

Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network

Citation for published version (APA):

van der Heyden, B., van de Worp, W. R. P. H., van Helvoort, A., Theys, J., Schols, A. M. W. J., Langen, R. C. J., & Verhaegen, F. (2020). Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network. *Journal of Applied Physiology*, 128(1), 42-49. <https://doi.org/10.1152/jappphysiol.00465.2019>

Document status and date:

Published: 01/01/2020

DOI:

[10.1152/jappphysiol.00465.2019](https://doi.org/10.1152/jappphysiol.00465.2019)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

RESEARCH ARTICLE | *Translational Control of Muscle Mass*

Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network

 Brent van der Heyden,^{1*} Wouter R. P. H. van de Worp,^{2*} Ardy van Helvoort,^{2,3} Jan Theys,⁴ Annemie M. W. J. Schols,² Ramon C. J. Langen,^{2*} and Frank Verhaegen^{1*}

¹Department of Radiation Oncology (MAASTRO), GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center⁺, Maastricht, The Netherlands; ²Department of Respiratory Medicine, NUTRIM-School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Center⁺, Maastricht, The Netherlands; ³Health and Science Department, Danone Nutricia Research, Utrecht, The Netherlands; and ⁴Department of Precision Medicine, GROW-School for Oncology and Developmental Biology, Faculty of Health, Medicine and Life Sciences, Maastricht University Medical Center⁺, Maastricht, The Netherlands

Submitted 10 July 2019; accepted in final form 5 November 2019

van der Heyden B, van de Worp WR, van Helvoort A, Theys J, Schols AM, Langen RC, Verhaegen F. Automated CT-derived skeletal muscle mass determination in lower hind limbs of mice using a 3D U-Net deep learning network. *J Appl Physiol* 128: 42–49, 2020. First published November 7, 2019; doi:10.1152/jappphysiol.00465.2019.—The loss of skeletal muscle mass is recognized as a complication of several chronic diseases and is associated with increased mortality and a decreased quality of life. Relevant and reliable animal models in which muscle wasting can be monitored noninvasively over time are instrumental to investigate and develop new therapies. In this work, we developed a fully automatic deep learning algorithm for segmentation of micro cone beam computed tomography images of the lower limb muscle complex in mice and subsequent muscle mass calculation. A deep learning algorithm was trained on manually segmented data from 32 mice. Muscle wet mass measurements were obtained from 47 mice and served as a data set for model validation and reverse model validation. The automatic algorithm performance was ~150 times faster than manual segmentation. Reverse validation of the algorithm showed high quantitative metrics (i.e., a Dice similarity coefficient of 0.93, a Hausdorff distance of 0.4 mm, and a center of mass displacement of 0.1 mm), substantiating the robustness and accuracy of the model. A high correlation ($R^2 = 0.92$) was obtained between the computed tomography-derived muscle mass measurements and the muscle wet masses. Longitudinal follow-up revealed time-dependent changes in muscle mass that separated control from lung tumor-bearing mice, which was confirmed as cachexia. In conclusion, this deep learning model for automated assessment of the lower limb muscle complex provides highly accurate noninvasive longitudinal evaluation of skeletal muscle mass. Furthermore, it facilitates the workflow and increases the amount of data derived from mouse studies while reducing the animal numbers.

NEW & NOTEWORTHY This deep learning application enables highly accurate noninvasive longitudinal evaluation of skeletal muscle mass changes in mice with minimal requirement for operator involvement in the data analysis. It provides a unique opportunity to increase and analyze the amount of data derived from animal studies automatically while reducing animal numbers and analytical workload.

* B. van der Heyden and W. R. P. H. van de Worp contributed equally to this article and R. C. J. Langen and F. Verhaegen contributed equally to this article.

Address for reprint requests and other correspondence: R. Langen, Dept. of Respiratory Medicine, Maastricht Univ. Medical Center⁺, P.O. Box 5800, 6202 AZ Maastricht, the Netherlands (e-mail: r.langen@maastrichtuniversity.nl).

artificial intelligence; muscle segmentation; μ CBCT

INTRODUCTION

Skeletal muscle is the largest tissue in the body and comprises ~40% of the total body mass in normal weight humans (13). It allows voluntary movement, including mobility and breathing, but also contributes to numerous metabolic processes, including substrate disposal and thermogenesis (9, 25). Skeletal muscle is one of the most dynamic and plastic tissues of the body, responding to changes in functional or metabolic demand, most obviously in response to increased or reduced physical activity, resulting in either hypertrophy or atrophy (27). However, these characteristics also enable skeletal muscle to be used as an energy and protein reserve that can be deployed during periods of starvation or during acute and chronic catabolism, as seen during trauma or infections and cancer cachexia, leading to accelerated muscle loss (30).

Loss of skeletal muscle mass, especially in the locomotor muscles, is a recognized complication of several chronic diseases, including chronic obstructive pulmonary disease, chronic heart failure, and chronic kidney disease, and of most malignant cancers (35). Muscle wasting is well known to be associated with increased patient mortality and decreased quality of life (1, 21, 28). As muscle wasting is an unmet medical need, the development of effective therapeutic strategies involves preclinical research, which requires reliable assessment of muscle mass in animal models.

In the clinical setting, several imaging modalities, such as computed tomography (CT), cone beam CT (CBCT), dual-energy X-ray absorptiometry (DEXA), or MRI, have been applied for years as practical and precise methods to assess body composition and skeletal muscle volumes (3, 4, 12, 14, 15, 17, 20, 26, 29). In contrast, skeletal muscle mass evaluation in preclinical studies is mostly based on assessment of muscle wet masses or muscle fiber cross-sectional area by postmortem histological examination. Both methods require muscle dissection and, consequently, are terminal experiments. This precludes longitudinal follow-up of muscle mass changes, and assessment of time-dependent changes in muscle mass therefore requires inclusion of multiple animal groups in the exper-

imental design. Recently, novel approaches to noninvasively determine muscle volume, including changes over time of individual experimental animals, using MRI and micro-CT were described (5, 6, 19, 23, 24).

Micro-CT is a dedicated X-ray-based imaging technology for small animals that allows for tissue density quantifications, which makes micro-CT a good image technique for longitudinal muscle mass assessment. To investigate muscle mass changes quantitatively, manual segmentation of the muscle volume on the CT reconstruction of the experimental animal is required, which in itself is a delicate and very time-consuming task for the researcher.

Many efforts have been made to develop automated CT segmentation techniques for patients, but, surprisingly, there have only been sparse reports on fully automatic soft tissue CT segmentation techniques in experimental animals. Baiker et al. (2) investigated an automatic reference atlas-based whole body segmentation method in micro-CT images using a digital mouse phantom. More recently, our group investigated organ-at-risk segmentation in mice before irradiation with a multi-atlas image deformation-based approach (31). Clearly, availability of automated muscle segmentation applications for mice would tremendously facilitate CT-based assessment of skeletal muscle mass in preclinical research. Moreover, it will strongly contribute to a reduction in the number of animals used for longitudinal studies that evaluate muscle mass changes, which is in line with the 3R principle embedded in national and international legislation and regulations on the protection of animals used for scientific purposes. To our knowledge, no literature exists on the fully automatic segmentation of skeletal muscle in mice. In this work, we developed and validated a new method to automatically segment skeletal muscle in mice using artificial intelligence techniques that could be widely applied in animal models used to study muscle atrophy or hypertrophy.

MATERIALS AND METHODS

A summarizing flowchart of the study setup is shown in Fig. 1. Further details are described in the subsequent paper sections.

Mouse Experimental and Imaging Procedures

Animal preparation. This work was conducted in accordance with institutional guidelines for the care and use of laboratory animals established by the Ethics Committee for Animal Experimentation of Maastricht University, in full compliance with national legislation on animal research following the European Directive 2010/63/EU for the use of animals for scientific purposes, and is part of a set of experiments to establish an orthotopic model of lung cancer cachexia (AV01070020174168). Male and female mice, at 9–10 wk old (129S2/SvPasCrl, Charles River Laboratories), were socially housed in a climate-controlled room (12:12 dark-light cycle with a constant room temperature of $21 \pm 1^\circ\text{C}$). Mice were given ad libitum access to food (AIN-93M) and water. After 1 wk of acclimatization, mice were randomly allocated to either sham control ($n = 15$) or tumor-bearing (TB) group ($n = 68$). All animals underwent surgery at the age of 12 wk at a standardized time window during their inactive period of the day. While anesthetized using a mixture of air and isoflurane (4% induction, 2% maintenance) and appropriate analgesia, sham control mice received an intrapulmonary injection with 15 μL of matrix (Matrigel, Corning), and the TB mice (orthotopic lung tumor model) received an intrapulmonary injection with 15 μL of Matrigel containing lung epithelium-derived adenocarcinoma cells (10). Mice were monitored daily. Based on an animal welfare scoring list or upon a loss of body mass $>20\%$, humane end points were applied. Furthermore, tumors were scored for characteristics, size, and progression. When a total tumor volume of 500 mm^3 was reached or in the case of signs of dyspnea because of tumor growth, pulmonary constriction, or metastasis, humane end points were applied.

Animal imaging. At baseline and weekly after surgery, μCBCT imaging was performed for all mice to assess lung tumor development (data not shown) and detect muscle volume changes over time. For this, mice were anesthetized as outlined above, placed in prone position with toes facing the flanks (foot and tibia angle $\pm 90^\circ$), and scanned using a μCBCT scanner (X-RAD 225Cx, Precision X-Ray Inc., North Branford, CT) at an X-ray tube potential of 50 kVp, X-ray tube current of 5.6 mA, and imaging time of 2 min (34). The imaging dose of 30 cGy was verified using a PTW TN300012 Farmer-type ionization chamber (PTW, Freiburg, Germany) according to the American Association of Physicists in Medicine TG-61 protocol (18). The μCBCT projection data was reconstructed using the pilot Feldkamp back projection algorithm with a voxel dimension of $100 \times 100 \times 100 \mu\text{m}^3$ (33).

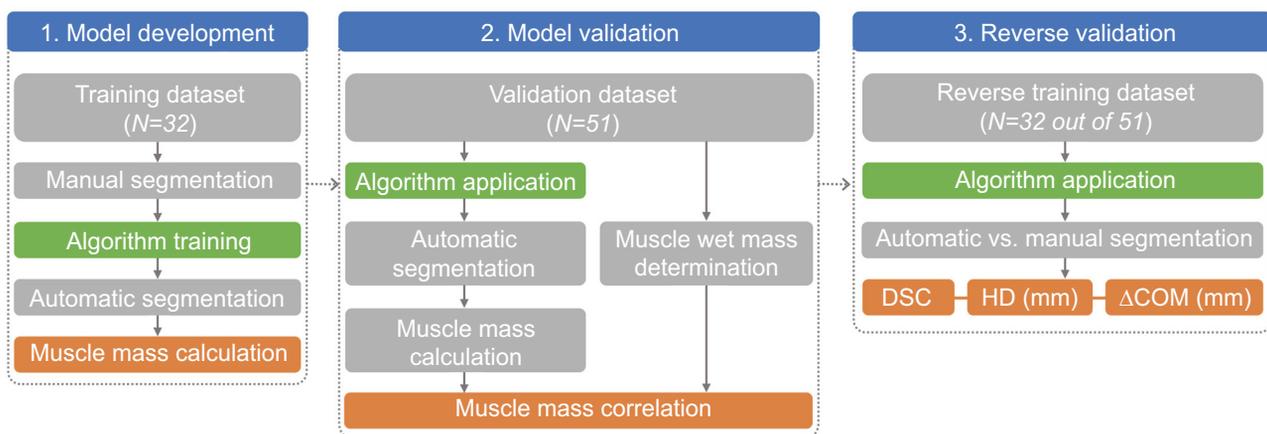


Fig. 1. Study setup flowchart. 1: in model development, the model was trained on a manually segmented data set of 32 mice. After training, the model segmented the lower limb muscle complex on micro cone beam computed tomography (μCBCT) scans and subsequently calculated the muscle mass using a predetermined conversion curve. 2: model validation was finally performed on a validation data set of 47 mice (4 out of 51 mice were excluded). As validation, the muscle mass determined by the model was correlated to muscle wet mass measurements. 3: a reverse validation approach was adopted to calculate three quantitative metrics between the automatic and the manual muscle segmentations. ΔCOM , center of mass displacement; DSC, Dice similarity coefficient; HD, 95th percentile Hausdorff distance.

At the end of the experiment, mice were scanned and subsequently euthanized using pentobarbital overdose to evaluate skeletal muscle mass. The soleus, plantaris, gastrocnemius, tibialis anterior, and extensor digitorum longus muscles including tendons were collected from both hind limbs using standardized dissection methods with a precision coefficient of 0.93 ± 0.05 . Subsequently, muscles were immediately weighed in pairs on an analytical balance with precision of ± 0.1 mg and a linearity of 0.2 mg (CP64, Sartorius, Goettingen, Germany). Baseline μ CBCT scans ($n = 32$) of TB mice were used to train the newly developed algorithm and 47 mice ($n = 14$ sham, $n = 33$ TB) were used to validate the performance of the automatic muscle segmentation algorithm. The 47 mice were included from a group of 51 mice, of which 4 mice met the exclusion criteria (based on deviation of the expected mass-ratio of individual muscles, indicative of imperfections in the dissection procedure). Because of logistic reasons, the time interval between imaging and the muscle collection (day of euthanasia) of the validation group was either 2 days ($n = 24$) or 0 days ($n = 23$).

Automatic Image Segmentation

Deep learning algorithm. As a training data set for our algorithm, the lower limb muscle complex (including the musculus gastrocnemius, soleus, plantaris, tibialis anterior, and extensor digitorum longus) was manually segmented on the reconstructed cross-sectional μ CBCT images that were acquired at the study baseline of 32 mice. The segmentation was performed by an experienced scientist with appropriate anatomical knowledge using the SmART-ATP software (Precision X-Ray Inc.) (32). During the manual segmentation of the training data set, the segmentation time was recorded to calculate the average manual segmentation time for one mouse. Images of five mice were randomly selected from the training data set and were segmented twice with a time interval of 9 mo between the segmentations to investigate the interobserver variability or reproducibility of the manual segmentations.

The deep learning algorithm [two-step three-dimensional (3D) U-Net model] was trained to segment the lower limb muscle complex volume on μ CBCT images [Hounsfield units (HU)] of mice (8, 36). This 3D convolutional neural network architecture consisted of three encoding layers and three decoding layers, used weighted cross entropy as loss function, and used a dropout ratio of 0.5. The neural network was applied in two successive steps and made use of Tensorflow (Python 2.7) in combination with the NVIDIA CUDA Deep Neural Network library (cuDNN) computational kernels. The algorithm was executed on a NVIDIA Quadro P6000 (24 GB) graphics processing unit (GPU).

An automatic preprocessing step was performed on the μ CBCT mice data set before the images were used by the deep learning algorithm. The μ CBCT volume was first cropped in the transversal plane to a $[256, 256, z]$ full resolution, where z is the number of reconstructed μ CBCT slices. Next, the full resolution μ CBCT volume was resized with a cubic interpolation method to x , y , and z dimensions of $[128, 128, z/2]$. A similar method, with nearest-neighbor interpolation, was applied on the manually segmented muscle volumes. In the final preprocessing step, both the down-sampled and full resolution μ CBCT volumes were normalized between -400 HU and $1,000$ HU.

In the first step, a 3D U-Net model was trained (350 epochs) to segment the muscle volume on the down-sampled and preprocessed μ CBCT images. The relative position of the automatically segmented muscle volume with respect to the down-sampled μ CBCT image was then used to extract a volume-of-interest with preset dimensions ($64 \times 64 \times 192$ pixels) in the full resolution μ CBCT data set. In the second step (350 epochs), the extracted full resolution volume-of-interest was used to train the second 3D U-Net.

Reverse validation. To evaluate the robustness and the consistency of the automatic segmentation algorithm, a reverse validation ap-

proach was applied. First, the “forward” automatic segmentation algorithm that was trained on the manually segmented data set (32 mice) was applied on a validation data set (51 mice). Here, the exclusion criterion was not adopted because the segmentations are not subject to experimental errors such as in the muscle wet mass experiments. In the reverse validation approach, a new model was trained based on the already automatically segmented muscle volumes. Here, the automatically segmented muscle volumes of 32 mice were randomly selected from the complete validation data set (51 mice) to have a similarly sized trained data set compared with the original forward segmentation algorithm. After training the “reverse” model, this model was applied to segment the training data set consisting of manual ground truth segmentations.

Finally, the automatic segmentation from the reverse algorithm was compared with the manual segmentation using three quantitative parameters, including the Dice similarity coefficient (DSC), 95th percentile Hausdorff distance (HD; in mm) and the center of mass displacement (Δ COM; in mm). The DSC indicates the volumetric overlap between the manual and the automatic muscle segmentation, the HD calculates the maximum distance of a point in the manual segmentation to the nearest point in the automatic segmentation considering the voxel dimensions, and the Δ COM calculates the 3D displacement of the mass centers of the manual and automatic segmentation in a Cartesian coordinate system.

μ CBCT to mass density conversion. A μ CBCT to mass density (CT2MD) conversion curve was required to calculate mass densities in every reconstructed μ CBCT voxel. This CT2MD curve was obtained by scanning a cylindrical mini-phantom (diameter = 3 cm, length = 1 cm) (SmART Scientific Solutions BV, Maastricht, the Netherlands), which is composed of a solid water bulk, 2 air inserts, and 10 tissue-mimicking inserts of 3.5-mm diameter with known mass densities (e.g., adipose 0.95 g/cm³, soft tissue 1.06 g/cm³, and bone inserts 1.33 – 1.83 g/cm³). The mean CT numbers in HUs were calculated in circular regions-of-interest in the middle of the inserts and were then related to the certified mass densities of the cylindrical inserts in the form of a CT2MD conversion curve.

To determine muscle mass using μ CBCT imaging, the CT2MD conversion curve is applied on the reconstructed μ CBCT mice data set (HUs) to obtain a three-dimensional density matrix (g/cm³). After applying the CT2MD curve on the reconstructed μ CBCT mice data sets, the binary masks resulting from our automatic muscle segmentation algorithm was applied on the converted mass density volume, and the knowledge of the reconstructed voxel dimensions was used to calculate the muscle mass noninvasively. An image intensity-based thresholding technique was applied on the μ CBCT image data set to remove bone from the analysis.

RESULTS

Reverse Validation and Segmentation Reproducibility

Three quantitative metrics (DSC, HD, and Δ COM) ± 1 SD were calculated to evaluate the accuracy of the reverse automatic segmentation algorithm. On average (± 1 SD), 158 ± 8 axial μ CBCT slices were automatically segmented by the segmentation algorithm.

The DSC (± 1 SD) was equal to 0.93 ± 0.03 , the HD (± 1 SD) was equal to 0.4 ± 0.2 mm and the Δ COM (± 1 SD) was equal to 0.1 ± 0.1 mm, indicating a good agreement compared with the manual ground truth muscle segmentations. To investigate the manual segmentation reproducibility of the human observer, these metrics were also calculated between two manual segmentations of five mice. Here, the DSC was equal to 0.95 ± 0.01 , the HD was equal to 0.4 ± 0.1 mm, and the Δ COM was equal to 0.18 ± 0.03 mm.

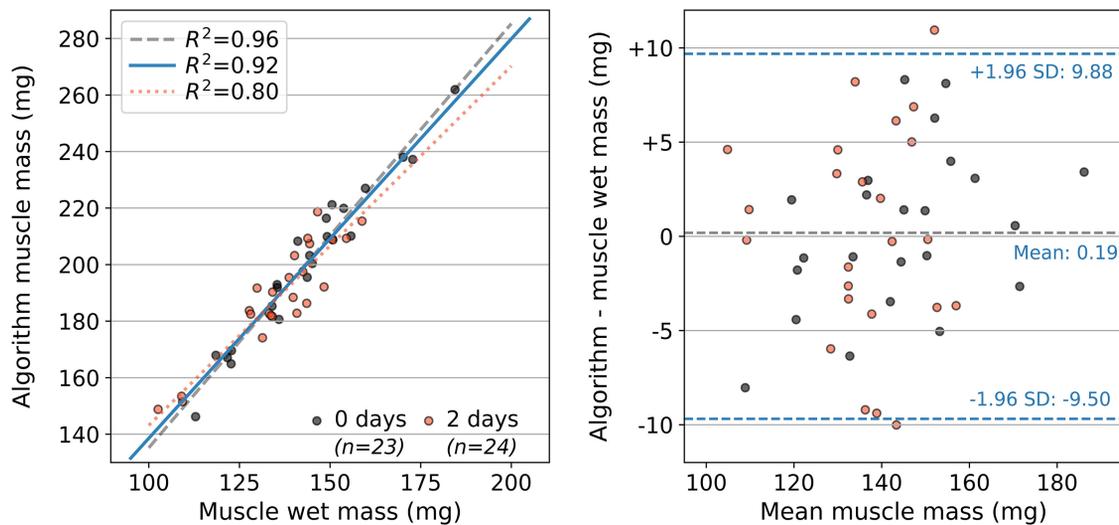


Fig. 2. *Left*: linear relationships between the muscle wet mass measurements and the muscle masses determined by the automatic segmentation algorithm ($R^2 = 0.96$, $y = 1.50x - 14.85$: subset weighted after 0 days; $R^2 = 0.92$, $y = 1.42x - 2.76$: subset weighted after 2 days; $R^2 = 0.80$, $y = 1.27x + 16.04$). *Right*: Bland-Altman plot between the algorithm muscle mass and the experimental muscle wet mass.

CT-Derived Muscle Mass Highly Correlates with Muscle Wet Masses

The combined muscle wet masses of the lower limb were determined in a validation set consisting of 47 mice and compared with the muscle masses calculated by the automatic segmentation algorithm (Fig. 2, *left*). A trendline (blue) was fitted through all data points by linear regression ($R^2 = 0.92$). When separated for the time span between μ CBCT scan and muscle wet mass assessment, the correlation further increased for the 23 mice for which muscle tissue was collected directly after imaging ($R^2 = 0.96$), compared with 24 mice of which muscle tissue was excised 2 days after imaging ($R^2 = 0.80$).

The ratio between the muscle wet mass and the algorithm-based muscle mass was a constant factor of 0.72 ± 0.03 and can be attributed to the excision of a select set muscles from the hind limb musculature that can excised with high accuracy. After factor correction of the algorithm-defined muscle mass, a Bland-Altman plot was constructed to analyze the agreement between the two methods (Fig. 2, *right*). No proportional bias was found, meaning that the two methods agree equally through the range of measurements.

Longitudinal Assessment of Muscle Mass

To study muscle mass changes over time and detect the onset of muscle wasting, longitudinal evaluation of muscle mass is essential. To evaluate the automatic follow-up capabilities of our algorithm for individual mice, muscle masses were assessed for repetitive analyses of 6 randomly selected mice (3 sham control mice and 3 TB mice) in the validation data set. The muscle masses shown in Fig. 3 were calculated relatively to the automatic measurement at baseline (*day 0*). These results show that the algorithm is capable of distinguishing muscle mass responses between sham control mice and TB mice developing muscle wasting.

One TB mouse and one sham control mouse were randomly chosen to evaluate the automatic contouring with manual contouring. In Supplemental material, animations in GIF format are provided to visualize the automatic muscle segmenta-

tion and the manual muscle segmentation of these two mice in the axial, coronal, and sagittal μ CBCT viewing planes (all Supplemental material is available at <https://doi.org/10.6084/m9.figshare.8832392.v2>).

Figure 4 shows the automatic muscle segmentation (orange contour) and the manual muscle segmentation (yellow contour) on the μ CBCT scan of these mice at the baseline measurement and at the last timepoint of the study. Because the mice were scanned at different days under different positioning conditions, a rigid deformation was applied with the Elastix software for visualization purposes (16). For these two mice, the DSCs, HDs, and Δ COMs were calculated between the automatic and the manual segmentation at the first and last scan, respectively. For the control mouse, the DSCs were equal to 0.94 and 0.95, the HDs were equal to 0.4 mm and 0.3 mm, and the Δ COMs were equal to 0.3 mm and 0.2 mm, respectively. For the TB

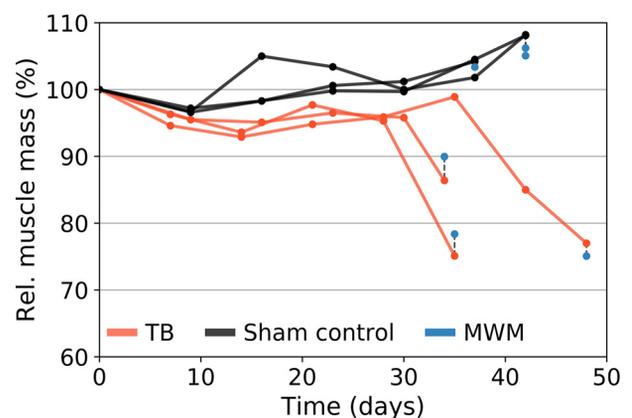


Fig. 3. Follow-up curves determined by our automatic muscle segmentation algorithm for three tumor-bearing (TB) mice (orange) and three sham control mice (black). The orange and black markers indicate the time of scanning, and the lines that connect the markers are linearly interpolated. The blue markers present the experimental muscle wet mass (MWM) at the last timepoint after applying the fitted linear relationship between the MWM and the algorithm muscle mass. The absolute (calculated) muscle mass at the baseline measurement varied slightly (time = 0 days, 220.3 ± 18.1 mg).

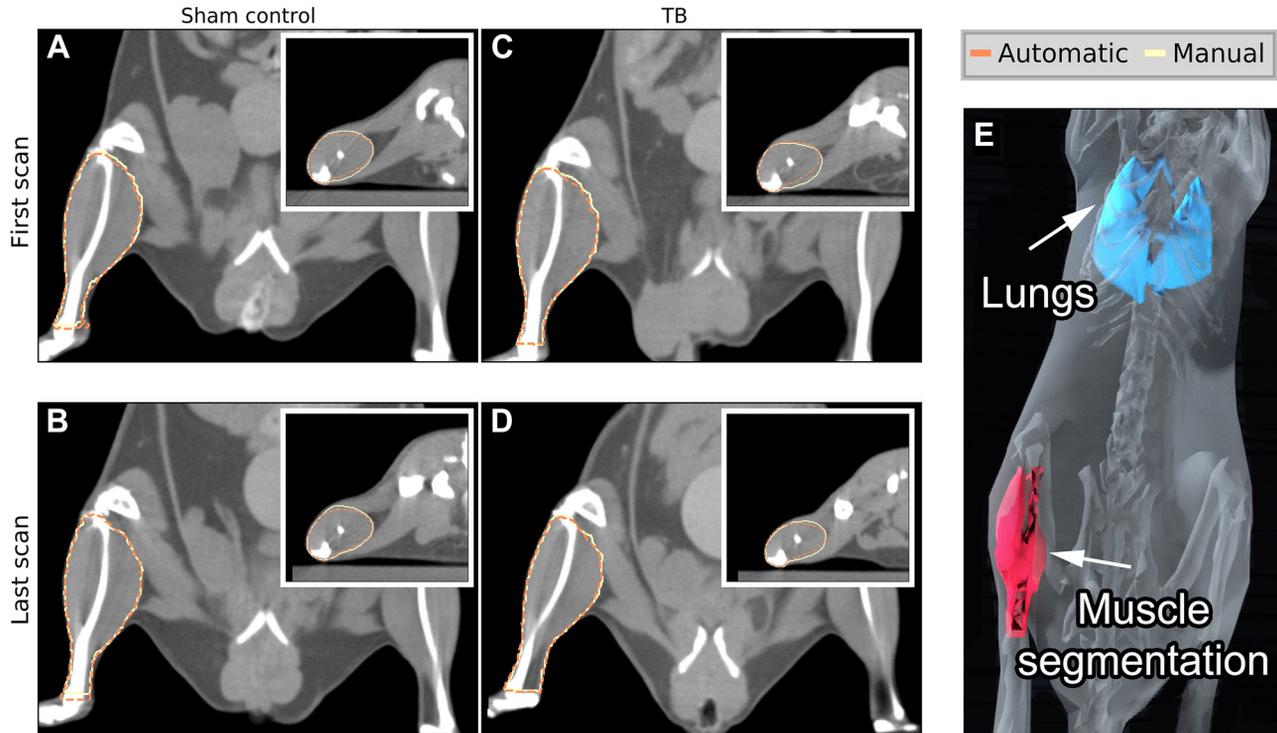


Fig. 4. Automatic (orange) and manual (yellow) hind limb muscle segmentation on a micro cone beam computed tomography (μ CBCT) slice for a sham control mouse and a tumor-bearing (TB) mouse for the scan at the baseline measurement (A and C) and the last timepoint (B and D). Axial μ CBCT slices are shown as insets in A–D. A three-dimensional model (E) of an example muscle segmentation is shown as illustration. In Supplemental material, animations of all cross-sectional μ CBCT slices are provided for both mice.

mouse, the DSCs were equal to 0.95 and 0.94, the HDs were equal to 0.3 mm and 0.4 mm, and the Δ COMs were equal to 0.2 mm and 0.4 mm, respectively.

Muscle Segmentation Algorithm

The segmentations generated by the forward two-step, 3D U-Net-based muscle segmentation algorithm were used to compare against muscle wet mass measurements. The first step of the 3D U-Net was trained in 5 h and 31 min, and the second step was trained in 6 h and 44 min, resulting in a total duration of 12 h and 15 min. Performing the two-step, 3D U-Net-based automatic segmentation algorithm required an average calculation time of 6 s for one mouse, whereas the average manual segmentation time for one mouse was 17 min. Consequently, compared with manual segmentation, the machine-based automatic method reduces the time needed to segment hind limb skeletal muscle by more than a factor of 150.

DISCUSSION

In this paper, we present a new method to noninvasively assess hind limb muscle mass in mice for cross-sectional and longitudinal purposes. It employs a deep learning algorithm trained to segment the lower muscle complex volume on μ CBCT images. This algorithm or model showed a high correlation with the actual muscle wet mass measurements. Furthermore, this deep learning application enables longitudinal evaluation of skeletal muscle mass changes in experimental mouse models, significantly reducing animal numbers and analytical workload.

Biomedical image segmentation methods are commonly trained on manually segmented data sets and validated against manual segmentations on a validation data set. After automatic segmentation, statements about the algorithm's performance are then based on specific parameters (e.g., DSC, HD, or Δ COM) compared with the manual segmentation. Although we calculated these three quantitative parameters solely for a random sham control mouse and a random TB mouse at two timepoints, we found a good agreement with the manual segmentations. The interobserver variability and reproducibility of the manual segmentations evaluated in the image data sets of five mice showed similar results (DSC = 0.95, HD = 0.4 mm, Δ COM = 0.2 mm) compared with the quantitative metrics calculated between the manual and the automatic segmentation (DSC = 0.95, HD = 0.4 mm, and Δ COM = 0.4 mm).

In a more comprehensive quantitative analysis, we adopted a different approach to investigate whether a model can be trained on a data set that was originally segmented by the algorithm. This means that the result of the forward automatic segmentation algorithm was good enough to train a new deep learning model that in turn can segment μ CBCT images with high accuracy. Using this reverse approach, we achieved mean high DSC scores of 0.93 and the HD and Δ COM both had a submillimeter difference below 0.4 mm.

As experimental model validation, a good linear correlation ($R^2 = 0.92$) was found between the experimental wet muscle masses and the mass determined by our algorithm. The linear relationship was not equal to the unity curve, which can be explained. First, not all soft tissues were extracted for the

muscle wet mass measurements. Excision of the hind limb musculature was restricted to the gastrocnemius, soleus, plantaris, tibialis anterior, and extensor digitorum longus muscles of both legs, as these can be collected intact and in a highly standardized manner. Consequently, remaining muscles, such as the flexor digitorum longus and the ankle plantar flexors and dorsiflexors, which contribute significantly to the lower hindlimb muscle mass (7), contributed to muscle mass calculated by the algorithm but not the muscle wet weight measurements. However, in the training data set and therefore also the automatic muscle segmentation, the whole muscle volume was segmented (in one leg), because the distinction between different muscle types requires enhanced image contrast, which is not feasible with current μ CBCT image quality. The muscle mass determined by our CT segmentation approach is therefore higher, as it represents all muscle mass and not only the set of collected wet muscle masses. The high correlation between these two methods suggests that the relative differences between mice are maintained regardless of whether total lower hindlimb muscle mass or the cumulative wet mass of the standardized collected muscles are used.

Another source of uncertainty was the time between the last scan and the actual wet muscle mass measurements (day of euthanasia), which was ideally as short as possible (0 days), although this was not always possible for logistical reasons. The data extracted from the mice that were scanned at the day of euthanasia showed a clearly higher linear correlation ($R^2 = 0.96$) compared with the mice that were analyzed within 2 days after imaging ($R^2 = 0.80$). The latter may be attributed to continuous changes in muscle mass, including muscle wasting in the TB mice.

Previous studies, such as Ceelen et al. (6), already reported a good linear correlation between the CT-derived mass and the wet muscle mass for mice ($R^2 = 0.85$), but here the CT-based muscle segmentation protocol was performed manually. The correlation in that study was comparable but slightly lower than our linear fit based on the automatically determined CT-derived muscle masses ($R^2 = 0.92$). More recently, Pasetto et al. (24) evaluated muscle atrophy in mouse models using μ CT imaging. However, no CT segmentation was applied in their work; instead, muscle mass was evaluated differently according to two-dimensional derived parameters, including the perpendicular distance from the tibia half-length to the external hind limb muscle margin and the distance from the upper extremity of the tibia to the medial malleolus.

The precision, accuracy, and reproducibility of DEXA in mice has already been investigated for in vivo body parameters, such as total bone mineral density, total body bone mineral, fat mass, and bone-free lean tissue mass (11, 22). In Bunckinx et al. (4a), DEXA was even proposed as the reference standard for measuring muscle mass in patients. Although DEXA provides longitudinal follow-up data of bone-free lean tissue mass with low radiation exposure, this technique cannot distinguish separate muscle groups, such as the lower limb muscle complex, which is possible with the μ CBCT-based algorithm. The use of μ CBCT as imaging modality in TB mice models, in contrast to DEXA, can provide additional longitudinal follow-up information, such as the tumor volume, from the same scan. MRI is an alternative high-resolution imaging technique, providing functional imaging with high soft tissue contrast. MRI does not require an imaging dose, in contrast to

micro-CT, which is an important consideration when radiation dose is a critical factor in the study setup. Although MRI provides a better soft tissue contrast than micro-CT, the latter permits more rapid acquisitions at significantly lower cost. Additionally, it allows for tissue density quantifications, which makes micro-CT a suitable imaging technique for longitudinal muscle mass assessment.

In the muscle wet mass procedure, only muscles that are excisable in a highly standardized way were included, whereas the manual and automatic segmentation methods determined the entire lower hind limb muscle complex volume. The muscle complex was segmented as a whole because the μ CBCT image reconstruction lacks image contrast to distinguish between individual muscle groups. In future studies, multimodality image information, such as μ CBCT and MRI, could be combined with deformable image registration to investigate the segmentation feasibility of individual muscle groups, although the μ CBCT is required to convert image intensities to mass densities.

Lower limb muscle complex segmentation for multiple mice at different time points risks amounting into a time-consuming task in a preclinical study. To illustrate this point, in our experiments, the manual segmentation by an experienced biologist of the lower limb complex on the μ CBCT images required on average 20 min for one mouse. The animals were scanned at 7 different timepoints on average, resulting in 581 μ CBCT scans, which would have taken more than 190 h to segment using manual contouring. In contrast, the method described here required \sim 11 h to segment the training data set manually, around 12 h to train the algorithm on GPU, and only 6 s per automatic muscle segmentation. The latter would even enable real-time follow-up of muscle wasting during the in vivo experiment.

Prior to implementation of the algorithm in other settings, additional validation steps may be required. In this study, the deep learning model was trained on a μ CBCT image data set that was acquired with one specific imaging protocol at an X-ray tube potential of 50 kVp. However, further research is required to investigate whether the trained model is valid in imaging data sets that were acquired at different X-ray tube potential settings (e.g., with an additional scan at 90 kVp in dual-energy μ CBCT) or in imaging data sets that were acquired with different spectral filtrations (33). Our training data set consisted entirely of 129S2/SvPasCrl mice. As such, further investigations are required to evaluate how the trained model will segment the lower limb muscle in mice that have age-, strain-, or pathological model-related differences in muscle size compared with the mice that formed the training data set used in this study.

Nevertheless, we anticipate that the proposed algorithm is capable of segmenting muscle volumes in a variety of animals as long as the study setup (i.e., X-ray acquisition protocols and reconstruction settings) and animal characteristics are consistent over time, i.e., the period in which the training data set is generated and applied to the nonsegmented data set. Therefore, we expect that for every study that uses a different imaging protocol or experimental animal species, the creation of a manually segmented training data set and the retraining of the algorithm is required to achieve the best automatic segmentation outcome.

However, this algorithm, which can be retrained depending on the study setup, will facilitate the workflow tremendously because only the training data set has to be segmented manually.

Involving experimental models using μ CBCT or μ CT imaging in future work will be of interest to create a larger and more diverse training data set of different animal species and settings scanned with different scan protocols and even different scanners to train a model that can be widely applied without the need for (extensive) retraining.

In conclusion, a noninvasive automatic algorithm was developed using artificial intelligence (two-step 3D U-Net) to segment skeletal muscle tissue in the lower limb complex of mice. The performance of the algorithm was in good agreement with the actual muscle wet mass measurements. This experimentally validated algorithm enables highly accurate noninvasive and automated longitudinal evaluation of skeletal muscle mass changes in mice with minimal operator involvement in the data analysis. It provides a unique possibility to collect large amounts of data from the mice and to understand when and how muscle wasting starts and when it becomes life threatening. Furthermore, this method will allow a more accurate planning of experiments and reduce the number of mice needed for longitudinal experiments.

ACKNOWLEDGMENTS

The authors thank Miriam van Dijk (Danone Nutricia Research, the Netherlands) and Natasja Lieuwes (Maastricht University) for their valuable contribution to this work.

DISCLOSURES

The NVIDIA Quadro P6000 GPU used for this research was donated by the NVIDIA Corporation. This research project was cofinanced by Danone Nutricia Research.

AUTHOR CONTRIBUTIONS

B.v.d.H., W.R.P.H.v.d.W., R.C.J.L., and F.V. conceived and designed research; B.v.d.H. and W.R.P.H.v.d.W. performed experiments; B.v.d.H. and W.R.P.H.v.d.W. analyzed data; B.v.d.H., W.R.P.H.v.d.W., R.C.J.L., and F.V. interpreted results of experiments; B.v.d.H. and W.R.P.H.v.d.W. prepared figures; B.v.d.H. and W.R.P.H.v.d.W. drafted manuscript; B.v.d.H., W.R.P.H.v.d.W., A.v.H., J.T., A.M.W.J.S., R.C.J.L., and F.V. edited and revised manuscript; B.v.d.H., W.R.P.H.v.d.W., A.v.H., J.T., A.M.W.J.S., R.C.J.L., and F.V. approved final version of manuscript.

REFERENCES

1. Anker SD, Ponikowski P, Varney S, Chua TP, Clark AL, Webb-Peploe KM, Harrington D, Kox WJ, Poole-Wilson PA, Coats AJ. Wasting as independent risk factor for mortality in chronic heart failure. *Lancet* 349: 1050–1053, 1997. doi:10.1016/S0140-6736(96)07015-8.
2. Baiker M, Milles J, Dijkstra J, Henning TD, Weber AW, Que I, Kaijzel EL, Löwik CW, Reiber JH, Lelieveldt BP. Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data. *Med Image Anal* 14: 723–737, 2010. doi:10.1016/j.media.2010.04.008.
3. Borga M, West J, Bell JD, Harvey NC, Romu T, Heymsfield SB, Dahlqvist Leinhard O. Advanced body composition assessment: from body mass index to body composition profiling. *J Investig Med* 66: 1–9, 2018. doi:10.1136/jim-2018-000722.
4. Breiman RS, Beck JW, Korobkin M, Glenny R, Akwari OE, Heaston DK, Moore AV, Ram PC. Volume determinations using computed tomography. *AJR Am J Roentgenol* 138: 329–333, 1982. doi:10.2214/ajr.138.2.329.
- 4a. Buckinx F, Landi F, Cesari M, Fielding RA, Visser M, Engelke K, Maggi S, Dennison E, Al-Daghri NM, Allepaerts S, Bauer J, Bautmans I, Brandi ML, Bruyere O, Cederholm T, Cerreta F, Cherubini A, Cooper C, Cruz-Jentoft A, McCloskey E, Dawson-Hughes B, Kaufman JM, Laslop A, Petermans J, Reginster JY, Rizzoli R, Robinson S, Rolland Y, Rueda R, Vellas B, Kanis JA. Pitfalls in the measurement of muscle mass: a need for a reference standard. *J Cachexia Sarcopenia Muscle* 9: 269–278, 2018. doi:10.1002/jcsm.12268.
5. Ceelen JJM, Schols AMWJ, Kneppers AEM, Rosenbrand RPHA, Drożdż MM, van Hoof SJ, de Theije CC, Kelders MCJM, Verhaegen F, Langen RCJ. Altered protein turnover signaling and myogenesis during impaired recovery of inflammation-induced muscle atrophy in emphysematous mice. *Sci Rep* 8: 10761, 2018. doi:10.1038/s41598-018-28579-4.
6. Ceelen JJM, Schols AMWJ, van Hoof SJ, de Theije CC, Verhaegen F, Langen RCJ. Differential regulation of muscle protein turnover in response to emphysema and acute pulmonary inflammation. *Respir Res* 18: 75, 2017. doi:10.1186/s12931-017-0531-z.
7. Charles JP, Cappellari O, Spence AJ, Hutchinson JR, Wells DJ. Musculoskeletal geometry, muscle architecture and functional specialisations of the mouse hindlimb. *PLoS One* 11: e0147669, 2016. doi:10.1371/journal.pone.0147669.
8. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. *arXiv* 1606.06650, 2016.
9. Evans DJ, Murray R, Kissebah AH. Relationship between skeletal muscle insulin resistance, insulin-mediated glucose disposal, and insulin binding. Effects of obesity and body fat topography. *J Clin Invest* 74: 1515–1525, 1984. doi:10.1172/JCI11565.
10. Gibbons DL, Lin W, Creighton CJ, Rizvi ZH, Gregory PA, Goodall GJ, Thilaganathan N, Du L, Zhang Y, Pertsemidis A, Kurie JM. Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes Dev* 23: 2140–2151, 2009. doi:10.1101/gad.1820209.
11. Halldorsdottir S, Carmody J, Boozer CN, Leduc CA, Leibel RL. Reproducibility and accuracy of body composition assessments in mice by dual energy x-ray absorptiometry and time domain nuclear magnetic resonance. *Int J Body Compos Res* 7: 147–154, 2009.
12. Heier CR, Gueron AD, Korotcov A, Lin S, Gordish-Dressman H, Fricke S, Sze RW, Hoffman EP, Wang P, Nagaraju K. Non-invasive MRI and spectroscopy of mdx mice reveal temporal changes in dystrophic muscle imaging and in energy deficits. *PLoS One* 9: e112477, 2014. doi:10.1371/journal.pone.0112477.
13. Janssen I, Heymsfield SB, Wang ZM, Ross R. Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *J Appl Physiol* (1985) 89: 81–88, 2000. doi:10.1152/jappl.2000.89.1.81.
14. Karlsson A, Rosander J, Romu T, Tallberg J, Grönqvist A, Borga M, Dahlqvist Leinhard O. Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. *J Magn Reson Imaging* 41: 1558–1569, 2015. doi:10.1002/jmri.24726.
15. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol* 54: 2–10, 2016. doi:10.1016/j.semcdb.2015.09.001.
16. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging* 29: 196–205, 2010. doi:10.1109/TMI.2009.2035616.
17. Levine JA, Abboud L, Barry M, Reed JE, Sheedy PF, Jensen MD. Measuring leg muscle and fat mass in humans: comparison of CT and dual-energy X-ray absorptiometry. *J Appl Physiol* (1985) 88: 452–456, 2000. doi:10.1152/jappl.2000.88.2.452.
18. Ma CM, Coffey CW, DeWerd LA, Liu C, Nath R, Seltzer SM, Seuntjens JP; American Association of Physicists in Medicine. AAPM protocol for 40–300 kV x-ray beam dosimetry in radiotherapy and radiobiology. *Med Phys* 28: 868–893, 2001. doi:10.1118/1.1374247.
19. Marcuzzo S, Zucca I, Mastropietro A, de Rosbo NK, Cavalcante P, Tartari S, Bonanno S, Preite L, Mantegazza R, Bernasconi P. Hind limb muscle atrophy precedes cerebral neuronal degeneration in G93A-SOD1 mouse model of amyotrophic lateral sclerosis: a longitudinal MRI study. *Exp Neurol* 231: 30–37, 2011. doi:10.1016/j.expneurol.2011.05.007.
20. Martin L, Hopkins J, Malietzis G, Jenkins JT, Sawyer MB, Brisebois R, MacLean A, Nelson G, Gramlich L, Baracos VE. Assessment of computed tomography (CT)-defined muscle and adipose tissue features in relation to short-term outcomes after elective surgery for colorectal cancer: a multicenter approach. *Ann Surg Oncol* 25: 2669–2680, 2018. doi:10.1245/s10434-018-6652-x.

21. Mytelka DS, Li L, Benoit K. Post-diagnosis weight loss as a prognostic factor in non-small cell lung cancer. *J Cachexia Sarcopenia Muscle* 9: 86–92, 2018. doi:10.1002/jcsm.12253.
22. Nagy TR, Clair AL. Precision and accuracy of dual-energy X-ray absorptiometry for determining in vivo body composition of mice. *Obes Res* 8: 392–398, 2000. doi:10.1038/oby.2000.47.
23. Park JS, Vohra R, Klussmann T, Bengtsson NE, Chamberlain JS, Lee D. Non-invasive tracking of disease progression in young dystrophic muscles using multi-parametric MRI at 14T. *PLoS One* 13: e0206323, 2018. doi:10.1371/journal.pone.0206323.
24. Pasetto L, Olivari D, Nardo G, Trolese MC, Bendotti C, Piccirillo R, Bonetto V. Micro-computed tomography for non-invasive evaluation of muscle atrophy in mouse models of disease. *PLoS One* 13: e0198089, 2018. doi:10.1371/journal.pone.0198089.
25. Periasamy M, Herrera JL, Reis FCG. Skeletal muscle thermogenesis and its role in whole body energy metabolism. *Diabetes Metab J* 41: 327–336, 2017. doi:10.4093/dmj.2017.41.5.327.
26. Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, Baracos VE. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol* 9: 629–635, 2008. doi:10.1016/S1470-2045(08)70153-0.
27. Schiaffino S, Dyar KA, Ciciliot S, Blaauw B, Sandri M. Mechanisms regulating skeletal muscle growth and atrophy. *FEBS J* 280: 4294–4314, 2013. doi:10.1111/febs.12253.
28. Schols AM, Broekhuizen R, Weling-Scheepers CA, Wouters EF. Body composition and mortality in chronic obstructive pulmonary disease. *Am J Clin Nutr* 82: 53–59, 2005. doi:10.1093/ajcn/82.1.53.
29. Shih R, Wang Z, Heo M, Wang W, Heymsfield SB. Lower limb skeletal muscle mass: development of dual-energy X-ray absorptiometry prediction model. *J Appl Physiol (1985)* 89: 1380–1386, 2000. doi:10.1152/jappl.2000.89.4.1380.
30. Straub RH, Cutolo M, Buttgereit F, Pongratz G. Energy regulation and neuroendocrine-immune control in chronic inflammatory diseases. *J Intern Med* 267: 543–560, 2010. doi:10.1111/j.1365-2796.2010.02218.x.
31. van der Heyden B, Podesta M, Eekers DB, Vaniqui A, Almeida IP, Schyns LE, van Hoof SJ, Verhaegen F. Automatic multiatlas based organ at risk segmentation in mice. *Br J Radiol* 92: 20180364, 2019. doi:10.1259/bjr.20180364.
32. van Hoof SJ, Granton PV, Verhaegen F. Development and validation of a treatment planning system for small animal radiotherapy: SmART-Plan. *Radiother Oncol* 109: 361–366, 2013. doi:10.1016/j.radonc.2013.10.003.
33. Vaniqui A, Schyns LEJR, Almeida IP, van der Heyden B, Podesta M, Verhaegen F. The effect of different image reconstruction techniques on pre-clinical quantitative imaging and dual-energy CT. *Br J Radiol* 92: 20180447, 2019. doi:10.1259/bjr.20180447.
34. Verhaegen F, Dubois L, Gianolini S, Hill MA, Karger CP, Lauber K, Prise KM, Sarrut D, Thorwarth D, Vanhove C, Vojnovic B, Weersink R, Wilkens JJ, Georg D. ESTRO ACROP: Technology for precision small animal radiotherapy research: Optimal use and challenges. *Radiother Oncol* 126: 471–478, 2018. doi:10.1016/j.radonc.2017.11.016.
35. von Haehling S, Anker SD. Prevalence, incidence and clinical impact of cachexia: facts and numbers-update 2014. *J Cachexia Sarcopenia Muscle* 5: 261–263, 2014. doi:10.1007/s13539-014-0164-8.
36. Wang C, MacGillivray T, Yang G, Newby D. A two-stage Unet framework for multi-class segmentation on full resolution image. *arXiv* 1804.04341v1, 2018.

