estimator decreases as the number of sections increases (1/n), which is faster than if sections were independent, in which case the rate of decrease would only be $1/\sqrt{n}$ (9). In our study, using a sample of every fourth section provides a CE of 20%, higher than the CVbiol among patients, i.e., 17%. Increasing the sample to every second section provides a CE of 6%, which is considerably lower than CVbiol. Because the Cavalieri estimator is more sensitive to changes in the number of sections used than to the total number of test points counted, counting >100–200 points will not add any significant precision to the required volume estimate.

According to the Cavalieri method, the following conditions have to be met to ensure unbiasedness: 1) a random start of the series of sections, 2) scan of the
whole object of interest, and 3) random placement of the transparent test grid on the images for area estimation. The first two requirements can be ensured by placing the first and last section outside the object. Failure to meet these conditions will cause bias, which cannot be corrected (9).

From a physiological point of view, CSA of all fibers at right angles to their long axis is related to the amount of tension that a muscle can produce (14). The anatomic CSA corresponds to the physiological CSA in muscles with parallel fibers, because all sections cut the fibers at right angles. In pinnate muscles, however, the anatomic CSA will cut a limited number of fibers and thereby not correspond to the physiological CSA (8, 24). Furthermore, CSA differs with respect to a given proximo-distal location along the leg and, therefore, does not necessarily correspond to strength. Estimation of muscle volume is a way of overcoming these problems.

In the present study, we found a stronger relationship between power and volume compared with CSA and circumference for both muscle groups. The relationship between volume and power of the plantar flexors but not the dorsiflexors has been studied by Alway et al. (1). The measurements were biased to an unknown degree, because assumptions were made about the form of the muscles in the segments between MRI scans. Furthermore, total volume was not estimated because only three regions of the plantar flexors were evaluated. Correlations between volume and power were not made, and direct comparison between CSA and volume as measurements of muscle mass was not performed.

Although close relationships could be established between size and power of the muscle groups, in our study, some variation in muscle power remains unexplained. One reason for this could be the variability introduced by determining maximal voluntary muscle power. Assessment of voluntary power requires full cooperation by all participating subjects. If subjects perform submaximally or their positioning changes, variation in power unrelated to muscle size is introduced. The contribution of such variation in the present study cannot be determined quantitatively; however, in our laboratory standardization of the examination techniques has resulted in high reproducibility (2). This includes standardized instructions before and during testing. In addition, after testing, the CV for torque values throughout the movement of the eight repetitions is calculated. If the CV exceeds 10%, submaximal performance is suspected and the subject is retested. Large variation in torque at isokinetic dynamometry has also been used to detect hysterical paresis (17). Variation in voluntary muscle power may also be explained by inability to fully activate the muscles. Using the twitch-interpolation technique, Belanger and McComas (3) observed that not all healthy subjects were able to activate their plantar flexor motor units fully, whereas this was not the case for the tibialis anterior muscle. This difference may explain our finding of a weaker correlation between power and size for the plantar flexors compared with the dorsiflexors.

Another explanation may be a larger interindividual variation in fiber orientation and muscle fiber composition of the plantar flexors compared with the dorsiflexors. Differences in intrinsic power in relation to gender may also contribute to the residual variation in the volume-power relationship. In the present study, however, we did not observe any difference in intrinsic muscle power between men and women. This is in accordance with other studies (16, 28, 29), supporting the notion that differences in power between genders primarily is due to quantitative and not qualitative differences in muscles.

In conclusion, volume of the ankle dorsiflexors and plantar flexors can be estimated with high precision and little effort, by applying stereological methods and MRI. Maximal muscle power was more closely related to volume than CSA or circumference. In future studies relating muscle size and function, we suggest that muscle volume should be estimated by applying stereological techniques to MRI.

R. Sangill is acknowledged for excellent technical assistance.

Address for reprint requests and other correspondence: P. Gadeberg, Dept. of Neurology, Aarhus Univ. Hospital, Nørrebrogade 44, 8000 Aarhus C, Denmark (E-mail: paula@akhphd.au.dk).

Received 1 July 1998; accepted in final form 12 January 1999.

REFERENCES


