Published in final edited form as:

IEEE Trans Biomed Eng. 2016 April ; 63(4): 805–813. doi:10.1109/TBME.2015.2474305.

Image-based Tissue Distribution Modeling for Skeletal Muscle Quality Characterization

S. Makrogiannis,

Mathematical Sciences Dept., Delaware State University, Dover, DE 19901-2277, USA

K. W. Fishbein,

National Institute on Aging, National Institutes of Health, Baltimore, MD 21224-6825, USA

A. Z. Moore,

National Institute on Aging, National Institutes of Health, Baltimore, MD 21224-6825, USA

R. G. Spencer, and

National Institute on Aging, National Institutes of Health, Baltimore, MD 21224-6825, USA

L. Ferrucci

National Institute on Aging, National Institutes of Health, Baltimore, MD 21224-6825, USA

S. Makrogiannis: smakrogiannis@desu.edu

Abstract

The identification and characterization of regional body tissues is essential to understand changes that occur with aging and age-related metabolic diseases such as diabetes and obesity and how these diseases affect trajectories of health and functional status. Imaging technologies are frequently used to derive volumetric, area, and density measurements of different tissues. Despite the significance and direct applicability of automated tissue quantification and characterization techniques, these topics have remained relatively under-explored in the medical image analysis literature. We present a method for identification and characterization of muscle and adipose tissue in the mid-thigh region using MRI. We propose an image-based muscle quality prediction technique that estimates tissue-specific probability density models and their eigenstructures in the joint domain of water- and fat-suppressed voxel signal intensities along with volumetric and intensity-based tissue characteristics computed during the quantification stage. We evaluated the predictive capability of our approach against reference biomechanical muscle quality measurements using statistical tests and classification performance experiments. The reference standard for muscle quality is defined as the ratio of muscle strength to muscle mass. The results show promise for the development of non-invasive image-based muscle quality descriptors.

Index Terms

Tissue identification and characterization; probabilistic modeling; magnetic resonance imaging (MRI)

I. Introduction

Changes in body composition occur in all aging individuals and have profound influences on risk for diseases and decline of physical function. Assessment of body composition is critical in the study of age-related and metabolic diseases [1], [2]. Loss in muscle strength and mass associated with aging - termed as sarcopenia - is a correlate of poor function and a risk factor for limited mobility. This pathological condition has important health and economic implications. In the East Boston cohort of the Established Populations for Epidemiologic Studies of the Elderly, it was observed that 38% of the men and 59% of the women over 65 years of age had difficulty stooping, and 24% of men and 29% of the women had difficulty lifting their arm over the shoulder height [3]. It is estimated that a large portion of such disability is attributed to muscle impairment. Direct healthcare costs due to sarcopenia in the U.S. in 2000 were estimated to exceed \$18.5 billion dollars [4]. It has been shown that age-related loss in muscle strength is caused in part by a decline in muscle mass [5]. However, it is becoming clear that decline in muscle strength that occurs with aging is much larger than what would be expected by changes in muscle mass alone [6]. To explain this discrepancy, investigators have explored the relationship between skeletal muscle composition and structure, and functional measures of muscle strength. For example, it is widely acknowledged that fat infiltration in the muscle may contribute to muscle weakness, metabolic diseases and frailty [7]. Therefore, image analysis techniques that could capture tissue characteristics associated with muscle quality are attracting considerable interest.

Medical imaging technologies provide a non-invasive and accurate means for body composition assessment. The more frequently utilized imaging modalities in related clinical studies are computed tomography (CT) and magnetic resonance (MR) imaging [8], [1], [9], [10], [11]. MR imaging can distinguish different types of soft tissue, especially if tissue-selective imaging techniques are used [8], [12]. CT has also been used to characterize muscle quality as a correlate to muscle strength [13], [1] mainly because the voxel intensities are measured in the standardized Hounsfield unit (HU) scale and due to the relatively fast image acquisition speed. Despite the increasing interest in assessment of muscle quality by medical imaging, the specific literature on this topic is limited. Most of the published studies have employed supervised work-flows, including manual segmentation and selection of ROIs, to quantify tissue volumes and tissue intensity averages. More importantly, there have been, to the best of our knowledge, no previous reports on automated muscle characterization with MRI.

In this report, we propose a method for assessment of muscle quality from tissue-selective MRI. This method is applied to volumetric non-suppressed (NS), water-suppressed (WS) and fat-suppressed (FS) images acquired at the mid-thigh level. An example of our data is displayed in Figure 1. Tissue segmentation follows the main stages of the method proposed in [11]. After the segmentation of subcutaneous adipose tissue (SAT) and bone, we estimated muscle and inter-muscular adipose tissue (IMAT) normal density models in the joint domain of WS and FS voxel intensities, and produced a muscle quality descriptor using subspace decomposition of muscle and statistical measures of variability. Our premise is that this set of features can be used to assess muscle quality. Here, we test the method in its

ability to discriminate between individuals with high and low muscle quality. The reference standard for muscle quality is a functional measure defined as the ratio of muscle strength to muscle mass. Our experiments suggest that the proposed descriptor extracts information that is relevant for prediction of muscle quality.

II. Tissue Identification

The objective of this stage is to identify regional adipose tissue, muscle and bone compartments as seen in Figure 2 and to extract information related to the volume of different tissues and their distribution of intensities.

In the first stage, we identified the mid-thigh region-of-interest (ROI) along the superior-inferior axis to restrict our analysis within the anatomical site of interest. We then applied the N4ITK inhomogeneity correction algorithm [14] on the set of non-, water- and fat-suppressed images. This algorithm reduces the low frequency intensity nonuniformity, or bias, that is present in MRI data and is more pronounced for scanners with higher B_0 field, such as 3 Tesla and above. In this method, authors first assumed an image formation model with multiplicative bias field of the form $v(\omega) = u(\omega)f(\omega)$, where, v is the acquired image, u is the ideal image, f is the bias field, and $\omega \in \Omega$ is the vector of spatial coordinates for an N -dimensional image, hence $\Omega \subseteq \mathbb{Z}^N$. The logarithmic form of this model is $v(\hat{\omega}) = \hat{u}(\hat{\omega}) + \hat{f}(\hat{\omega})$ $v(\hat{\omega}) = \log v(\omega)$, where $\hat{u}(\hat{\omega}) = \log u(\omega)$ and $\hat{f}(\hat{\omega}) = \log f(\omega)$. The objective is to restore the original image $\hat{u}(\hat{\omega}) = v(\hat{\omega}) - \hat{f}(\hat{\omega})$. A numerical solution is found in an iterative scheme $\hat{u}^n = \hat{v}^{n-1} - \hat{f}_e^r$ at the n th iteration. The residual bias field $\hat{f}_e^r$ is approximated using B-splines $S^*\{\cdot\}$ in multiple resolutions by $\hat{f}_e^r = S^*\{\hat{u}^{n-1} - E[\hat{u}^{n-1}]\}$.

Next, we linearly co-registered the FS and WS images into the NS spatial domain to correct for possible subject motion using the the method proposed in [15]. Because the optimization functions in image registration are in general non-convex, the authors proposed cost function apodization and a global-local optimization method. We used mutual information-based cost function to address multi-modal variations.

After the co-registration we applied linear intensity transformations to standardize the MR intensity dynamic range and reduce the inter-subject variability. We identified the fat region using centroid clustering [16] on the NS volumetric image and used the average fat intensities of NS, FS and WS volumes as reference values to linearly scale the voxel intensities and standardize their dynamic range. Our hypothesis is that the fat tissue intensities are mostly dominated by the subcutaneous fat, whose MR intensity values are largely invariant to the physiological factors therefore can be used as a reference point to define linear intensity transformations.

After selecting one leg using the spatial image coordinates, we used parametric active contour models to extract the subcutaneous fat of the selected leg in the non-suppressed image. In short, these physics-based models [17] guide the deformation of a parametric contour toward the boundary between the subcutaneous fat and the muscle under the influence of image-based, internal and external forces seeking energy minimization. The authors in [17] defined an energy functional of the form

$E_{total} = \int_0^1 [E_{int}(v(s)) + E_{im}(v(s)) + E_{ext}(v(s))] ds$, for a parametric closed curve representation $v(s) = (x(s), y(s))$ with $s \in [0, 1]$ and $(x, y) \in \Omega$, where $\Omega \subseteq \mathbb{Z}^2$ is the image spatial domain. The image-based energy E_{im} is computed from the image gradient magnitude so that the snake is attracted to image edges, hence $E_{im}(v(s)) = -|\nabla(G_\sigma * I)|^2$, where G_σ denotes a Gaussian kernel with σ standard deviation. The internal energy E_{int} penalizes stretching and bending, $E_{int}(v(s)) = (1/2)(\alpha|v'(s)| + \beta|v''(s)|)$, where α, β are the penalty weights. We used a balloon external force E_{ext} to guide the snake from the internal SAT surface to its muscle interface. The minimizing snake is the solution of the Euler equation that is found numerically using a discrete system [17]. Here we implemented active contour segmentation in a multi-scale hierarchical scheme. That is, starting from a coarser scale an approximate solution is found that is used as initial condition for finer scale segmentation. Following this strategy we achieved faster convergence.

Next, we removed the bone region as follows: we first applied multiple Otsu thresholding to identify the air, fat and muscle groups on the non-suppressed image and then ranked the regions according to their circularity to select the bone marrow. Finally, we used morphological operations with adaptive recursion to label voxels previously unclassified due to the partial volume effect between bone marrow and cortical bone as cortical bone.

III. Tissue Distribution Models

We hypothesize that statistical models of MR muscle and IMAT voxel intensity distributions can define a predictive descriptor of fatty infiltration in skeletal muscle. We first modeled the joint WS and FS intensities by a 2-component Gaussian mixture model (*GMM*) corresponding to muscle and IMAT. We then computed the eigenstructure of joint water- and fat- suppressed signal intensity distributions from MRI data. The parametric probability models together with the computed eigenvectors and eigenvalues define a multi-parametric descriptor of fat infiltration in mid-thigh muscle.

The normal density parametric model of a voxel intensity sample X in the joint intensity domain $[FS, WS]$ for each tissue type denoted by N_{tissue} is defined as

$$N_{tissue}(X|\omega_{tissue}, M_{tissue}, C_{tissue}) = \frac{1}{(2\pi)^{D/2} |C_{tissue}|^{1/2}} \cdot \exp(-1/2(X - M_{tissue})^T C_{tissue}^{-1} (X - M_{tissue})) \quad (1)$$

where $tissue \in \mathcal{T} = \{muscle, IMAT\}$, ω_{tissue} denotes each tissue class,

$M_{tissue} = [\mu_{tissue}^{FS}, \mu_{tissue}^{WS}]^T$, $C_{tissue} = [\sigma_{tissue}^{(i,j)}]$ with $(i, j) \in \{WS^2, WS \times FS, FS \times WS, FS^2\}$ and $D = 2$ is the dimensionality of $[FS, WS]$ domain. The muscle and IMAT distributions produce the observed Gaussian mixture

$$p(X|M_{tissue}, C_{tissue}) = \sum_{tissue \in \mathcal{T}} [a_{tissue} \cdot N_{tissue}(X|\omega_{tissue}, M_{tissue}, C_{tissue})]. \quad (2)$$

We used the Expectation-Maximization (E-M) algorithm to fit the Gaussian mixture model and find $N_{tissue}(X|\omega_{tissue}, M_{tissue}, C_{tissue})$ and a_{tissue} , $\forall tissue \in \mathcal{T}$. We note that $a_{tissue} =$

$P(\omega_{tissue})$, where $P(\omega_{tissue})$ denotes the prior probability of each tissue. Given the tissue density models N_{tissue} and priors $P(\omega_{tissue})$ we classified each voxel into the IMAT or muscle group using the likelihood ratio-based discriminant function

$g(X) = \frac{N_{muscle}(X|\omega_{muscle}, M_{muscle}, C_{muscle})}{N_{IMAT}(X|\omega_{IMAT}, M_{IMAT}, C_{IMAT})} - \frac{P(\omega_{IMAT})}{P(\omega_{muscle})}$ according to the Bayesian framework. The estimated tissue distribution models correspond to ellipsoids in the two-dimensional sample domain.

In the next step we aim to obtain a transformed feature space in which to define measures of compactness of computed tissue distribution models under the hypothesis that the dispersion of samples is caused by fat infiltration into the muscle. A solution in the least-square sense is obtained by finding the unit vectors $\mathbf{u}_{tissue}$ that maximize $\mathbf{u}_{tissue}^T C_{tissue} \mathbf{u}_{tissue}$, subject to the constraint $\|\mathbf{u}_{tissue}\| = 1$ [16] with $tissue \in \mathcal{T} = \{muscle, IMAT\}$. Using Lagrangian multiplier optimization, with λ_{tissue} being the undetermined multiplier, we seek to optimize

$$L(\mathbf{u}_{tissue}, \lambda_{tissue}) = \mathbf{u}_{tissue}^T C_{tissue} \mathbf{u}_{tissue} - \lambda_{tissue} (\mathbf{u}_{tissue}^T \mathbf{u}_{tissue} - 1). \quad (3)$$

By setting $\frac{\partial L}{\partial \mathbf{u}_{tissue}} = 0$, we arrive at the eigenvalue problem

$$C_{tissue} \mathbf{u}_{tissue} = \lambda_{tissue} \mathbf{u}_{tissue}. \quad (4)$$

After solving (4), the covariance matrices C_{tissue} of muscle and IMAT can be factored as

$$C_{tissue} = \sum_{i=1}^D \hat{\lambda}_{tissue}^i \hat{\mathbf{u}}_{tissue}^i \hat{\mathbf{u}}_{tissue}^{i^T}, \quad (5)$$

where $\hat{\lambda}_{tissue}^i$, $\hat{\mathbf{u}}_{tissue}^i$ are the computed eigenvalues and eigenvectors for each tissue density model. The eigenvectors define the orientations of the ellipsoids approximating the tissue density distributions in the joint $[FS, WS]$ domain. The corresponding eigenvalues are proportional to the length of the principal axes. Moreover, the determinants of these covariance matrices are equal to the product of the principal axis lengths

$$|C_{tissue}| = \prod_{i=1}^D \hat{\lambda}_{tissue}^i \geq 0. \quad (6)$$

Figure 3 helps to explain the rationale behind the proposed tissue distribution modeling. This figure illustrates the tissue intensity distributions of three thigh MRI samples corresponding to low, medium and high muscle quality (MQ) reference measurements computed as the ratio of muscle strength to muscle volume. Muscle strength is a functional variable obtained by isokinetic dynamometry and muscle volume is computed from supervised muscle delineation and anthropometric measures as explained in the experiments section. We display the scatterplots of labeled muscle and IMAT samples after E-M clustering in the joint domain of WS and FS image intensities, the corresponding tissue

distribution models, and the scatterplot of the muscle samples overlaid on the muscle distribution model using an enlarged vertical and horizontal scale for visualization purposes.

We draw three central conclusions from this analysis. First, we note that the lengths of both principal axes increase with decreasing muscle quality. We attribute this to the infiltration of skeletal muscle with fat - a major contributor to muscle quality decline - that increases the inhomogeneity of muscle intensity distribution. If a larger proportion of muscle mass is fat, then a lower proportion of the mass consists of contractile proteins, which explains the low muscle quality. Second, the orientation of principal axes differs somewhat between the different levels of muscle quality because of varying correlation between the sample components (sometimes caused by intensity inhomogeneities or other quality artifacts), so it may be beneficial to identify the dispersion along principal axes rather than in the original domain of FS and WS intensities. Equivalently, this method performs decomposition of the voxel intensities into subspaces of reduced inter-variable correlation, thus producing more accurate estimation. Third, this approach is amenable to MRI analysis because it computes relative distances in the feature space of standardized intensities as described in Section II. This representation is therefore not dependent on the absolute MRI voxel intensities, which may vary between different scans.

IV. Tissue Characterization

In this stage we developed a machine learning system to classify subjects into high and low muscle quality based on MRI-derived features produced by the tissue quantification and distribution modeling stages.

Our technique produced the following features; i) tissue distribution model parameters including the means, covariances, and eigen-space parameters of muscle and inter-muscular adipose tissue described in Section III, ii) $\{NS, FS, WS\}$ scaled intensity means as described in Section II, computed over the muscle, subcutaneous fat, inter-muscular adipose tissue, and bone regions, and iii) average areas over the ROI, and area fraction measurements of the muscle, subcutaneous fat, inter-muscular adipose tissue, and bone. The resulting dimensionality of the joint domain is 40. The complete list of features can be found in Appendix A.

Starting with the full feature domain, we applied a dimensionality reduction stage. We first implemented a Fisher's distance ranking-based feature selection approach that we denote by *FDR*. This method computes the Fisher's distance (or F-score) for each feature. Greater distances correspond to features that produce increased separation; therefore we next sorted distances in descending order and selected the highest scores according to a user-defined threshold. In addition, we employed the hierarchical dimensionality reduction technique [16] denoted by *HDR*. This method computes the correlation between all feature pairs and iteratively merges the pair of maximum correlation, thus reducing redundancy in the feature space. Both *FDR* and *HDR* can be considered correlation-based methods, but *FDR* selects the more discriminant features after univariate tests, whereas *HDR* reduces the redundancy by removing features that are very highly correlated with other features in our space. The third approach that we utilized was the minimum redundancy-maximum relevance (*MRMR*)

feature selection method [18]. This method uses mutual information to measure the entropy produced by each feature set and appropriately selects the set of features that minimizes the redundancy and maximizes the relevance.

We utilized support vector machines [19] to build our classification models. Our decision was mainly driven by the wide applicability and good performance of support vector machines in variable classification problems. We tested the hypothesis that our image-based analysis can characterize muscle quality by cross-validation of our classifier against labeled samples of high and low muscle quality.

V. Experiments and Discussion

We validated the performance of our descriptor against reference biomechanical muscle quality measurements computed from clinical variables collected in the same study. Maximum quadriceps muscle strength was defined as the highest value of torque from the left leg in up to three consecutive measures of concentric knee extensor strength at an angular velocity of 0.52rad/s using an isokinetic dynamometer (Kin-Com model 125E, ver. 3.2; Chattanooga Group, Chattanooga, TN). Cross-sectional thigh muscle area was measured from mid-thigh CT slices with 10mm slice thickness acquired on a Somatom CT scanner (Somatom Sensation 10; Siemens, Malvern, PA), and quantified using a supervised quantification workflow (Geanie software, ver. 2.1; BonAlyze, Jyvaskyla, Finland). The

muscle quality measure was defined as $MQ = \frac{T_{KE}}{CSA_{MTM} \cdot H}$, where T_{KE} is the knee extensor torque, CSA_{MTM} is the mid-thigh muscle cross-sectional area, and H is the participant's height [2].

To determine a threshold for separating the subjects between low and high muscle quality, we first divided the MQ range of our measurements into 100 intervals, trained and tested our classifier using the complete set of features in a repeated 10-fold cross-validation experiment, and plotted the true positive rate (TPR) and true negative rate (TNR) for each MQ value as displayed in Figure 4. We then selected as threshold between the low and high muscle quality subjects the point of intersection between TPR and TNR for minimal classification error. The resulting threshold is $63.1 (N \cdot m/cm^3)$.

We performed tissue quantification over a subset of MRI data acquired on an Achieva 3T scanner (Philips Healthcare, Best, The Netherlands) for the Baltimore Longitudinal Study of Aging (BLSA), which is an ongoing, prospective study of longitudinal physical and cognitive changes associated with aging and age-related diseases [20]. Three scans were acquired sequentially using a 3D multi-slice FFE pulse sequence, with parameters: repetition delay $TR = 25\text{ms}$, echo time $TE = 3.8\text{ms}$, and flip angle 25° . The first scan was acquired without fat or water suppression while the second and third scans were acquired with water- and fat-suppression, respectively, using the SPIR (selective pulse inversion recovery) technique provided on the Philips platform. The entire thigh region was imaged with 80 5-mm -thick slices with in-plane resolution of $0.98 \times 0.98\text{mm}^2$ and the field of view was $440 \times 300 \times 400\text{mm}^3$. The participants remained at rest before and during MRI acquisition according to the imaging protocol.

In our experiments we defined cube-shaped ROIs with the same orientation as the volumetric MRI. The ROI centerpoint is determined to be the mid-point between the femur head and patella of the left leg, which was also the centerpoint of the volumetric MRI during acquisition. The superior-inferior length of ROI was fixed at $25mm$. The ROI in-plane dimensions are equal to the MRI in-plane field of view noted in the previous paragraph. An ROI example is displayed in Figure 2. We applied our tissue identification and characterization techniques to the part of each MRI scan defined by these ROIs. Furthermore, in order to reduce the effect of gender differences on our analysis, we selected a cohort of female participants. Our final dataset consisted of 64 female subjects with age 67.44 ± 3.15 (95% CI) years.

To estimate the group separation produced by each feature, we first calculated p -values produced by t-tests and Fisher's distances. Table I contains the list of statistically significant features with $p < 0.05$, the corresponding p -values, and the Fisher's distances. Out of the 14 features with statistically significant p -values, 11 originated from the proposed tissue density modeling in the original or the transformed feature space, but none of these features was related to absolute tissue area measurements. Table II displays the pairs of highly correlated features that were iteratively merged to perform dimensionality reduction according to *HDR*. This technique identified increased correlation between covariance and eigenvalue features, and intensity means and *GMM* parameters, as expected. The limitation of *HDR* is that it does not explicitly seek the most discriminant features. Table III lists the features selected by *MRMR* in our classification experiments. Out of the 12 selected features, 5 features originated from our intensity distribution model parameters either in the standardized or the transformed domain, 5 features were mean estimates in the standardized intensity domain, and the remaining 2 were the mean IMAT area computed over the ROI - that is a volumetric measure - and the muscle area fraction, which is defined as the fraction of mean muscle area over the sum of the mean muscle area and mean IMAT area. It is interesting that the last two tissue area-related features have also been identified as possible predictors of muscular degeneration.

In addition, we trained and tested a support vector machine-based classifier to separate low muscle quality from high muscle quality participants in the same feature domain. Classification accuracy was estimated in a stratified 10-fold cross-validation framework. The stratification was employed to ensure that the number of subjects assigned to each class was approximately equal in the training and testing phases. Furthermore, the 10-fold cross-validation technique reduced the possible effect of over-fitting. Each cross-validation experiment was repeated 100 times to reduce variability due to randomized fold selection. Moreover, we repeated our above experiments using leave-one-out cross-validation. Classification performance metrics produced by the different domains are summarized in Tables IV and V.

To estimate classification performance we used the following measures derived from the contingency table.

True Positive Rate (*TPR*), also called Recall, or Sensitivity – Ratio of true positives TP to the number of detected positives P given by $TPR = \frac{TP}{P}$.

True Negative Rate (*TNR*), also called Specificity – Ratio of true negatives *TN* to the number of negatives *N* given by $TNR = \frac{TN}{N}$.

Classification Accuracy (*ACC*) – Ratio of sum of true positives and true negatives to the total number of positives and negatives: $ACC = \frac{TP+TN}{P+N}$.

Precision (*Pr*) – Ratio of true positives to the sum of true positives and false positives $Pr = \frac{TP}{TP+FP}$. This equals the Positive Predictive Value.

G_{mean} – Geometric mean of *TPR* and *TNR*, $G_{mean} = \sqrt{TPR \cdot TNR}$.

$F_{measure}$ – Calculated according to $F_{measure} = \frac{2}{\frac{1}{Pr} + \frac{1}{TPR}}$.

The repeated 10-fold cross-validation results in Table IV-presented as mean $\pm$ 95% confidence interval half width - indicated that our analysis technique can separate low from high muscle quality subjects with an accuracy of $71.4 \pm 0.6\%$ in the complete feature space. Among the reduced feature spaces, the *MRMR* method produced the highest classification accuracy at $70.9 \pm 0.9\%$. The hierarchical dimensionality reduction technique with a correlation threshold at 95%, which iteratively merged features with greater than 95% correlation, yielded classification accuracy of $69.3 \pm 0.7\%$ in a domain reduced to 29 features. Feature selection by Fisher's distance ranking using a threshold of 0.35 produced accuracy of $69.2 \pm 0.6\%$ and dimensionality $D = 7$. We experimentally determined that an F-score threshold of 0.35 indicates reasonable separation between the two compared groups. Other classification metrics that balance the true positive and true negative rates, such as G_{mean} and the $F_{measure}$, in general produce similar results as *ACC*. Similarly to our univariate tests, features selected by *MRMR* were mostly related to *FS* intensity averages, covariances in the [*FS*, *WS*] domain, and associated eigen-components.

The leave-one-out cross-validation experiments yielded some interesting results (Table V). The *MRMR* technique produced the highest classification accuracy at 78.1%, followed by the complete feature domain and Fisher's distance ranking at 73.4%, and hierarchical dimensionality reduction at 71.9%. We also computed ROC performance for the leave-one-out cross-validation, as displayed in Figure 5. We note that *MRMR* produced the largest area under the curve (*AUC*) among the tested feature selection approaches. The superior performance of *MRMR* may be attributed to the fact that it simultaneously reduces the redundancy and optimizes for relevance of the feature set as explained in [18].

We observe that the sets of more discriminant features selected by the Student's t-test and *MRMR* dimensionality reduction, shown in Tables I and III, lend credence to our premise that the proposed tissue intensity distribution models contribute to the prediction of low muscle quality. Also, our feature selection and classification results reported in Tables III, IV and V, suggest that the MR intensity standardization using SAT reference intensities largely suppresses the effects of inter-subject variability on the dynamic range of MR voxel intensities, therefore producing discriminant features. Nevertheless, the standardized intensities may still change when different MRI acquisition parameters, such as flip angle and *TR*, are used. In addition, our multi-variate classification experiments indicated an

encouraging separation between groups of low and high muscle quality, further supporting the relationship between characteristics of muscle composition and function.

The fact that the proposed methodology is automated, is by itself an advantage that greatly reduces the need for manual delineation of tissues and quantification. When considering the computational cost of this approach, the computational time for segmentation and feature calculation over an ROI of 5 slices is $56.6 \pm 0.8 \text{ sec}$. The feature selection and prediction procedures at the on-line stage -that is, excluding the off-line training- are very fast: $10^{-4} \pm 3.5 \times 10^{-6} \text{ sec}$ for the complete domain, $6 \times 10^{-4} \pm 2.7 \times 10^{-5} \text{ sec}$ for *FDR*, $2.4 \times 10^{-3} \pm 2.8 \times 10^{-5} \text{ sec}$ for *HDR*, and $1.1 \times 10^{-3} \pm 2.2 \times 10^{-5} \text{ sec}$ for *MRMR*. These computational times were obtained from our Matlab implementation on a Linux-based platform with an 8-core Intel Xeon x86 64 CPU at 3.3 GHz and with 8 GB of RAM. We used the Matlab wrapper of the libsvm library for support vector machine classification. The segmentation code can be significantly optimized for time and space once it is implemented in a compiled language such as C/C++.

As an alternative means of acquiring distinct fat and water images for determination of muscle quality, one could employ the so-called Dixon technique [21], [22], [23]. Dixon sequences achieve fat-water separation based on the chemical shift difference between fat and water and the consequent modulations of the echo train generated by a multiecho sequence, and do not rely upon chemical species suppression as in the current study. The current study was performed in the absence of this specialized capability, but nevertheless demonstrates our ability to perform muscle quality assessments and correlations. These outcome measures are essentially independent of whether the method we chose, or a Dixon method, was used for data acquisition.

Although the assessment of muscle quality could be estimated in clinical practice by simpler measures such as DEXA and hand grip, those simpler measures do not provide any clue on the origin and pathophysiology of the decline in muscle quality with aging and diseases. On the contrary, particular imaging characteristics that are associated with lower muscle quality may help understanding the underlying mechanisms of the decline in muscle quality with aging, and therefore suggest possible interventions, and provide biomarkers for tracking their effectiveness. For example, the findings of this study suggest that muscle fat infiltration is one non-exhaustive mechanism that leads to reduced muscle quality. Thus, intervention that prevents muscle fat infiltration is likely to be effective in preventing the decline in muscle quality with aging. The proposed analysis could also be utilized to track the effect of interventions that may theoretically improve muscle quality, such as resistance exercise. It is well known that resistance training in the first week is associated with substantial increment in strength without any change in mass. It would be interesting in the future to verify that the descriptor of muscle quality computed by this technique captures the changes in the muscle over the first week that account for the increase in strength.

We should also note that the MRI techniques used for our analysis are sensitive to body composition, but do not provide data on other contributors to muscle quality decline, such as denervation and bioenergetic alterations, due to either mitochondrial dysfunction or metabolic changes.

VI. Conclusion

We have presented an MRI-based muscle and IMAT characterization model that utilizes information produced by tissue quantification. Our central hypothesis is that we can use the MRI-based muscle quality signatures to predict biomechanical properties of the mid-thigh, namely the muscle quality index MQ . In our method, we estimate statistical properties of $[NS, FS, WS]$ MR intensity distributions with emphasis on Gaussian mixture models of muscle and IMAT voxel intensities in the joint $[FS, WS]$ domain and the eigendecomposition of the covariance matrices. To test our hypothesis, we validated the separation capability of our technique against reference data produced by independent clinical measures of muscle strength. Our technique produced encouraging separation between low and high muscle quality subjects.

Acknowledgments

This research was supported by the National Institute of General Medical Sciences of the National Institutes of Health (NIH) under Award Number SC3GM113754 and by the Intramural Research Program of National Institute on Aging, NIH. We also acknowledge the support of the Center for Research and Education in Optical Sciences and Applications (CREOSA) of Delaware State University funded by NSF CREST-8763. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Goodpaster BH, Carlson CL, Visser M, Kelley DE, Scherzinger A, Harris TB, Stamm E, Newman AB. Attenuation of skeletal muscle and strength in the elderly: The health ABC study. *J Appl Physiol*. Jun; 2001 90(6):2157–2165. [PubMed: 11356778]
2. Moore AZ, Caturegli G, Metter EJ, Makrogiannis S, Resnick SM, Harris TB, Ferrucci L. Difference in muscle quality over the adult life span and biological correlates in the Baltimore Longitudinal Study of Aging. *Journal of the American Geriatrics Society*. Feb; 2014 62(2):230–236. [PubMed: 24438020]
3. Foley, DJ.; Berkman, LF.; Branch, LG.; Farmer, ME.; Wallace, RB. Established Populations for Epidemiologic Studies of the Elderly: Resource Data Book. 1986. NIH Publ. 86-2443
4. Janssen I, Shepard DS, Katzmarzyk PT, Roubenoff R. The healthcare costs of sarcopenia in the United States. *J Am Geriatr Soc*. Jan; 2004 52(1):80–85. [PubMed: 14687319]
5. Lynch NA, Metter EJ, Lindle RS, Fozard JL, Tobin JD, Roy TA, Fleg JL, Hurley BF. Muscle quality. i. Age-associated differences between arm and leg muscle groups. *J Appl Physiol*. Jan; 1999 86(1):188–194. [PubMed: 9887130]
6. Goodpaster BH, Park SW, Harris TB, Kritchevsky SB, Nevitt M, Schwartz AV, Simonsick EM, Tylavsky FA, Visser M, Newman AB. The loss of skeletal muscle strength, mass, and quality in older adults: The Health, Aging and Body Composition study. *J Gerontol A Biol Sci Med Sci*. Oct; 2006 61(10):1059–1064. [PubMed: 17077199]
7. Delmonico MJ, Harris TB, Visser M, Park SW, Conroy MB, Velasquez-Mieyer P, Boudreau R, Manini TM, Nevitt M, Newman AB, Goodpaster BH. Health, Aging, and Body. Longitudinal study of muscle strength, quality, and adipose tissue infiltration. *Am J Clin Nutr*. Dec; 2009 90(6):1579–1585. [PubMed: 19864405]
8. Schick F, Machann J, Brechtel K, Stempfer A, Klumpp B, Stein DT, Jacob S. MRI of muscular fat. *Magn Reson Med*. Apr; 2002 47(4):720–727. [PubMed: 11948733]
9. Positano V, Christiansen T, Santarelli MF, Ringgaard S, Landini L, Gastaldelli A. Accurate segmentation of subcutaneous and inter-muscular adipose tissue from MR images of the thigh. *J Magn Reson Imaging*. Mar; 2009 29(3):677–684. [PubMed: 19243051]
10. Karampinos DC, Baum T, Nardo L, Alizai H, Yu H, Carballido-Gamio J, Yap SP, Shimakawa A, Link TM, Majumdar S. Characterization of the regional distribution of skeletal muscle adipose

tissue in type 2 diabetes using chemical shift-based water/fat separation. *J Magn Reson Imaging*. Apr; 2012 35(4):899–907. [PubMed: 22127958]

11. Makrogiannis S, Serai S, Fishbein KW, Schreiber C, Ferrucci L, Spencer RG. Automated quantification of muscle and fat in the thigh from water-, fat-, and nonsuppressed MR images. *Journal of Magnetic Resonance Imaging*. May; 2012 35(5):1152–1161. [PubMed: 22170747]
12. Ruan XY, Gallagher D, Harris T, Albu J, Heymsfield S, Kuznia P, Heshka S. Estimating whole body intermuscular adipose tissue from single cross-sectional magnetic resonance images. *J Appl Physiol*. Feb; 2007 102(2):748–754. [PubMed: 17053107]
13. Yoshizumi T, Nakamura T, Yamane M, Islam AH, Menju M, Yamasaki K, Arai T, Kotani K, Funahashi T, Yamashita S, Matsuzawa Y. Abdominal fat: standardized technique for measurement at CT. *Radiology*. Apr; 1999 211(1):283–286. [PubMed: 10189485]
14. Tustison NJ, Avants BB, Cook PA, Zheng Yuanjie, Egan A, Yushkevich PA, Gee JC. N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*. Jun; 2010 29(6):1310–1320. [PubMed: 20378467]
15. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. Oct; 2002 17(2):825–841. [PubMed: 12377157]
16. Duda, RO.; Hart, PE.; Stork, DG. *Pattern Classification*. Wiley-Interscience Publication; 2000.
17. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*. Mar; 1998 7(3):359–369. [PubMed: 18276256]
18. Ding CHQ, Peng H. Minimum redundancy feature selection from microarray gene expression data. *J Bioinformatics and Computational Biology*. Feb; 2005 3(2):185–206. [PubMed: 15852500]
19. Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov*. Jun; 1998 2(2):121–167.
20. Shock, NW.; Greulich, RC.; Costa, PT.; Andres, R.; Lakatta, EG.; Arenberg, D.; Tobin, JD. *Normal Human Aging: The Baltimore Longitudinal Study on Aging*. NIH Publication; Washington, D.C: 1984.
21. Hernando, Diego; Liang, Zhi-Pei; Kellman, Peter. Chemical shift-based water/fat separation: a comparison of signal models. *Magn Reson Med*. Sep; 2010 64(3):811–822. [PubMed: 20593375]
22. Karampinos, Dimitrios C.; Yu, Huanzhou; Shimakawa, Ann; Link, Thomas M.; Majumdar, Sharmila. T-corrected fat quantification using chemical shift-based water/fat separation: application to skeletal muscle. *Magn Reson Med*. Nov; 2011 66(5):1312–1326. [PubMed: 21452279]
23. Triplett, William T.; Baligand, Celine; Forbes, Sean C.; Willcocks, Rebecca J.; Lott, Donovan J.; DeVos, Soren; Pollaro, Jim; Rooney, William D.; Lee Sweeney, H.; Bnnemann, Carsten G.; Wang, Dah-Jyuu; Vandenborne, Krista; Walter, Glenn A. Chemical shift-based mri to measure fat fractions in dystrophic skeletal muscle. *Magn Reson Med*. Jul; 2014 72(1):8–19. [PubMed: 24006208]

Appendix A. Complete Feature Domain

Feature Group	Acronyms
Areas	M-AA: Muscle Area Average SAT-AA: Subcutaneous Adipose Tissue Area Average IMAT-AA: Inter-Muscular Adipose Tissue Area Average B-AA: Bone Area Average
Area Fractions	SAT-AF: Subcutaneous Adipose Tissue Area Fraction M-AF: Muscle Area Fraction (scaled using M and IMAT) M-AF2: Muscle Area Fraction (scaled using M, IMAT and SAT)

Feature Group	Acronyms
	IMAT-AF: Inter-Muscular Adipose Tissue Area Fraction
Standardized Intensities	<p>M-NS-SIA: Muscle Non-Suppressed Intensity Average</p> <p>SAT-NS-SIA: Subcutaneous Adipose Tissue Non-Suppressed Scaled Intensity Average</p> <p>IMAT-NS-SIA: Inter-Muscular Adipose Tissue Non-Suppressed Scaled Intensity Average</p> <p>B-NS-SIA: Bone Non-Suppressed Scaled Intensity Average</p> <p>M-FS-SIA: Muscle Fat-Suppressed Scaled Intensity Average</p> <p>SAT-FS-SIA: Subcutaneous Adipose Tissue Fat-Suppressed Scaled Intensity Average</p> <p>IMAT-FS-SIA: Inter-Muscular Adipose Tissue Fat-Suppressed Scaled Intensity Average</p> <p>B-FS-SIA: Bone Fat-Suppressed Scaled Intensity Average</p> <p>M-WS-SIA: Muscle Water-Suppressed Scaled Intensity Average</p> <p>SAT-WS-SIA: Subcutaneous Adipose Tissue Water-Suppressed Scaled Intensity Average</p> <p>IMAT-WS-SIA: Inter-Muscular Adipose Tissue Water-Suppressed Scaled Intensity Average</p> <p>B-WS-SIA: Bone Water-Suppressed Scaled Intensity Average</p>
Intensity Distributions in [FS, WS] Domain	<p>M-FS-EMA: Muscle Fat-Suppressed E-M Average</p> <p>M-WS-EMA: Muscle Water-Suppressed E-M Average</p> <p>M-FSFS-COV: Muscle Fat-Suppressed-Fat-Suppressed Covariance</p> <p>M-FSWS-COV: Muscle Fat-Suppressed-Water-Suppressed Covariance</p> <p>M-WSWS-COV: Muscle Water-Suppressed-Water-Suppressed Covariance</p> <p>IMAT-FS-EMA: Inter-Muscular Adipose Tissue Fat-Suppressed E-M Average</p> <p>IMAT-WS-EMA: Inter-Muscular Adipose Tissue Water-Suppressed E-M Average</p> <p>IMAT-FSFS-COV: Inter-Muscular Adipose Tissue Fat-Suppressed-Fat-Suppressed Covariance</p> <p>IMAT-FSWS-COV: Inter-Muscular Adipose Tissue Fat-Suppressed-Water-Suppressed Covariance</p> <p>IMAT-WSWS-COV: Inter-Muscular Adipose Tissue Water-Suppressed-Water-Suppressed Covariance</p>
Intensity Distributions in Transformed Domain	<p>M-CM-EVE11: Muscle Cov. Matrix EigenVector (1, 1)</p> <p>M-CM-EVE12: Muscle Cov. Matrix EigenVector (1, 2)</p> <p>M-CM-EVA1: Muscle Cov. Matrix EigenValue (1)</p> <p>M-CM-EVA2: Muscle Cov. Matrix EigenValue (2)</p> <p>M-CM-DET: Muscle Cov. Matrix Determinant</p> <p>IMAT-CM-EVE11: Inter-Muscular Adipose Tissue Cov. Matrix EigenVector (1, 1)</p> <p>IMAT-CM-EVE12: Inter-Muscular Adipose Tissue Cov. Matrix EigenVector (1, 2)</p> <p>IMAT-CM-EVA1: Inter-Muscular Adipose Tissue Cov. Matrix EigenValue (1)</p>

Feature Group	Acronyms
	IMAT-CM-EVA2: Inter-Muscular Adipose Tissue Cov. Matrix EigenValue (2) IMAT-CM-DET: Inter-Muscular Adipose Tissue Cov. Matrix Determinant

Fig. 1. Example of a non-suppressed volumetric MR image from our dataset, centered at the mid-thigh region. The tissue identification and characterization technique is applied to the designated mid-thigh region.

Fig. 2.

Top: Fat-suppressed (FS) and water-suppressed (WS) images that we use in our method. The region of interest (ROI) is outlined by the red box that can be seen in orthoplanar and 3D views. It is centered at mid-thigh and it covers the whole cross-sectional region of the left thigh. Bottom: Segmented muscle, subcutaneous adipose tissue (SAT) and inter-muscular adipose tissue (IMAT) regions produced by our tissue quantification algorithm. In this visualization we removed the anterior part of the thigh to display all internal tissues and their identified regions.

$MQ \left(\frac{N \cdot m}{cm^3} \right)$

NS MRI

Scatterplots

Tissue Models

Muscle Model

37.07 (low)

69.73 (medium)

92.64 (high)

Fig. 3.

(From left to right) Examples of MRI scans at the mid-thigh site for subjects exhibiting low, medium, and high muscle quality (MQ); scatterplots show muscle and IMAT voxel intensities in WS and FS scans; tissue classification and estimated muscle and fat distribution models in the $[WS, FS]$ domain; scatterplots and distribution models of the muscle group only.

Fig. 4.

True positive rate (*TPR*) and true negative rate (*TNR*) versus muscle quality (*MQ*) for classification using all features and repeated 10-fold cross-validation. The threshold for generating classes of low and high muscle quality is defined as the point of intersection of *TPR* and *TNR*.

Fig. 5.

Receiver-operating-curve (ROC) performance estimation for leave-one-out cross-validation of different feature selection methods. We note that MRMR produces the largest area-under-the-curve among the tested approaches.

TABLE I

List of features with statistically significant p -values and corresponding Fisher's distances.

Feature	p-value	Fisher's distance
M-FS-SIA	0.000299	0.479
IMAT-FS-SIA	0.00056	0.455
B-FS-SIA	0.0031	0.387
M-FS-EMA	0.0054	0.361
IMAT-FS-EMA	0.0045	0.371
M-FSFS-COV	0.034	0.273
IMAT-FSFS-COV	0.0085	0.351
IMAT-FSWS-COV	0.030	0.279
M-CM-EVE11	0.020	0.315
M-CM-EVE12	0.027	0.289
M-CM-EVA1	0.034	0.272
IMAT-CM-EVE12	0.041	0.260
IMAT-CM-EVA1	0.0085	0.351
IMAT-CM-DET	0.030	0.282

TABLE II

Pairs of features merged by the *HDR* algorithm.

Feature 1	Feature 2
IMAT-AA	M-AF
M-CV-EVA1	M-FSFS-COV
IMAT-CV-EVA1	IMAT-FSFS-COV
M-WS-EMA	M-WS-SIA
IMAT-WS-EMA	IMAT-WS-SIA
IMAT-CV-EVE12	IMAT-CV-EVE11
M-CV-EVE12	M-CV-EVE11
M-WSWS-COV	M-WS-EMA
IMAT-FS-EMA	M-FS-EMA
IMAT-FS-SIA	M-FS-SIA
IMAT-WS-EMA	IMAT-NS-SIA

TABLE III

List of features selected by the *MRMR* algorithm.

Feature
IMAT-FS-SIA
IMAT-AA
B-NS-SIA
IMAT-WSWS-COV
M-AF
IMAT-CM-EVE12
M-CM-EVE11
IMAT-FSFS-COV
M-WS-SIA
B-FS-SIA
B-WS-SIA
M-CM-EVE12

TABLE IV

Estimates of true positive rate (TPR), true negative rate (TNR), classification accuracy (ACC), precision Pr , G_{mean} , and $F_{measure}$ for separation between low and high muscle quality subjects using repeated 10-fold cross-validation. D represents the dimensionality (number of features) considered in each classification scheme.

Feature Domain	Complete	FDR	HDR	MRMR
D	40	7	29	12
$TPR(\%)$	69.6 ± 0.8	55.7 ± 0.6	69.0 ± 0.9	68.2 ± 1.1
$TNR(\%)$	73.1 ± 0.9	82.6 ± 1.0	69.5 ± 1.0	73.6 ± 1.4
$ACC(\%)$	71.4 ± 0.6	69.2 ± 0.6	69.3 ± 0.7	70.9 ± 0.9
$Pr(\%)$	72.2 ± 0.7	76.5 ± 1.0	69.4 ± 0.8	72.4 ± 1.1
$G_{mean}(\%)$	71.3 ± 0.6	67.8 ± 0.5	69.2 ± 0.7	70.7 ± 0.9
$F_{measure}(\%)$	70.9 ± 0.7	64.4 ± 0.6	69.2 ± 0.7	70.1 ± 0.9

TABLE V

Estimates of true positive rate TPR , true negative rate TNR , classification accuracy ACC , precision Pr , G_{mean} , $F_{measure}$, and the area under the curve (AUC) of ROC for classification of low and high muscle quality subjects using leave-one-out cross-validation. D represents the dimensionality (number of features) considered in each classification scheme.

Feature Domain	Complete	FDR	HDR	MRMR
D	40	7	29	12
$TPR(\%)$	77.4	71.0	74.2	64.5
$TNR(\%)$	69.7	75.8	69.7	90.9
$ACC(\%)$	73.4	73.4	71.9	78.1
$Pr(\%)$	70.6	73.3	69.7	87.0
$G_{mean}(\%)$	73.5	73.3	71.9	76.6
$F_{measure}(\%)$	73.8	72.1	71.9	74.1
AUC	0.694	0.703	0.662	0.767