Original Research

Automatic and Quantitative Assessment of Regional Muscle Volume by Multi-Atlas Segmentation Using Whole-Body Water–Fat MRI

Anette Karlsson, MSC,[1,2]* Johannes Rosander, MSC,[3] Thobias Romu, MSC,[1,2] Joakim Tallberg, Med Student,[2] Anders Grönqvist, MSC,[2,4] Magnus Borga, Prof,[1,2] and Olof Dahlqvist Leinhard, PhD[2,5]

Purpose: To develop and demonstrate a rapid whole-body magnetic resonance imaging (MRI) method for automatic quantification of total and regional skeletal muscle volume.

Materials and Methods: The method was based on a multi-atlas segmentation of intensity corrected water–fat separated image volumes. Automatic lean muscle tissue segmentations were achieved by nonrigid registration of atlas datasets with 10 different manually segmented muscle groups. Ten subjects scanned at 1.5 T and 3.0 T were used as atlases, initial validation and optimization. Further validation used 11 subjects scanned at 3.0 T. The automated and manual segmentations were compared using intraclass correlation, true positive volume fractions, and delta volumes.

Results: For the 1.5 T datasets, the intraclass correlation, true positive volume fractions (mean ± standard deviation, SD), and delta volumes (mean ± SD) were 0.99, 0.91 ± 0.02, −0.10 ± 0.70L (whole body), 0.99, 0.93 ± 0.02, 0.01 ± 0.07L (left anterior thigh), and 0.98, 0.80 ± 0.07, −0.08 ± 0.15L (left abdomen). The corresponding values at 3.0 T were 0.97, 0.92 ± 0.03, −0.17 ± 1.37L (whole body), 0.99, 0.93 ± 0.03, 0.03 ± 0.08L (left anterior thigh), and 0.89, 0.90 ± 0.04, −0.03 ± 0.42L (left abdomen). The validation datasets showed similar results.

Conclusion: The method accurately quantified the whole-body skeletal muscle volume and the volume of separate muscle groups independent of field strength and image resolution.

Key Words: multi-atlas segmentation; muscles; registration; muscle volume; classification; MRI
J. Magn. Reson. Imaging 2014;00:000–000.

ACCURATE AND PRECISE MEASUREMENTS of the muscle volume are important for advancing understanding of different diseases, syndromes, and disorders such as muscular dystrophies (1,2), sport injuries (3), inflammatory myopathies (4), spinal cord injuries (5), or sarcopenia (6–8) (muscle loss due to aging). When diagnosing sarcopenia, muscle strength tests combined with muscle volume measurements are needed (6). Associated with aging and the progression of sarcopenia, the composition of the muscles also changes and an increased fat infiltration occurs (7). However, the impact of the higher fat content inside the muscles is not yet fully understood (6). For improved understanding of the prevalence, onset, and progress of sarcopenia, new methods, including an accurate technique for measuring muscle volume, are needed (6–8). Another example where detailed and accurate knowledge of muscle volume and muscle composition is important is in whiplash-associated disorders (WAD) (9). A higher fat infiltration in the neck muscles has been found in people with WAD compared to healthy controls (10). A higher fat concentration in the quadriceps muscle associated to the fibromyalgia syndrome has also recently been found (11).

The current "gold standard" imaging method for the determination of muscle mass and its distribution is dual energy x-ray absorptiometry (DXA), which is rapid and readily available. However, DXA uses ionizing radiation and only enables analysis of 2D projections of the body. Therefore, no detailed muscle group separation, or quantification of fat content within the muscle tissue, can be obtained using DXA.

[1]Department of Biomedical Engineering (IMT), Linköping University, Linköping, Sweden.
[2]Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden.
[3]Advanced MR Analytics (AMRA) AB, Linköping, Sweden.
[4]Department of Radiation Physics and Department of Medical and Health Sciences, Linköping University, Linköping, Sweden.
[5]Department of Medical and Health Sciences (IMH), Linköping University, Linköping, Sweden.

Potential conflicts of interest: Anette Karlsson, Johannes Rosander, Thobias Romu, Magnus Borga, and Olof Dahlqvist Leinhard are stock holders in Advanced MR Analytics AB.

*Address reprint requests to: A.K., Linköping University, 581 83 Linköping, Sweden. E-mail: anette.k.karlsson@liu.se

Contract grant sponsor: Research Council of Southeast Sweden (FORSS).
Received May 10, 2014; Accepted July 30, 2014.
DOI 10.1002/jmri.24726
View this article online at wileyonlinelibrary.com.

A more accurate analysis can be made using tomographic methods, ie, computed tomography (CT) and magnetic resonance imaging (MRI). Water–fat separated MRI, based on Dixon imaging techniques (12), enables high soft-tissue contrast, enabling detailed measurements of the muscle volumes and fat infiltration. Drawbacks associated with MRI are availability and cost (7). With current techniques, scanning the whole body with sufficient resolution for body composition analysis may be achieved in less than 10 minutes (13,14). However, the work involved when manually segmenting the muscle tissue within the whole body is far too large to be feasible in anything but very small studies. Even when using optimized semiautomatic methods, a single segmentation of the whole-body muscular system may take several days to complete. The development of robust automatic segmentation of muscle tissue is therefore needed in order to make MRI an alternative for studying muscle tissue volume in large-scale studies.

Methods that automatically define and measure muscle tissue volume in MR images often use approaches based on morphological operations (15–17). The measurements are performed on parts of the body (15,16) or on the whole body's muscle tissue (17). One limitation of morphologically based approaches is that separation of adjacent muscles is extremely challenging. Baudin et al. (18) proposed one example for solving the segmentation of different muscles within the thigh by using random walks and prior knowledge.

Anatomical knowledge can also be incorporated into a segmentation method by atlases, ie, real or synthetic images with corresponding manually defined anatomical labels. The segmented atlases can then be transferred to a new subject (target) by nonrigid registration of the atlas onto the target's images. However, due to large anatomical variation between subjects and technical difficulties such as placement of arms and legs during scanning, a single registration may not converge correctly everywhere. Therefore, most atlas-based techniques address limited parts of the body, such as the brain, which shows a relatively limited variation in shape and location of its anatomical structures (19). Although the variation is rather small, previous studies on atlas-based segmentation of the brain have still shown increased robustness to anatomical variation when using more than one atlas followed by a voting scheme (20,21).

Nonrigid registration methods maximize the similarity between two images. The results vary depending on the similarity measure used. Common measures are based on differences in image intensity or local phase. One example of an intensity-based method is the Demons algorithm (22). An example that uses a phase-based similarity measure is the morphon method (23). Phase-based methods are insensitive to slow intensity variations, which are common in MR images due to B_0 and radiofrequency (RF) inhomogeneity. An additional feature of the morphon is its ability to deform the prototypes on different scales with different degrees of regularization, an important feature for whole-body registration.

The aim of this work was both to present and to assess a method capable of quantifying the volume of the whole skeletal muscular system, based on whole body water–fat separated MR images, while also enabling separation of the muscle tissue into major muscle groups. The method should also be robust with respect to changes in image resolution.

MATERIALS AND METHODS

The first part of this section describes the method for automated muscle tissue quantification. The second part describes how the method was evaluated.

Automated Muscle Tissue Volume Quantification

The method is divided into four different steps:

1. Whole-body water–fat separated imaging with intensity inhomogeneity correction.
2. Nonrigid registration of multiple atlases to the acquired image volume, ie, the target.
3. Muscle tissue classification using a voting scheme based on the multiple atlas registrations.
4. Muscle tissue volume quantification derived from the tissue classification and the local fat signal.

In Fig. 1, a flow chart of the method is provided. Details regarding the four steps are described below.

Step 1: Whole Body Intensity Corrected Water–Fat Separated Imaging

A prerequisite for this method is water–fat separated images with quantitative fat information. In this study, two-point Dixon imaging (12) with phase-sensitive reconstruction (24–26) is used to acquire the water–fat separated images. There are, however, several alternative separation methods that could be used (27–32).

During postprocessing, the water and fat image pairs were intensity inhomogeneity-corrected using the method described previously (13,24). Briefly, this method calculates a quantitative fat image based on pure adipose tissue as an internal intensity reference, ie, the signal intensity level in a given fat image voxel is related to the intensity in pure adipose tissue, which is given the value 1, corresponding to 100% adipose tissue. In addition to creating a quantitative fat image, the operation also removed some of the inhomogeneity in the original water image. See Fig. 2 for an example of a water and fat image pair.

Step 2: Nonrigid Multiple Atlas Registration

The starting point for atlas-based segmentation is the generation of the atlas. An atlas is here defined as a water and fat image pair with corresponding labels defining the different muscle groups (Fig. 3). The atlas is then registered onto a target with the aim that it should be as similar to the target as possible. For this task the water images are used, since they display the least anatomical variation and contain the most information regarding muscle shapes. The multiscaled and phase-based morphon method (23) was selected for

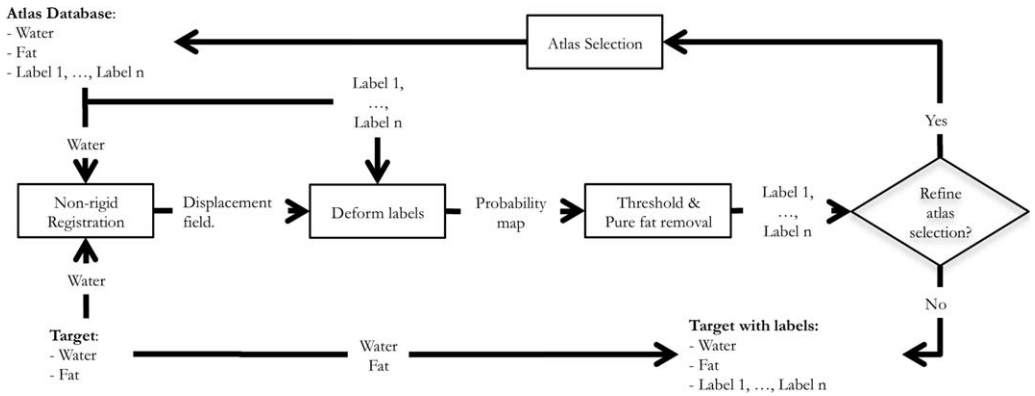

Figure 1. A flowchart describing the segmentation algorithm. All water volumes of the atlases in the database are individually registered onto the target's water volume. The labels of the atlases are deformed, and a probability map is acquired. This is followed by choosing the threshold and pure fat removal. Another iteration may be performed with refined atlas selection for improving the result.

the nonrigid registration. The morphon registration iteratively estimates a displacement field by applying directional quadrature filters on the prototype and the target and updates the displacement field such that the phase differences between the filter responses are minimized. After the convergence of the morphon, the resulting displacement field is used to transfer the anatomical knowledge, ie, the muscle labels, from the atlas to the new subject.

The robustness of the atlas-based segmentation is improved by registering multiple atlases to the target. Multiple atlas registration allows simultaneous use of atlases with different body composition, representing a plurality of anatomical variation, which may improve the segmentation performance. The registration is therefore made for each atlas and the suggested labels are added to each other on the target, forming a map, which after normalization can be interpreted as a probability map, as illustrated in Fig. 4. A value equal to 1 means that all atlases classify a voxel as muscular tissue, and a value equal to zero means that none of the atlases classifies that voxel as muscle tissue.

Step 3: Muscle Tissue Classification

The third step of the method is to classify the label of each voxel based on the resulting probability map from the multiple atlas registration. This is done by applying a threshold that determines how many atlases must agree to classify a voxel as a certain muscle. Too high a threshold would lead to an underestimation of the muscle volume, while too low a threshold would lead to an overestimation. Different muscle groups have different surroundings, so optimal thresholds for each group are used.

The optimal thresholds for each segmented muscle group are calculated by finding the threshold that maximizes the similarity between the manual segmentations and the automatic segmentation, based on the Similarity Index (SI), also known as the Dice

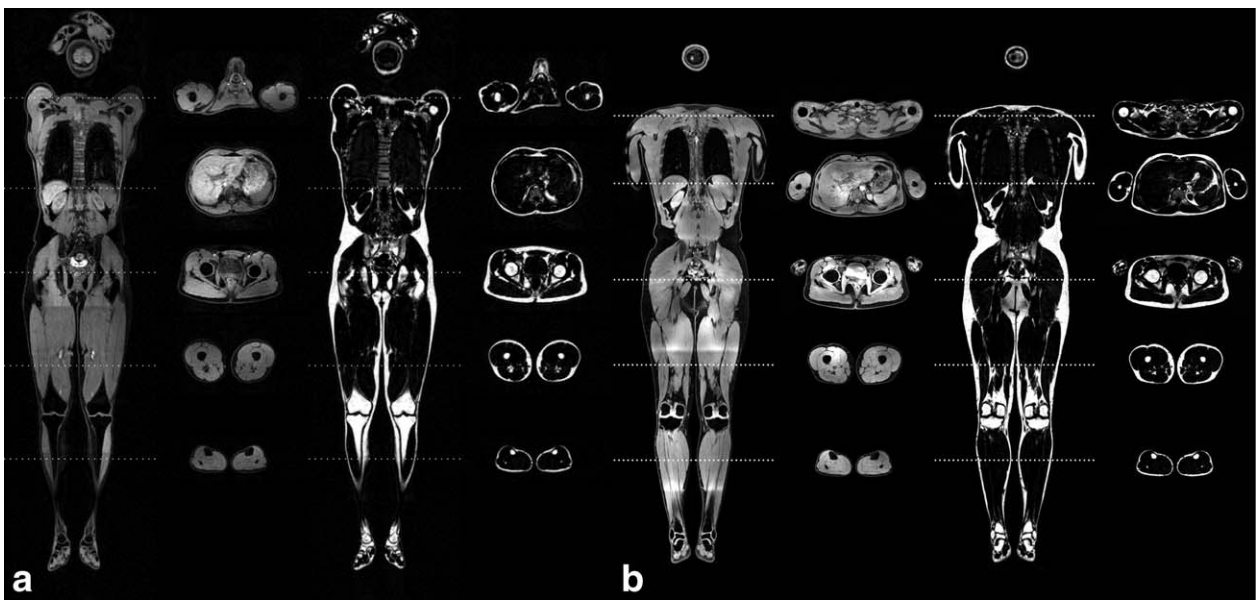

Figure 2. Coronal and transverse sections of an intensity-corrected water (left) and fat (right) image volume with (**a**) an image resolution of $3.5 \times 3.5 \times 3.5$ mm^3 (1.5 T) and (**b**) an image resolution of $1.75 \times 1.75 \times 1.75$ mm^3 (3.0 T).

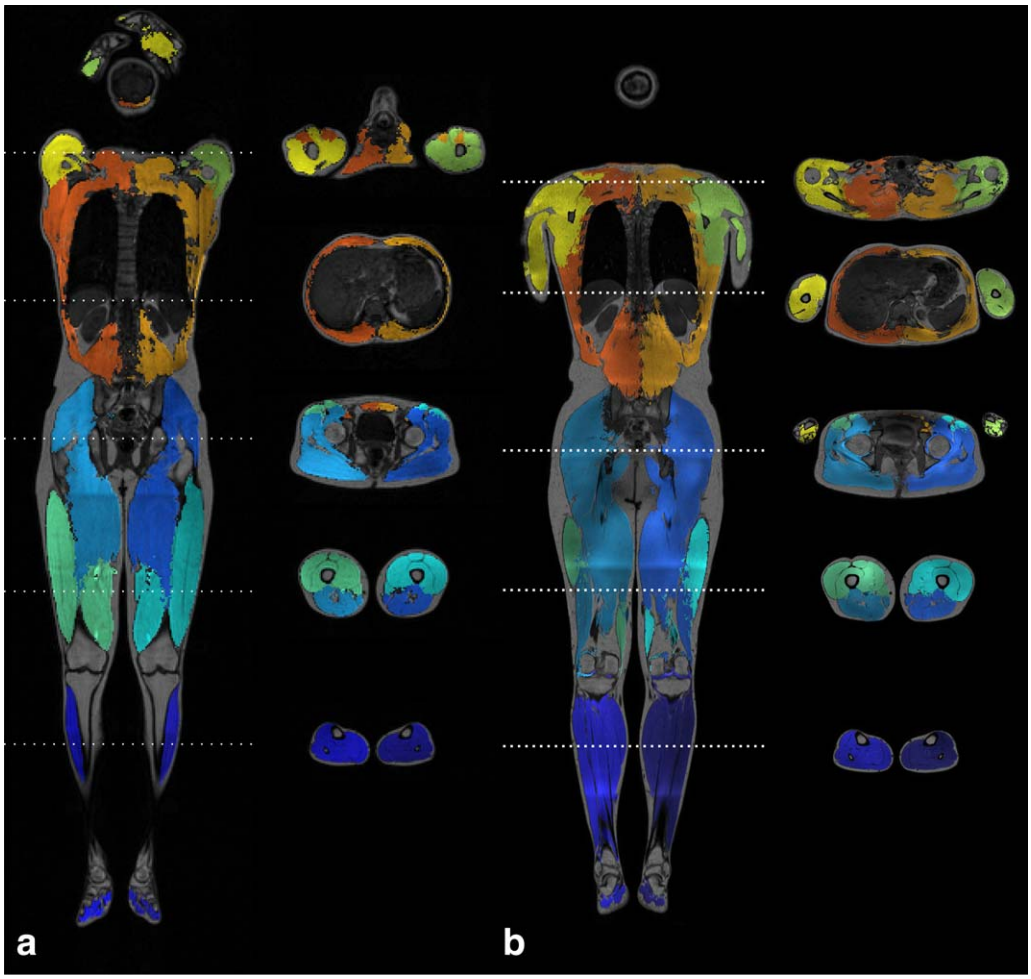

Figure 3. An atlas with the 10 different muscle labels shown using color overlay onto the intensity-corrected in-phase (water plus fat) image volume for (**a**) the 1.5 T scan and (**b**) the 3.0 T scan. The following labels were used in the atlas: Left and right lower leg, right and left anterior thigh, right and left posterior thigh (including gluteus maximus), left and right abdomen, and left and right arm.

coefficient (33). The threshold value that provides the highest mean SI (over all atlases used) is then used as the optimal threshold value.

The probability of achieving a good registration result is most probably higher if the atlas and the target are similar. An automated selection of atlases with muscle distributions similar to the target can improve the segmentation result and is here implemented as an iterative process. After the first iteration, where a random or general subset of atlases is used, a second iteration is performed where only the atlases with some desired feature, eg, the most similar whole body muscle volume, compared to the result from the first iteration, are used for a second round of voting.

Step 4: Muscle Tissue Volume Quantification

The final step to obtain the muscle volume is to combine the result from the muscle segmentation and classification steps with image information present in the target volume.

The first step of the muscle tissue volume quantification consists of removal of voxels not belonging to the body, ie, background removal. Only voxels inside a fuzzy body mask are included in order to reduce the influence of noise from voxels in the background. First, the water volume is normalized, as is done on the fat volume described earlier. Then a binary body mask is created by applying a threshold to the sum of the quantitative fat image and normalized water image at a level of 0.5. The fuzzy mask is created by setting the voxels within one voxel from the borders of the binary mask equal to the sum of the normalized water and fat images. By using this fuzzy mask, the partial volume effects on the volume measurement are minimized.

The volume of the segmented muscle mask, M_M, is calculated by:

$$M_M = \sum_{\forall\ voxels} (M_{AUT} \cdot m) \cdot V_{vox} \qquad [1]$$

where M_{AUT} is the automatic muscle segmentation, m is the fuzzy body mask, and V_{vox} is the volume of each voxel.

The next step for the quantification of muscle tissue volume corrects for fatty infiltration in M_{AUT}. This is

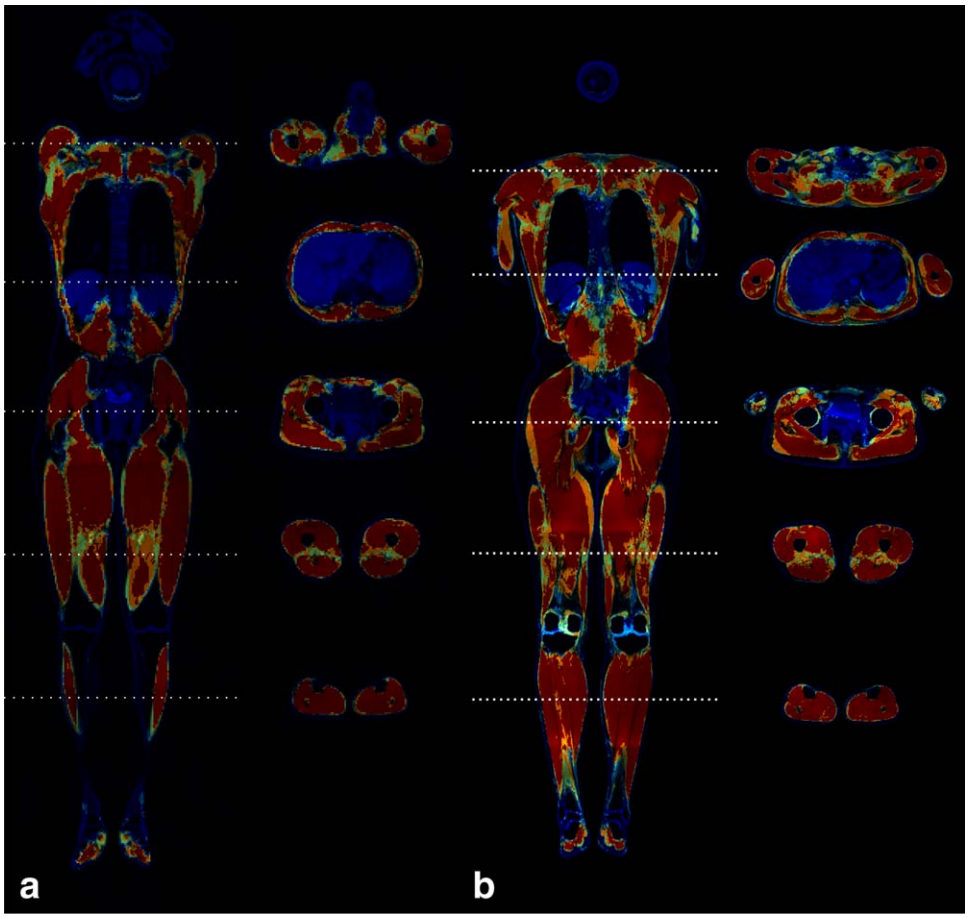

Figure 4. A whole-body muscle tissue probability map showing the result of the multi-atlas registration for 1.5 T (a) and 3.0 T (b). The color red indicates that all atlases segment the voxel as muscle tissue; blue indicates that no atlases segment the voxel as muscle tissue. The whole-body probability map was fused from the probability maps from 10 different muscle groups defined in a muscle atlas. At the borders, where overlaps occur, the highest probability value was used.

done in two different ways. The first definition is here called fat-free muscle volume, M_{FF}, where all voxels with fat concentration higher than 50% are removed from M_{AUT}.

$$M_{FF} = \sum_{\forall\ voxels} (M_{AUT}(f < 0.5) \cdot m) \cdot V_{vox} \qquad [2]$$

where f represents the intensity-corrected quantitative fat image. By this operation, subcutaneous tissue and large streaks of pure fat infiltration are removed from the muscle definition. This improves the robustness of the volume quantification, as muscle tissue in most regions is surrounded by subcutaneous tissue and errors in this interface are therefore corrected. The definition also enables the use of lower threshold values in the voting procedure, as the introduced error is corrected.

To avoid resolution-dependent muscle volume measurement and to also correct for diffuse muscle fat infiltration, a second definition is introduced measuring only the lean muscle volume, M_L:

$$M_L = \sum_{\forall\ voxels} ((1 - f) \cdot M_{AUT}) \cdot m \cdot V_{vox} \qquad [3]$$

Here the fat signal in each voxel classified as muscle tissue is subtracted from the final muscle volume estimate. In the presence of diffuse fat infiltration, M_L quantifies the amount of muscle tissue without including the diffuse fat infiltration. Similar to M_{FF}, M_L also removes regions containing pure adipose tissue, ie, fatty streaks and subcutaneous tissue.

Method Evaluation

Data Acquisition

This study was approved by the local ethics committee (Dnr: M88-07-T93-08, and Dnr: 2011/262-32). All volunteers participated after each giving written informed consent.

Ten healthy volunteers (six female and four male) were included for creation of atlases, calculation of optimal threshold values at 1.5 T and 3.0 T, and for initial method evaluation. They were scanned with a 1.5 T Philips Achieva MR-scanner and a 3.0 T Philips Ingenia MR-scanner (Philips Health Care, Best, the Netherlands) (see acquisition details, below). The time between the two different scans was less than 30 minutes. The ages ranged from 21 to 29 years and the mean age was 24.9 ± 2.4 (standard deviation, SD). The BMI ranged from 20.1 to 32.3 kg/m^2 and the mean BMI was 23.6 ± 3.8 (SD).

An additional 11 volunteers (four males, seven females) were included as validation data and were scanned in the 3.0 T scanner. The ages ranged from 33 to 54 years and the mean age was 43.6 ± 6.8 (SD). The BMI ranged from 19.7 to 32.3 and the mean BMI was 25.8 ± 3.6 (SD).

Table 1
Manual and Automatic Volumes (Mean, SD, and 95% Limits of Agreement) for Both Field Strengths and for All Muscle Groups Including the Total Body

Muscle Group	Field Strength	Operator	M_M Mean (SD) [l]		M_{FF} Mean (SD) [l]		M_L Mean (SD) [l]	
Left lower leg	1.5 T	Man	1.59	(0.19)	1.58	(0.19)	1.51	(0.19)
		Aut	1.62	(0.20)	1.55	(0.19)	1.50	(0.19)
	3.0 T	Man	1.59	(0.22)	1.57	(0.21)	1.50	(0.20)
		Aut	1.60	(0.20)	1.53	(0.19)	1.48	(0.19)
Right lower leg	1.5 T	Man	1.57	(0.18)	1.56	(0.18)	1.50	(0.18)
		Aut	1.60	(0.20)	1.54	(0.19)	1.50	(0.19)
	3.0 T	Man	1.58	(0.20)	1.56	(0.19)	1.51	(0.19)
		Aut	1.61	(0.19)	1.54	(0.18)	1.51	(0.18)
Left posterior thigh	1.5 T	Man	3.15	(0.60)	3.10	(0.60)	2.96	(0.61)
		Aut	3.33	(0.60)	3.13	(0.60)	3.00	(0.60)
	3.0 T	Man	3.19	(0.62)	3.14	(0.62)	2.99	(0.63)
		Aut	3.37	(0.62)	3.16	(0.62)	3.04	(0.63)
Right posterior thigh	1.5 T	Man	3.23	(0.59)	3.19	(0.60)	3.07	(0.61)
		Aut	3.42	(0.61)	3.22	(0.60)	3.11	(0.61)
	3.0 T	Man	3.36	(0.64)	3.29	(0.64)	3.17	(0.65)
		Aut	3.55	(0.68)	3.32	(0.66)	3.22	(0.68)
Left anterior thigh	1.5 T	Man	2.14	(0.45)	2.11	(0.45)	2.03	(0.45)
		Aut	2.19	(0.44)	2.10	(0.45)	2.03	(0.45)
	3.0 T	Man	2.10	(0.46)	2.07	(0.46)	1.99	(0.45)
		Aut	2.13	(0.45)	2.05	(0.45)	1.98	(0.45)
Right anterior thigh	1.5 T	Man	2.18	(0.46)	2.15	(0.46)	2.09	(0.46)
		Aut	2.24	(0.44)	2.14	(0.45)	2.09	(0.45)
	3.0 T	Man	2.08	(0.42)	2.06	(0.42)	2.01	(0.43)
		Aut	2.12	(0.42)	2.05	(0.42)	2.01	(0.44)
Left abdomen	1.5 T	Man	3.72	(0.82)	3.62	(0.82)	3.36	(0.81)
		Aut	4.01	(0.83)	3.69	(0.85)	3.46	(0.83)
	3.0 T	Man	3.48	(0.91)	3.35	(0.85)	3.10	(0.80)
		Aut	3.78	(0.96)	3.38	(0.83)	3.20	(0.82)
Right abdomen	1.5 T	Man	3.79	(0.81)	3.73	(0.82)	3.48	(0.82)
		Aut	4.06	(0.89)	3.78	(0.89)	3.56	(0.88)
	3.0 T	Man	3.42	(0.79)	3.34	(0.77)	3.13	(0.74)
		Aut	3.73	(0.88)	3.39	(0.81)	3.25	(0.81)
Left arm	1.5 T	Man	1.72	(0.49)	1.69	(0.49)	1.60	(0.47)
		Aut	1.66	(0.40)	1.53	(0.41)	1.47	(0.40)
	3.0 T	Man	2.21	(0.64)	2.16	(0.62)	2.04	(0.60)
		Aut	2.32	(0.64)	2.12	(0.60)	2.05	(0.58)
Right arm	1.5 T	Man	1.78	(0.50)	1.75	(0.50)	1.70	(0.50)
		Aut	1.76	(0.41)	1.63	(0.44)	1.60	(0.44)
	3.0 T	Man	2.28	(0.65)	2.21	(0.62)	2.16	(0.61)
		Aut	2.33	(0.69)	2.13	(0.63)	2.12	(0.62)
Whole body	1.5 T	Man	24.85	(4.94)	24.46	(4.97)	23.28	(4.98)
		Aut	26.18	(4.97)	24.57	(4.98)	23.16	(4.87)
	3.0 T	Man	25.28	(5.36)	24.74	(5.21)	23.59	(5.15)
		Aut	26.76	(5.70)	24.91	(5.31)	23.64	(5.23)

The acquired volumes were calculated using the muscle mask volume (M_M), the fat free muscle volume (M_{FF}), and the lean muscle volume (M_L) definitions.

Water–fat separated whole body images were acquired using the integrated quadrature body coil for the 1.5 T acquisition. At 3.0 T, the following coils were used: an integrated posterior phased array coil, a 16-channel phased array head and neck coil, and two flexible anterior phased array coils. The images were acquired using a 3D gradient echo sequence with opposite phase echo time of 2.3 msec and an in-phase echo time of 4.6 msec for the 1.5 T and 1.15 msec and 2.3 msec, respectively, for the 3.0 T. For the 1.5 T acquisition, the repetition time was 6.57 msec, the flip angle was 13°, and the acquired voxel resolution was $3.5 \times 3.5 \times 3.5$ mm^3. For the 3.0 T acquisition the repetition time was 3.78 msec with a flip angle of 10° and the acquired voxel resolution was $1.75 \times 1.75 \times 1.75$ mm^3. The image protocol was applied repeatedly, starting from the head, with a 30-mm and a 28-mm image stack overlap for 1.5 T and 3.0 T, respectively, until whole-body coverage was achieved. In the abdominal region, expiratory breath-hold acquisition was used for minimizing respiratory artifacts. The total scanning time was ~10 minutes for the 1.5 T and 25 minutes for the 3.0 T scanner. Water–fat separation of the in-phase and opposite-phase images was performed using phase sensitive reconstruction (24–26).

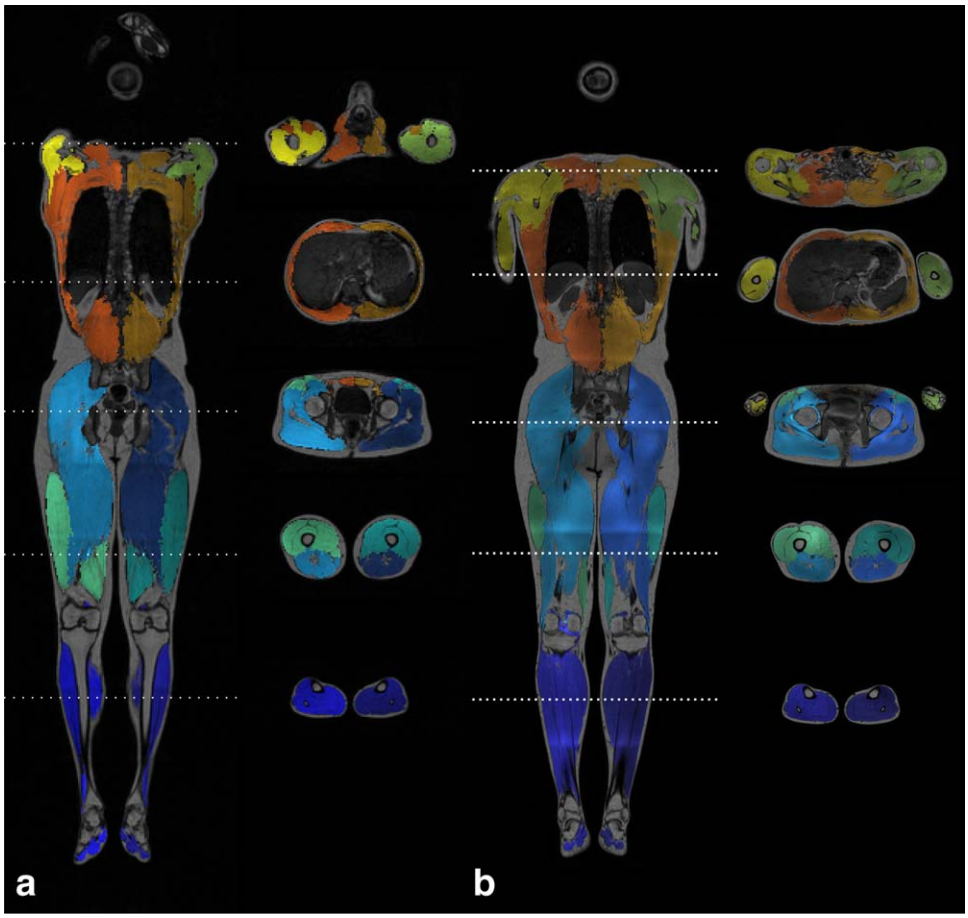

Figure 5. The resulting segmentation with the automatically labeled muscle groups shown in different colors at 1.5 T (**a**) and 3.0 T (**b**).

Atlas Creation

Anatomical muscle atlases were created by dividing the muscular system into 10 groups. The different muscle groups were labeled as follows: left lower leg, right lower leg, left posterior thigh, right posterior thigh, left anterior thigh, right anterior thigh, left abdomen, right abdomen, left arm, and right arm. The lower leg groups contained all muscles within the foot and the lower leg on the left and right side of the body, respectively. The posterior thigh groups contained the gluteus, hamstring, M. iliacus, and medial thigh muscles and the anterior thigh groups contained the M. sartorius and the quadriceps femoris. The arms contained all muscles belonging to the arm together with the rotator cuff and deltoid ideus. All other muscles were assigned the abdomen groups. Furthermore, a merged label, consisting of all identified muscle tissue, was used, representing the total skeletal muscle tissue volume. In Fig. 3 the definitions of the different labeled regions are shown.

The 10 muscle groups were segmented for all volunteers for both the 1.5 T and the 3.0 T acquisitions using a manual segmentation software (34). The software used manually placed foreground and background seed-points in the volume. Foreground points were placed in muscle tissue belonging to the desired region and background points were placed in other tissues. For each new point the software calculated a new and updated segmentation of the muscle tissue. This operation was repeated until the operator was satisfied with the muscle definition. The manual segmentations, together with the fat and the water images, comprised the 10 atlases in the multi-atlas segmentation. One trained operator (medical student in the late half of education) segmented all 10 datasets from the 1.5 T and another trained operator (M.S.C., with several years of experience of segmenting anatomical structures in MR images) segmented the 10 datasets from the 3.0 T.

Method Validation and Threshold Optimization

Two leave-one-out cross-validations were performed separately: one for the 1.5 T data and another one for the 3.0 T data. Leave-one-out cross-validation means that the method was applied using one subject as the target, considering its manual segmentation as ground truth, and the other nine as atlases. This was repeated for all 10 subjects, resulting in 10 targets with nine deformed labels per muscle group.

For each muscle group the optimal threshold value (according to SI) for classifying muscle tissue from the multi-atlas registration was calculated, both for the muscle masks, M_{AUT}, and the fat-free muscle masks, $M_{AUT}(f<0.5)$. It was not possible to calculate SI for the lean muscle volume since it is not interpretable as a logical mask.

The optimal mean SI for the fat-free muscle mask was higher than for the muscle mask in all muscle groups for both 1.5 T and 3.0 T. The optimal

Table 2
Delta Volumes (Mean, SD, 95% Limits of Agreement) Using the Fat Free Muscle Mask Definition and the Intraclass Correlation (ICC With 95% Limits of Agreement) for the Manual and the Automatic Fat Free Muscle Mask Volumes for Both Resolutions

			Delta Volume [L]					Intraclass Correlation			
					95% Limits				95% Limits		
Muscle Group		Type	Mean (SD)		of Agreement			ICC	of Agreement		
Left lower leg	Man - Aut	1.5 T	0.03	(0.08)	−0.12	-	0.18	0.92	0.74	-	0.98
		3.0 T	0.04	(0.08)	−0.13	-	0.20	0.91	0.70	-	0.98
	1.5 T - 3.0 T	Man	0.01	(0.08)	−0.14	-	0.16	0.93	0.74	-	0.98
		Aut	0.02	(0.05)	−0.08	-	0.12	0.97	0.88	-	0.99
Right lower leg	Man - Aut	1.5 T	0.02	(0.06)	−0.10	-	0.14	0.95	0.84	-	0.99
		3.0 T	0.02	(0.08)	−0.15	-	0.18	0.90	0.66	-	0.97
	1.5 T - 3.0 T	Man	0.00	(0.07)	−0.14	-	0.14	0.94	0.76	-	0.98
		Aut	0.00	(0.05)	−0.10	-	0.09	0.97	0.89	-	0.99
Left posterior thigh	Man - Aut	1.5 T	−0.03	(0.08)	−0.19	-	0.13	0.99	0.96	-	1.00
		3.0 T	−0.03	(0.10)	−0.22	-	0.17	0.99	0.95	-	1.00
	1.5 T - 3.0 T	Man	−0.04	(0.07)	−0.18	-	0.10	0.99	0.96	-	1.00
		Aut	−0.04	(0.04)	−0.12	-	0.05	1.00	0.97	-	1.00
Right posterior thigh	Man - Aut	1.5 T	−0.03	(0.06)	−0.15	-	0.08	0.99	0.97	-	1.00
		3.0 T	−0.03	(0.12)	−0.27	-	0.21	0.98	0.94	-	1.00
	1.5 T - 3.0 T	Man	−0.11	(0.12)	−0.35	-	0.13	0.97	0.74	-	0.99
		Aut	−0.11	(0.07)	−0.25	-	0.04	0.98	0.54	-	1.00
Left anterior thigh	Man - Aut	1.5 T	0.01	(0.07)	−0.13	-	0.16	0.99	0.95	-	1.00
		3.0 T	0.03	(0.08)	−0.13	-	0.18	0.99	0.95	-	1.00
	1.5 T - 3.0 T	Man	0.04	(0.08)	−0.13	-	0.21	0.98	0.92	-	1.00
		Aut	0.05	(0.03)	0.00	-	0.11	0.99	0.64	-	1.00
Right anterior thigh	Man - Aut	1.5 T	0.01	(0.05)	−0.09	-	0.11	0.99	0.98	-	1.00
		3.0 T	0.01	(0.06)	−0.12	-	0.13	0.99	0.96	-	1.00
	1.5 T - 3.0 T	Man	0.09	(0.08)	−0.06	-	0.25	0.96	0.55	-	0.99
		Aut	0.09	(0.04)	0.01	-	0.17	0.98	0.16	-	1.00
Left abdomen	Man - Aut	1.5 T	−0.08	(0.15)	−0.38	-	0.22	0.98	0.93	-	1.00
		3.0 T	−0.03	(0.42)	−0.85	-	0.79	0.89	0.61	-	0.97
	1.5 T - 3.0 T	Man	0.27	(0.23)	−0.19	-	0.72	0.92	0.30	-	0.98
		Aut	0.32	(0.13)	0.05	-	0.58	0.93	−0.01	-	0.99
Right abdomen	Man - Aut	1.5 T	−0.04	(0.22)	−0.48	-	0.39	0.97	0.88	-	0.99
		3.0 T	−0.05	(0.38)	−0.80	-	0.69	0.89	0.63	-	0.97
	1.5 T - 3.0 T	Man	0.39	(0.20)	0.00	-	0.79	0.86	−0.04	-	0.98
		Aut	0.39	(0.19)	0.01	-	0.77	0.89	−0.03	-	0.98
Left arm	Man - Aut	1.5 T	0.15	(0.14)	−0.11	-	0.42	0.91	0.28	-	0.98
		3.0 T	0.03	(0.13)	−0.23	-	0.30	0.98	0.91	-	0.99
	1.5 T - 3.0 T	Man	−0.47	(0.17)	−0.80	-	−0.14	0.71	−0.05	-	0.94
		Aut	−0.59	(0.24)	−1.06	-	−0.12	0.54	−0.07	-	0.89
Right arm	Man - Aut	1.5 T	0.13	(0.12)	−0.11	-	0.36	0.94	0.47	-	0.99
		3.0 T	0.08	(0.19)	−0.30	-	0.45	0.95	0.82	-	0.99
	1.5 T - 3.0 T	Man	−0.45	(0.20)	−0.83	-	−0.07	0.72	−0.06	-	0.94
		Aut	−0.51	(0.22)	−0.94	-	−0.07	0.64	−0.07	-	0.92
Whole body	Man - Aut	1.5 T	−0.10	(0.70)	−1.47	-	1.27	0.99	0.97	-	1.00
		3.0 T	−0.17	(1.37)	−2.85	-	2.50	0.97	0.88	-	0.99
	1.5 T - 3.0 T	Man	−0.27	(0.91)	−2.05	-	1.51	0.98	0.94	-	1.00
		Aut	−0.35	(0.69)	−1.70	-	1.01	0.99	0.96	-	1.00

The delta volume and the ICC using the fat free muscle mask definition for the 1.5 T and 3.0 T manual and automatic results are presented.

threshold was therefore considered the thresholds acquired from the fat-free masks. In the second iteration, the four atlases with the most similar total body muscle volume (calculated using the M_{FF} definition) were selected, and at least two of them had to agree in order for a voxel to be classified as muscle tissue.

To validate the method further, nine of the 3.0 T volunteers were used as atlases to segment the 11 validation datasets using the same predetermined thresholds. The result was compared with manual segmentations, performed by a medical student in the late half of education.

Statistics

The intraclass correlation (IIC) was calculated at a 95% confidence level using SPSS v. 22 (IBM, Armonk, NY). MatLab_R2013b (MathWorks, Natick, MA) was used for all other calculations. The volume and SD together with 95% limits of agreement were calculated for each muscle group and for all three muscle volume definitions.

The accuracy of the measurements, ie, ground truth segmentations compared to automated segmentation results, was evaluated by calculating the delta volume

Table 3
True Positive Volume Fraction (TPVF), False Positive Volume Fraction (FPVF), and the False Negative Volume Fraction (FNVF) for 1.5 T and 3.0 T for All Muscle Groups Using the Fat Free Muscle Mask (M_{FF}) Definition

Muscle Group	Field Strength	TPVF Mean	(SD)	FNVF Mean	(SD)	FPVF Mean	(SD)
Left lower leg	1.5 T	0.93	(0.03)	0.07	(0.03)	0.05	(0.02)
	3.0 T	0.93	(0.04)	0.07	(0.04)	0.05	(0.01)
Right lower leg	1.5 T	0.93	(0.02)	0.07	(0.02)	0.05	(0.02)
	3.0 T	0.93	(0.04)	0.07	(0.04)	0.06	(0.02)
Left posterior thigh	1.5 T	0.93	(0.01)	0.07	(0.01)	0.08	(0.02)
	3.0 T	0.93	(0.01)	0.07	(0.01)	0.08	(0.03)
Right posterior thigh	1.5 T	0.93	(0.01)	0.07	(0.01)	0.08	(0.01)
	3.0 T	0.93	(0.02)	0.07	(0.02)	0.08	(0.02)
Left anterior thigh	1.5 T	0.93	(0.02)	0.07	(0.02)	0.06	(0.01)
	3.0 T	0.93	(0.03)	0.07	(0.03)	0.06	(0.02)
Right anterior thigh	1.5 T	0.93	(0.02)	0.07	(0.02)	0.06	(0.01)
	3.0 T	0.93	(0.03)	0.07	(0.03)	0.06	(0.01)
Left abdomen	1.5 T	0.80	(0.07)	0.20	(0.07)	0.12	(0.05)
	3.0 T	0.90	(0.04)	0.10	(0.04)	0.09	(0.02)
Right abdomen	1.5 T	0.82	(0.06)	0.18	(0.06)	0.11	(0.03)
	3.0 T	0.88	(0.05)	0.12	(0.05)	0.09	(0.03)
Left arm	1.5 T	0.88	(0.03)	0.12	(0.03)	0.14	(0.03)
	3.0 T	0.84	(0.06)	0.16	(0.06)	0.17	(0.06)
Right arm	1.5 T	0.86	(0.04)	0.14	(0.04)	0.14	(0.03)
	3.0 T	0.83	(0.05)	0.17	(0.05)	0.18	(0.05)
Whole body	1.5 T	0.91	(0.02)	0.09	(0.02)	0.07	(0.01)
	3.0 T	0.92	(0.03)	0.08	(0.03)	0.07	(0.02)

(with mean, SD, and 95% limits of agreement), together with true positive volume fraction (TPVF), false negative volume fraction (FNVF), and false positive volume fraction (FPVF) (35). Due to the large amount of voxels inside the image that did not belong to the set of labeled muscle voxels, the definition of FPVF was slightly altered to take all false positive voxels outside of the ground truth mask divided by the number of voxels inside the ground truth mask. (Udupa et al [2006] defined this as taking all voxels outside divided by the total number of voxels outside the mask (35).)

RESULTS

The results for the paired 1.5 T and 3.0 T data are presented in Tables (1–3). In Table 4, the result from the validation data is summarized.

The mean muscle volumes were very similar regardless of operator (manual or automatic) and of scanner modality/resolution for all muscle groups and all three muscle definitions (Table 1). M_M gave the largest estimated volumes while M_L gave the smallest muscle volume for all muscle groups, scanners, and operators. There was no large difference between the M_M and M_{FF} volumes for the manual operator. The

Table 4
Delta Volumes (Mean, SD, 95% Limits of Agreement) Using the Fat Free Muscle Mask Definition and the Intraclass Correlation (ICC With 95% Limits of Agreement) for the Manual and the Automatic Fat Free Muscle Mask Volume Using 9 Volunteers as Atlases and 11 Volunteers as Targets

Muscle Group	Delta Volume [l] Mean	(SD)	95% Limits of Agreement		Intraclass Correlation ICC	95% Limits of Agreement	
Left lower leg	0.05	(0.04)	−0.03	0.12	0.99	0.76	1.00
Right lower leg	0.07	(0.06)	−0.05	0.19	0.97	0.66	0.99
Left posterior thigh	0.19	(0.06)	0.08	0.30	0.98	0.07	1.00
Right posterior thigh	0.10	(0.09)	−0.07	0.28	0.99	0.85	1.00
Left anterior thigh	−0.06	(0.06)	−0.18	0.07	0.99	0.91	1.00
Right anterior thigh	0.02	(0.04)	−0.07	0.10	1.00	0.99	1.00
Left abdomen	0.04	(0.27)	−0.48	0.57	0.97	0.89	0.99
Right abdomen	0.01	(0.24)	−0.45	0.48	0.97	0.91	0.99
Left arm	0.40	(0.43)	−0.44	1.25	0.75	0.13	0.93
Right arm	0.38	(0.29)	−0.19	0.94	0.86	0.08	0.97
Whole body	0.76	(0.83)	−0.86	2.39	0.99	0.90	1.00

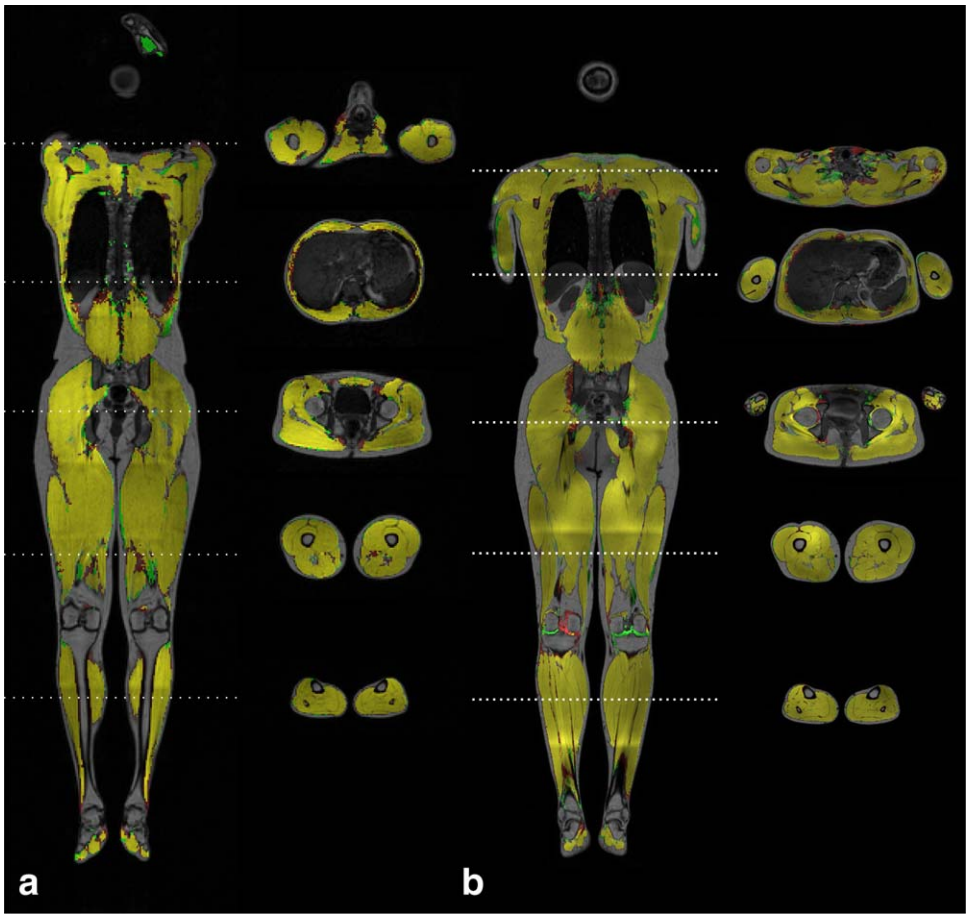

Figure 6. An image showing the result of the whole-body muscle tissue segmentation compared to the ground truth manual segmentation at 1.5 T (**a**) and 3.0 T (**b**). The yellow overlay indicates true positive voxels, the red overlay indicates false positive voxels, and the green overlay indicates false negative voxels. True negative voxels are shown in grayscale.

automatic operator resulted in a decrease in volume between M_M and M_{FF}. The automatic operator also overestimated the M_M volumes for both 1.5 T and 3.0 T compared to the manual operators.

A typical result of the multi-atlas segmentation is shown in Fig. 5 for both 1.5 T (Fig. 5a) and 3.0 T (Fig. 5b). The segmentation robustly classified all muscle groups defined by the atlas in the paired 1.5 T and 3.0 T data as well as in the validation data. This is also in agreement with the low mean and SD calculated for the delta volume in all defined muscle groups, seen in Table 2 and Table 4. However, a higher variability in the segmentation of the arms was observed.

In Fig. 6, true positive, false positive, and false negative voxels are shown in yellow, red, and green, respectively. Except for the outer border of the muscles, the only regions with inconsistent segmentations were found at complex shapes such as the joints and the pelvic region. A high mean TPVF equal to 0.93 ± 0.02 and 0.93 ± 0.03 (mean $\pm$ SD) was observed for the muscle groups in the lower part of the body for 1.5 T and 3.0 T, respectively. In the upper part of the body, the mean TPVF was 0.84 ± 0.05 for 1.5 T and $0.86 \pm 0.05$for 3.0 T (mean $\pm$ SD). On a whole-body level the TPVF was 0.91 ± 0.02 for 1.5 T and 0.92 ± 0.03 for 3.0 T (mean $\pm$ SD) (see details in Table 3).

There was a strong ICC between the automatic and the manual segmentation (see Tables 2 and 4). The ICC was above 0.89 for M_{FF} and M_L in all muscle groups for the paired 1.5 T and 3.0 T data. In the validation data, ICC was above 0.97 for all muscle groups except for the arms.

DISCUSSION

The validation of the presented method showed high accuracy compared to manual ground truth segmentations and that the method was able to provide muscle volume estimates independent of the different field strengths used and the eight-fold resolution differences of the data. This was seen both in the high TPVF values that is similar to other methods specialized solely to specific muscle groups (15) and by the low limits of agreements seen for most muscle groups as well as for the whole-body muscle tissue volume. The validation also showed high reproducibility, as the differences between automated measurements at 1.5 T and 3.0 T were smaller than the corresponding manual comparison for most muscle groups. The validation data indicates a high robustness of the method, as similar delta volumes and ICC values were obtained comparing automatic and manual segmentations.

A major advantage with the presented method is that it is compatible with compartmental fat quantification techniques, based on the same water–fat separated images. For example, the volume of visceral adipose tissue, subcutaneous tissue, and total adipose tissue volume can be quantified (13,36). By combining the techniques, deeper knowledge about the interactions between muscle and body metabolism, body lipid storage, disease progression, and intervention effects can be achieved.

The method showed a larger variability in the upper part of the body compared to the lower part. This is in accordance with other whole-body measurements (17). In the 1.5 T data, an increased variability was observed in the resulting arm segmentations due to high variability in arm positioning. Consistent placement of the arms within the field of view (FOV) is difficult in whole-body MRI. There was less variability in the 3.0 T setup, where the arms were positioned along the side of the subjects, although this strategy is problematic in obese subjects. Here, eight subjects were lean (BMI between 20–25 kg/m^2) and had no problem fitting the FOV in the left–right direction. A larger variability was observed in the arms for the validation data, where only 4 out of 11 subjects were lean. Improved strategies for consistent arm positioning are therefore needed.

Age and BMI differences do not seem to affect the result of the methods performance. Even though lean young subjects were used as atlases and the validation data had slightly larger and much older subjects, the method showed high accuracy. This is in agreement with previous published results using this method where 11 atlases were used to segment 10 normal (BMI = 18–25), 10 overweight (BMI = 25–30), and 10 obese (BMI >30) subjects (37).

Each of the steps in the method can be optimized separately. Here we have presented one successful way of combining the different modules in order to perform automatic muscle quantification that works at two different field strengths with different acquired resolutions, signal to noise ratio levels, repetition times, and echo times.

A general advantage of the method is that it is initialized using manually defined atlases. If another definition set is desired, a new atlas with the new anatomical regions of interest can easily be defined. Here, a framework allowing straightforward validation of such new atlases is presented, as the performance of a new atlas definition cannot be guaranteed in advance.

The muscle mask extraction was optimized based on SI of the fat-free muscle, M_{FF}, definition. This is not optimal if quantitative information about fat infiltration inside the muscle is desired. For such an application, much emphasis must be put on minimizing the FPVF, thereby minimizing the influence from surrounding subcutaneous adipose tissue. The segmentation will then provide a region of interest within the muscle tissue rather than a region describing the complete muscle. M_{FF} offers a solution for this problem as long as the diffuse fat infiltration is less than 50%. However, some intramuscular lipid depots tend to be thicker than the voxel dimension and will be excluded using this definition. The lean muscle volume definition cannot easily be extended for quantification of fat infiltration of the muscle tissue. Generally, multipoint techniques would be advantageous as they enable T_2*-corrected fat quantification, but the two-point approach was selected due to its excellent scan time performance.

The detailed knowledge about the skeletal muscular system provided by this method is particularly useful in longitudinal interventional studies. The rescan variability using the method on a single scanner is lower than the variability between the different field strengths observed here (37), and will allow early detection of the intervention effects on the volume of the muscular system.

Further developments and validations of the presented method are needed in order for it to be feasible in larger longitudinal clinical studies. First, validation needs to be performed in patients suffering from muscular diseases and dystrophies. Second, disease-optimized atlases with more detailed muscle group definitions need to be developed and validated, as the definitions will vary depending on the clinical question. Finally, an extension of the method to include also quantification of intramuscular adipose tissue would be beneficial.

In conclusion, this article presents a novel method that accurately, reproducibly, and automatically quantifies total muscle volume and individual muscle groups, based on a rapid MR scan. The method's performance is similar to that of the manual operator. Although further validation is needed, we have shown its independence of scanner field strength and resolution, which is an important prerequisite for large-scale multicenter, or longitudinal studies, for investigation of muscle disorders such as sarcopenia.

ACKNOWLEDGMENTS

The authors thank Prof. Peter Lundberg and Associate Professor Anneli Peolsson at the Department of Medical and Health Sciences at Linköping University for recruiting the study subjects and writing the ethical board applications.

REFERENCES

1. Emery AE. The muscular dystrophies. Lancet 2002;359:687–695.
2. Willis TA, Hollingsworth KG, Coombs A, et al. Quantitative muscle MRI as an assessment tool for monitoring disease progression in LGMD2I: a multicentre longitudinal study. PLoS One 2013;8: e70993.
3. Foley S, Ding C, Cicuttini F, Jones G. Physical activity and knee structural change: a longitudinal study using MRI. Med Sci Sports Exerc 2007;39:426–434.
4. Garcia J. MRI in inflammatory myopathies. Skeletal Radiol 2000; 29:425–438.
5. Qin W, Bauman WA, Cardozo C. Bone and muscle loss after spinal cord injury: organ interactions. Ann N Y Acad Sci 2010;1211: 66–84.
6. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis: report of the

European Working Group on Sarcopenia in Older People. Age Ageing 2010;39:412–423.
7. Cruz-Jentoft AJ, Morley JE (eds.). Sarcopenia. Somerset, NJ: Wiley-Blackwell; 2012.
8. Fielding RA, Vellas B, Evans WJ, et al. Sarcopenia: an undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International Working Group on Sarcopenia. J Am Med Direct Assoc 2011;12: 249–256.
9. Elliott JM, Kerry R, Flynn T, Parrish TB. Content not quantity is a better measure of muscle degeneration in whiplash. Man Ther 2013:1–5.
10. Elliott J, Jull G, Noteboom JT, Darnell R, Galloway G, Gibbon WW. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders: a magnetic resonance imaging analysis. Spine 2006;31:E847–855.
11. Gerdle B, Forsgren MF, Bengtsson A, et al. Decreased muscle concentrations of ATP and PCR in the quadriceps muscle of fibromyalgia patients—a 31P-MRS study. Eur J Pain 2013;17:1205–1215.
12. Dixon W. Simple proton spectroscopic imaging. Radiology 1984; 153:189–194.
13. Dahlqvist Leinhard O, Johansson A, Rydell J, et al. Quantitative abdominal fat estimation using MRI. In: Proc Int Conf Pattern Recog (ICPR) 2008;art.no. 4761764.
14. Kullberg J, Johansson L, Ahlstrom H, et al. Automated assessment of whole-body adipose tissue depots from continuously moving bed MRI: a feasibility study. J Magn Reson 2009;30:185–193.
15. Broderick BJ, Dessus S, Grace PA, ÓLaighin G. Technique for the computation of lower leg muscle bulk from magnetic resonance images. Med Eng Phys 2010;32:926–933.
16. Brunner G, Nambi V, Yang E, et al. Automatic quantification of muscle volumes in magnetic resonance imaging scans of the lower extremities. Magn Reson Imaging 2011;29:1065–1075.
17. Wald D, Teucher B, Dinkel J, et al. Automated quantification of adipose and skeletal muscle tissue in whole-body MRI data for epidemiological studies. Proc Spie 2012;8315.
18. Baudin P-Y, Azzabou N, Carlier PG, Paragios N. Prior knowledge, random walks and human muscle segmentation. Nice, France. Berlin, Heidelberg: Springer; 2012. p 569–576.
19. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage 1999;9: 179–194.
20. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. NeuroImage 2009;46:726–738.
21. Bai W, Shi W, O'Regan DP, et al. A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: application to cardiac MR images. IEEE Trans Med Imaging 2013;32:1302–1315.
22. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal 1998;2:243–260.
23. Knutsson H, Mats A. Morphons: segmentation using elastic canvas and paint on priors. Genoa, Italy; 2005.
24. Romu T, Borga M, Dahlqvist Leinhard O. MANA — multi scale adaptive normalized averaging. In: Proc Int Symp Biomed Imaging (ISBI) 2011;art.no. 5872424:361–364.
25. Rydell J, Johansson A, Dahlqvist Leinhard O, et al. Three dimensional phase sensitive reconstruction for water/fat separation in MR imaging using inverse gradient. In: Proc 16th Annual Meeting ISMRM, Toronto; 2008.
26. Rydell J, Knutsson H, Pettersson J, et al. Phase sensitive reconstruction for water/fat separation in MR imaging using inverse gradient. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Brisbane, Australia; 2007.
27. Berglund J, Ahlstrom H, Johansson L, Kullberg J. Two-point Dixon method with flexible echo times. Magn Reson Med 2011; 65:994–1004.
28. Berglund J, Kullberg J. Three-dimensional water/fat separation and T2* estimation based on whole-image optimization—application in breathhold liver imaging at 1.5 T. Magn Reson Med 2012; 67:1684–1693.
29. Hernando D, Haldar JP, Sutton BP, Ma J, Kellman P, Liang ZP. Joint estimation of water/fat images and field inhomogeneity map. Magn Reson Med 2008;59:571–580.
30. Ma J. Dixon techniques for water and fat imaging. J Magn Reson Imaging 2008;28:543–558.
31. Reeder SB, Pineda AR, Wen Z, et al. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): application with fast spin-echo imaging. Magn Reson Med 2005;54:636–644.
32. Yu H, McKenzie CA, Shimakawa A, et al. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. J Magn Reson Imaging 2007;26:1153–1161.
33. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans Med Imaging 1994;13:716–724.
34. Malmberg F, Lindblad J, Nyström I. Sub-pixel segmentation with the image foresting transform. Playa del Carmen, Mexico; 2009.
35. Udupa JK, Leblanc VR, Zhuge Y, et al. A framework for evaluating image segmentation algorithms. Comput Med Imaging Graphics 2006;30:75–87.
36. Ludwig UA, Klausmann F, Baumann S, et al. Whole-body MRI-based fat quantification: a comparison to air displacement plethysmography. J Magn Reson Imaging 2014 [Epub ahead of print].
37. Thomas MS, Newman D, Leinhard OD, et al. Test-retest reliability of automated whole body and compartmental muscle volume measurements on a wide bore 3T MR system. Eur Radiol 2014.