MAGMA (1997) **5**, 93–98

Volumetric measurement of human calf muscle from magnetic resonance imaging

M.A. Elliott,[1]* G.A. Walter,[1] H. Gulish,[2] A.S. Sadi,[2] D.D. Lawson,[2] W. Jaffe,[2] E.K. Insko,[1] J.S. Leigh,[1] and K. Vandenborne[2]

Departments of Radiology[1] and Rehabilitation Medicine,[2] University of Pennsylvania, Philadelphia, Pennsylvania, USA

Muscle mass is a determining factor in skeletal muscle function and is affected by inactivity, immobilization, disease, and aging. The aim of this study was to develop an objective and time-efficient method to quantify the volume and cross-sectional area of human calf muscles using three-dimensional magnetic resonance images. We have estimated the errors incurred in muscle volume measurements arising from artifacts known to occur in magnetic resonance imaging (MRI). The largest source of error was due to partial volume effects, which resulted in overestimation of phantom volumes ranging from 145 to 900 cc by 6% to 13%. The magnitude of this effect has been shown to increase with decreasing object size and decreasing spatial resolution. We have presented a straightforward correction for this effect, which has reduced the volume measurement error to less than 4% for all cases. Through the use of computer simulations, the correction algorithm has been shown to be independent of object shape and orientation. To reduce user subjectivity, a semiautomated computer program has been developed to segment MRI data for particular muscle groups. Images from seven human subjects were analyzed by the program, yielding muscle volumes of 154.2 ± 23.2, 281.2 ± 35.8, and 432.2 ± 83.7 for the lateral gastrocnemius, medial gastrocnemius, and soleus, respectively.

Keywords: segmentation, image processing software, computer assisted, correction, cross-sectional area, atrophy.

INTRODUCTION

Muscle atrophy is a common clinical phenomenon observed as a result of inactivity or disuse [1,2]. Short-term muscle atrophy is observed after immobilization, bed rest, a change in lifestyle, or a specific disease process. One of the most important causes of atrophy, however, is aging. Over the age of 60, muscle mass decreases at the rate of approximately 1% per year in both men and women [3–5]. Quantitative determination of muscle volume and cross-sectional area can prove useful in monitoring the progression of atrophy and the therapeutic effects of intervention. The aim of this study was to develop an objective and time-efficient method for quantification of the volume and cross-sectional area of human calf muscles, specifically the medial gastrocnemius, lateral gastrocnemius, and soleus.

Magnetic resonance imaging is the method of choice for *in vivo* volume measurements because of its high sensitivity to soft tissue and its noninvasive nature. For the measurement of volumes with adequate precision, high-resolution image data sets are needed. Moreover, Gibbs et al. [6] have shown that three-dimensional (3D) imaging schemes allow for more accurate volume determination than do two-dimensional (2D), multislice experiments. As a result, high-resolution proton MRI over volumes the size of the human calf will typically consist of several million data points. Fortunately, the past few years have seen the development of fast imaging sequences [7] that allow acquisition of such large data sets in reasonable amounts of time. Another experimental constraint for determination of muscle volumes is the need to distinguish muscle tissue from nonmuscle tissue, primar-

* *To whom all correspondence should be addressed, at B1 Stellar-Chance Labs, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA.*
Date received: January 10, 1997; date accepted: January 18, 1997.

ily fat. A variety of techniques exist to generate such contrast in MRI, including T_2 weighting, inversion recovery, and chemically selective saturation [8,9].

Several artifacts can lead to errors in the estimation of tissue volumes from magnetic resonance imaging data. The authors are aware of three major sources of such artifacts: (1) Gibbs ringing due to data truncation, (2) nonlinear or poorly calibrated magnetic field gradients, and (3) partial volume filling of surface voxels. For accurate volume estimates from 3D MRI data, each of these sources of error has been examined. In the case of partial volume effects, the degree of error was found to be unacceptably high, and a correction method is described.

To obtain volumetric information on human skeletal muscle, we have developed an interactive computer program, EXTRACTOR, which is used to segment three-dimensional MRI data sets according to distinct muscle groups. The program permits the interactive selection of standard image-processing operations, including pixel intensity thresholding, seed growing, boundary tracing, and morphological erosion and dilation [10]. Segmentation of a data set can be performed either as a sequence of 2D operations on each slice or with 3D operations on the entire volume or specified subvolume.

METHODS

Human subjects

MRIs were performed on seven human subjects, four male and three female. The subjects ranged from 19 to 50 years of age, with a mean age of 32.14. Images of the right leg extending from the malleolus to midthigh were obtained as described below.

Phantoms

For validation of the accuracy of our volumetric measurements, six roughly spherical phantoms consisting of balloons filled with tap water were imaged with the protocol described below. The phantom volumes ranged from 145 to 900 cc, which approximately spans the range of the muscles of the human calf [11]. The experiments were repeated for different orientations of the phantoms relative to the radiofrequency (RF) coil and the main magnetic field.

MRI data acquisition

Proton density magnetic resonance images of the human subjects and the phantoms were obtained with a 3D fast-gradient echo sequence on a 1.5T GE Signa (General Electric Medical Systems, Milwaukee, WI) whole-body system. The data were acquired with a birdcage extremity coil with TR = 100 ms, TE = 10 ms, and a flip angle of 30°. The images were acquired in three sets, each with an encoding matrix of 256 × 256 × 28 and a field of view (FOV) of 16 × 16 × 19.6 cm. This yielded a nominal voxel size of 2.73 mm^3. The field of view was chosen using a coronal localizing image. Each acquisition took 9 minutes, resulting in a total experiment time of approximately 40 minutes.

Given the large data sets involved, the lengthy acquisition schemes of T_2-weighted and inversion recovery imaging make those techniques unfeasible. Consequently, chemically selective fat suppression, which adds very little to the time length of the experiment, was employed to enhance the definition between muscle groups and to exclude infiltrated fat from muscle volume calculations.

Data analysis

A widgetized, user-guided computer program was developed to enable the extraction of individual muscle volumes from the 3D MRI data. The operator is allowed to select from a standard menu of image-processing operations to manipulate a data set. The sequence of operations used in this study proceeds as follows. First, an image intensity threshold is selected to identify pixels that contain signal from muscle tissue and to reject pixels with low signal intensity. Next, a combination of boundary tracing and morphological erosion is performed to isolate the contiguous set of voxels that comprises a particular muscle. At the spatial resolution limits imposed by practical time considerations, we were *not* able to isolate completely the muscles of the lower leg without user-driven boundary tracing. After a muscle has been isolated, a seed-growing algorithm is used to retain only contiguous voxels within a single muscle. Voxel counting, with a calibrated gradient system providing the voxel volume, is used to convert pixel volumes and areas to physical units (e.g., cm^3).

Correction for partial volume effects

Volume estimates obtained by simple counting of all voxels within an object can result in substantial error due to partial filling of surface voxels [12,13]. This arises from the fact that the voxels comprising the surface of an object on average contain less than a full voxel volume worth of the object. Counting surface voxels equally with interior voxels will result in overestimation of the object's total volume. A variety of algorithms can correct for this error [12–16]. We have implemented the following straightforward solution:

1. A morphologic erosion with a cubic erosion kernel 3 voxels wide is performed on the object. This

effectively removes any voxel in the object which neighbors a voxel not in the object. The result is a mask for the interior voxels of the object.
2. The Boolean difference between the original volume and the eroded volume is taken, yielding a mask for the surface voxels of the object.
3. The total volume (in voxel volume units) of the object is computed as: volume = 1.0 · (# of interior voxels) + (filling fraction) · (# of surface voxels).

The "filling fraction" term is the average of the fraction of all surface voxels that is occupied by the object. The pixel intensity threshold used to segment the object determines the value of this term. If a threshold of zero were used (i.e., any voxel containing any amount of the object is included), the filling fraction would be ½ because, on average, each surface voxel is half-filled. Signal-to-noise ratio (SNR) constraints, among others, dictate the need for higher intensity threshold levels. Accordingly, we estimate the filling fraction value to be as follows:

$$\text{filling fraction} = \frac{1}{2}\left\{\frac{(\text{max pixel intensity}) + (\text{threshold pixel intensity})}{\text{max pixel intensity}}\right\}.$$

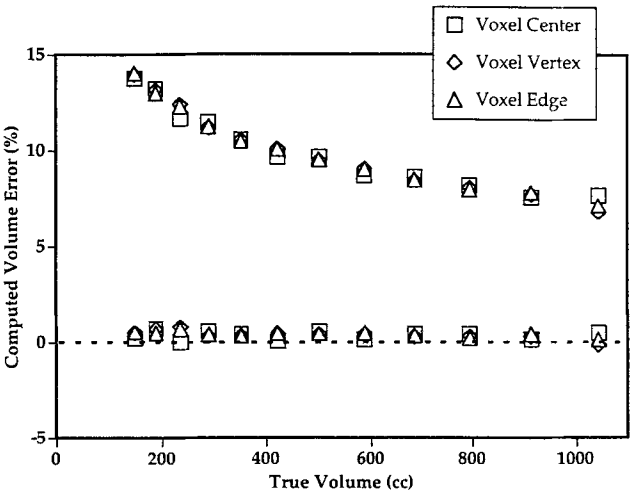

Fig. 1. Computer simulation results for spherical objects imaged with the voxel resolution used in the MRI experiments. Calculations were done for spheres centered at a voxel center, centered at a voxel edge, and centered at a voxel vertex. Computed volume error as a percentage of actual volume is shown for simple voxel counting (*top traces*) and after application of the partial volume correction described in the text (*bottom traces*).

RESULTS

Computer simulation of partial volume correction

The accuracy of volume determination from MRI data was assessed using computer simulations. Spherical objects with volumes ranging from 150 to 1000 cc were simulated when digitized with the resolution described in the MRI methods above. The simulations were repeated for spheres with their centers in three alignments relative to the voxel grid. The object centers (i) coincided with the center of a voxel, (ii) coincided with a voxel vertex, and (iii) coincided with a voxel edge. In all simulations, perfect gradient calibration was assumed (i.e., voxel volumes were precisely known). Figure 1 shows the resulting volume measurements that would be obtained if equally weighted voxel counting were used. All volumes were overestimated, ranging in error from approximately 7% to 15%. Moreover, the error was essentially independent of object position relative to the voxel grid. The degree of overestimation decreased as the *relative* resolution increased (i.e., the number of voxels spanning the object increased). The partial volume correction algorithm detailed above was then applied to the simulation results. With this correction applied all volume estimates were correct to within 1%.

The simulation was repeated for ellipsoidal objects whose major axis to minor axis ratios ranged from 1.5:1 to 2.5:1. The degree of volume overestimation increased with elliptical eccentricity and ranged from approximately 10% to 29%. The partial volume correction algorithm was able to reduce the error to less than 2.5% independent of ellipsoidal shape (see Fig. 2).

Phantom measurements

The phantom MRI data were analyzed with the EXTRACTOR computer program. In good agreement with the simulation results, the uncorrected volume estimates ranged in error from 6% to 13%. The application of the partial volume correction algorithm to the phantom data is shown in Fig. 3. In all cases, the corrected volumes were closer to the true values, with percentage errors reduced by at least a factor of 3. The maximum error of the corrected volumes was 4% for the smallest phantom (145 cc) and was less than 1% for the phantoms with volumes of 345 cc and more.

Human measurements

Trained operators segmented the images for the muscle groups described above. The average time required to segment an individual muscle from all three data sets spanning the calf was approximately 45 minutes. To assess the degree of interoperator variability, different operators processed nine muscles from three subjects (three soleus, three lateral gastrocnemius, and three medial gastrocnemius). The volumes determined by each operator are plotted against each other in Fig. 4. The maximum variation in estimated muscle

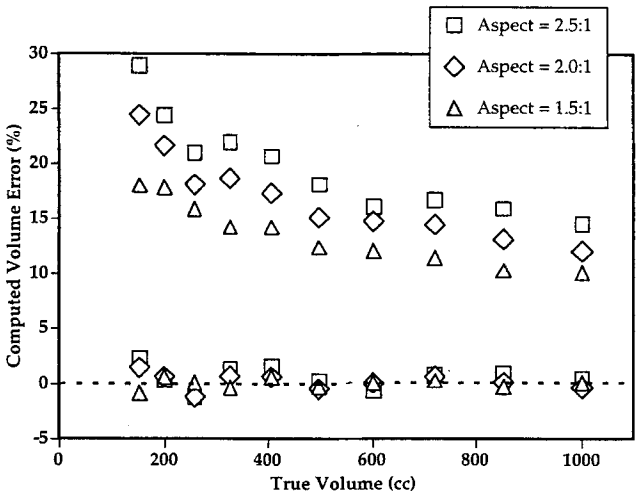

Fig. 2. Computer simulation results for ellipsoidal objects imaged with the voxel resolution used in the MRI experiments. Calculations were performed for ellipsoids with major-to-minor-axis ratios of 2.5:1, 2:1, and 1.5:1. Computed volume error as a percentage of actual volume is shown for simple voxel counting (*top traces*) and after application of the partial volume correction (*bottom traces*).

volume between operators was 4.3%, and the correlation coefficient was 0.996.

Table 1 shows the mean muscle volume of all analyzed subjects broken down by muscle group (n = 7). The average percentage of change in volume estimation due to the partial volume correction was −8.8%, −8.0%, and −7.2% for the lateral gastrocnemius, medial gastrocnemius, and soleus, respectively. In agreement with the phantom results, the percentage of change in estimated volume increased with decreasing object size (Fig. 5).

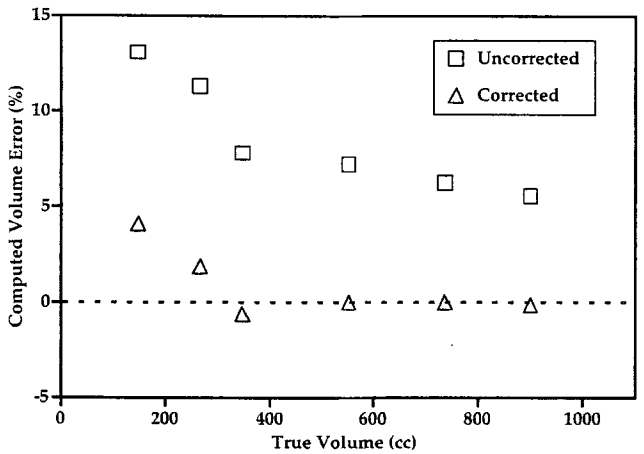

Fig. 3. Volume estimates from the phantom MRI experiments. Volumes were obtained using the EXTRACTOR computer program using simple voxel counting (*top trace*) and after application of partial volume correction (*bottom trace*).

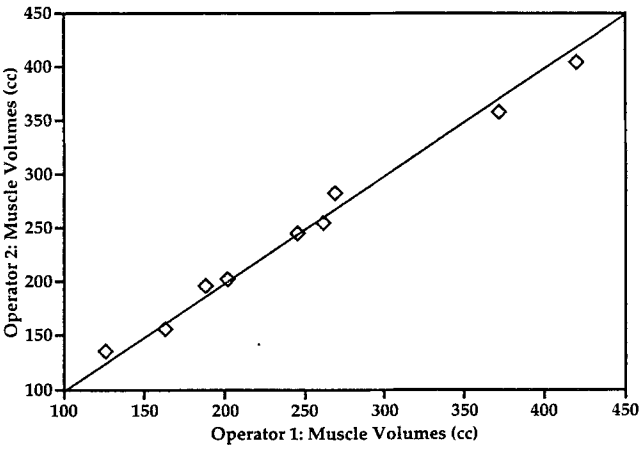

Fig. 4. Comparison of interoperator variability in muscle volume determination. Muscle volumes as determined by different operators are plotted against each other for the lateral gastrocnemius, medial gastrocnemius, and soleus muscles of three subjects.

DISCUSSION

Through the use of computer simulations and phantom experiments, we have shown that simple voxel counting can result in significant errors in volume calculation. The primary source of this error derives from the inclusion of voxels on the surface of an object which are only partially filled with signal. The magnitude of the error increases with decreasing object size because of the greater proportion of all voxels contained in the object which are surface voxels. The effect is minimized for spherical objects, which have the lowest ratio of surface area to volume. As expected, the effect increases for elongated objects. For imaging schemes with unequal voxel dimensions (i.e., noncubic voxels), the effect can depend on object ori-

Table 1. Muscle volumes as determined by simple voxel counting and after partial volume correction for the muscles of the human triceps surae.

Muscle	Voxel counting volume (cc)	Partial volume correction volume (cc)	% Change
LG	169.1 (24.4)	154.2 (23.2)	−8.8
MG	305.7 (37.4)	281.2 (35.8)	−8.0
Soleus	465.7 (87.8)	432.2 (83.7)	−7.2

Abbreviations: LG—lateral gastrocnemius; MG—medial gastrocnemius.
Values are means (SD).

entation, as the number of voxels spanning the object will be different. For minimization of the error in volume estimation, the number of voxels contained by the object should be maximized. In general, this can be accomplished by coinciding the highest resolution dimension with the longest axis of the object.

Our investigations demonstrate that partial volume effects can lead to errors as large as 15% in the volume measurement of the human triceps surae using high-resolution 3D MRI. To counteract this effect, a straightforward correction has been implemented, which is applied postacquisition to the reconstructed MR image. The correction algorithm is robust, reducing the error from 13% to less than 4% in phantom studies, as well as demonstrating insensitivity to object shape in computer simulations. Application of the correction algorithm in human studies has resulted in an average reduction of estimated muscle volumes ranging from 8.8% for the lateral gastrocnemius to 7.2% for the soleus. The magnitude of the volume correction is smaller in the *in vivo* results as compared to the phantom and the simulation results because a higher pixel intensity threshold was used in the *in vivo* experiments. Lower SNR and imperfect suppression of fat necessitated higher threshold values for the *in vivo* data.

Partial volume effects can be quite complex. Hughes et al. [17] found that partial volume filling caused underestimation of volumes from computed tomographic (CT) data. This is not contradictory to our finding, but rather shows the effect when the selected pixel intensity threshold is high. In this case, more than half of an object's surface voxels will fail to be included in the segmented volume, and the volume will be underestimated. Consequently, the partial volume correction described in this communication performs a weighting of surface voxels based on the threshold value employed in the segmentation. In practice, the threshold should be selected as low as possible without including an unacceptable amount of contamination from noise or non–muscle tissue filled voxels.

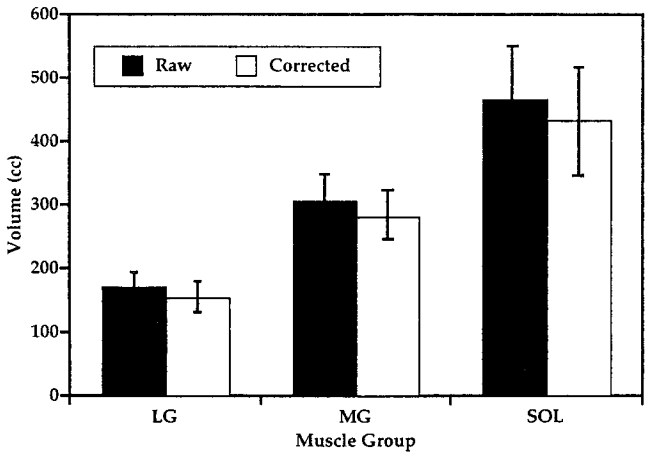

Fig. 5. Average volumes of the lateral gastrocnemius, medial gastrocnemius, and soleus muscles determined by voxel counting (*raw*) and partial volume correction (*corrected*) for $n = 7$ subjects. Error bars show the standard deviation of each population.

User subjectivity is often another major source of precision error in image quantitation. To minimize this effect, we have developed a semiautomated computer program to segment magnetic resonance images. While requiring human input, primarily in the form of boundary tracing, a sufficient amount of automation has been incorporated in the process to reduce user bias in the measurement of human calf muscles to less than 5%.

Calibration of the magnetic field gradients of the MR scanner is a prerequisite for volumetric measurements using MRI. The accuracy with which the nominal dimensions of each voxel are known sets the upper limit on the accuracy with which volumes can be determined. In general, the minimum possible error in a volume measurement will be equal to the error between the expected voxel volume and the actual voxel volume. For example, if a volume is computed assuming a voxel size that differs by 5% from its actual size, then the estimated volume will be in error by at least 5%. This effect will be independent of the size of the object or the image resolution. The MRI scanner used in this study maintains a tolerance of 5% in its gradient calibration. This was deemed acceptable for this study and no correction for this potential source of error was attempted. This decision is supported by the size dependence of the errors in the phantom measurements, which is inconsistent with errors due to gradient miscalibration.

Truncation artifacts due to finite acquisition times can also result in significant errors in Fourier MR imaging [18]. These artifacts appear as "ringing" in the sharp edges of the reconstructed image and can be mistaken for additional regions of signal in otherwise signal-free areas of the image. Consequently, volume estimates from MRI can be further overestimated because of truncation artifacts. The magnitude of this effect increases with decreasing spatial resolution and can often be a problem in the phase-encoding dimensions of magnetic resonance images. Moreover, like partial volume effects, the relative magnitude of the effect increases with decreasing object size. However, at the resolution obtained in this study, truncation artifacts do not appear to result in volume estimation errors of more than about 5% for the smallest volumes

of interest. It seems probable to the authors that the residual error in the volume estimation of the two smallest phantoms (see Fig. 3) is due to truncation artifacts. Correction for this effect is far from trivial and, at present, no correction for this residual error has been undertaken.

The EXTRACTOR computer program was written in the IDL programming language manufactured by Research Systems, Inc., Boulder, Colorado, USA. The program is available free of charge via FTP from *ftp.mmrrcc.upenn.edu* (130.91.196.205) or from the World Wide Web at *http://www.mmrrcc.upenn.edu.*

ACKNOWLEDGMENTS

The authors thank Dr. J. Esterhai and Dr. E. Okereke for their essential contributions to this work. This research was supported by NIH Grants RR02305 and HD33738.

REFERENCES

1. Bassey EJ, Fiatarone MA, O'Neill EF, Kelly M, Evans WJ, Lipsitz LA (1992) Leg extensor power and functional performance in very old men and women. *Clin Sci* **82:** 321–327.
2. Veldhuizen JW, Verstappen FT, Vroemen JP, Kuipers H, Greep JM (1993) Functional and morphological adaptations following four weeks of knee immobilization. *Int J Sports Med* **14:** 283–287.
3. Evans WJ (1995) Effects of exercise on body composition and functional capacity of the elderly. *J Gerontol* **50A:** 147–150.
4. Wolfson L, Judge J, Whipple R, King M (1995) Strength is a major factor in balance, gait, and the occurrence of falls. *J Gerontol* **50A:** 64–67.
5. Lexell J (1995) Human aging, muscle mass, and fiber type composition. *J Gerontol* **50A:** 11–16.
6. Gibbs P, Buckley DL, Blackband SJ, Horsman A (1996) Tumour volume determination from MR images by morphological segmentation. *Phys Med Biol* **41:** 2437–2446.
7. Tkach J, Haacke EM (1988) A comparison of fast spin echo and fast gradient field echo sequences. *Magn Reson Imaging* **6:** 373–389.
8. Phoenix J, Betal D, Roberts N, Helliwell T, Edwards R (1996) Objective quantification of muscle and fat in human dystrophic muscle by magnetic resonance image analysis. *Muscle Nerve* **19:** 302–310.
9. Weinberger E, Shaw D, White K, Winters W, Stark J, Nazar-Stewart V, Hinks R (1995) Nontraumatic pediatric musculoskeletal MR imaging: comparison of conventional and fast-spin-echo short inversion time inversion recovery technique. *Radiology* **194:** 721–726.
10. Clarke LP, Velthuizen RP, Camacho MA, Heine JJ, Vaidyanathan M, Hall LO, Thatcher RW, Silbiger ML (1995) MRI segmentation: methods and applications. *Magn Reson Imaging* **13:** 343–368.
11. Fukunaga T, Roy RR, Shellock FG, Hodgson JA, Edgerton VR (1996) Specific tension of human plantar flexors and dorsiflexors. *J Appl Physiol* **80:** 158–165.
12. Kenedy DN, Filipek PA, Caviness VS (1989) Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging. *IEEE Trans Med Imaging* **8:** 1–7.
13. Johnson LA, Pearlman JD, Miller CA, Young TI, Thulborn KR (1993) MR quantification of cerebral ventricular volume using a semiautomated algorithm. *Am J Neuroradiol* **14:** 1373–1378.
14. Lancaster JL, Eberly D, Alyassin A, Downs JH, Fox PT (1992) A geometric model for measurement of surface distance, surface area, and volume from tomographic images. *Med Phys* **19:** 419–431.
15. Alyassin AM, Lancaster JL, Downs JH, Fox PT (1994) Evaluation of new algorithms for the interactive measurement of surface area and volume. *Med Phys* **21:** 741–752.
16. Hillel PG, Hastings DL (1993) A three-dimensional second-derivative surface-detection algorithm for volume determination on SPECT images. *Phys Med Biol* **38:** 583–600.
17. Hughes SW, D'Arcy TJ, Maxwell DJ, Saunders JE, Ruff CF, Chiu WSC, Sheppard RJ (1996) Application of a new discreet form of gauss' theorem for measuring volume. *Phys Med Biol* **41:** 1809–1821.
18. Wood ML, Henkelman RM (1985) Truncation artifacts in magnetic resonance imaging. *Magn Reson Med* **2:** 517–526.