Deep learning auto-segmentation of cervical skeletal muscle for sarcopenia analysis in patients with head and neck cancer

Mohamed A. Naser¹, Kareem A. Wahid¹, Aaron J. Grossberg², Brennan Olson³, Risham Jain², Dina El-Habashy¹,⁴, Cem Dede³, Vivian Salama¹, Moamen Abobakr¹, Abdallah S. R. Mohamed⁵, Renjie He¹, Joel Jaskari⁵, Jaakko Sahlsten⁵, Kimmo Kaski⁵ and Clifton D. Fuller¹*

¹Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, United States, ²Department of Radiation Medicine, Oregon Health & Science University, Portland, OR, United States, ³Medical Scientist Training Program, Oregon Health & Science University, Portland, OR, United States, ⁴Department of Clinical Oncology, Menoufia University Shibin El Kom, Shibin El Kom, Egypt, ⁵Department of Computer Science, Aalto University School of Science, Espoo, Finland

Background/Purpose: Sarcopenia is a prognostic factor in patients with head and neck cancer (HNC). Sarcopenia can be determined using the skeletal muscle index (SMI) calculated from cervical neck skeletal muscle (SM) segmentations. However, SM segmentation requires manual input, which is time-consuming and variable. Therefore, we developed a fully-automated approach to segment cervical vertebra SM.

Materials/Methods: 390 HNC patients with contrast-enhanced CT scans were utilized (300-training, 90-testing). Ground-truth single-slice SM segmentations at the C3 vertebral level were manually generated. A multi-stage deep learning pipeline was developed, where a 3D ResUNet auto-segmented the C3 section (33 mm window), the middle slice of the section was auto-selected, and a 2D ResUNet auto-segmented the auto-selected slice. Both the 3D and 2D approaches trained five sub-models (5-fold cross-validation) and combined sub-model predictions on the test set using majority vote ensembling. Model performance was primarily determined using the Dice similarity coefficient (DSC). Predicted SMI was calculated using the auto-segmented SM cross-sectional area. Finally, using established SMI cutoffs, we performed a Kaplan-Meier analysis to determine associations with overall survival.

Results: Mean test set DSC of the 3D and 2D models were 0.96 and 0.95, respectively. Predicted SMI had high correlation to the ground-truth SMI in males and females (r>0.96). Predicted SMI stratified patients for overall survival in males (log-rank p = 0.01) but not females (log-rank p = 0.07), consistent with ground-truth SMI.
Introduction

Sarcopenia – the excessive loss of skeletal muscle (SM) mass and function – is a common and debilitating phenomenon in head and neck cancer (HNC) patients (1). Weight loss is frequent in HNC due to nutritional deficiencies induced by tumor geometry affecting normal tissues (2) and/or side effects caused by therapeutic interventions (3). Although the link between treatment-associated weight loss and survival in HNC is unclear (4), sarcopenia has been strongly associated with oncologic outcomes and late radiation-induced toxicities (5–7). Notably, in a recent meta-analysis of HNC patients by Surov et al. (5), sarcopenia was significantly associated with lower overall survival (hazard ratio = 1.64, p < 0.00001) and disease-free survival (hazard ratio = 2.00, p < 0.00001). Therefore, sarcopenia prediction is of paramount importance in patients with HNC.

Sarcopenia can be identified using different diagnostic criteria (8). One quantitative method investigated in various studies is using a threshold based on the skeletal muscle index (SMI), the cross-sectional area of skeletal muscle measured on axial imaging normalized to the square of the patient’s height (9). The SMI is most commonly calculated and referenced using CT imaging of abdominal musculature (10–14). However, abdominal imaging is not available for all HNC patients. Importantly, Swartz et al. (15), van Rijn-Dekker et al. (6), and Olson et al. (16) have recently suggested the C3 cervical vertebra musculature cross-sectional area may also be used to quantify sarcopenia accurately.

Current approaches to generate C3 musculature segmentations needed for SMI calculation rely on either semi-automated or completely manual segmentation (6), which can be time-consuming, introduce unnecessary errors, and suffer from interobserver variability. A fully-automated approach would be an attractive alternative to the current manual/semi-automated standard. Deep learning, which has found success in medical image segmentation (17–20), may be an ideal choice for fully-automated segmentation of SM. Several recent studies have utilized deep learning methods for automated SM measurement based on abdominal CT scans with reasonable performance (21–26). However, to date, no studies have attempted to automate the SMI calculation workflow based on head and neck imaging.

The primary objective of this study was to develop a fully-automated approach to segment skeletal muscle at the C3 vertebral level for use in SMI calculations. These calculations could be directly used to determine sarcopenia status for predicting prognostic outcomes. To achieve this goal, we developed and implemented a two-stage deep learning system that utilizes 3D and 2D ResUNets to detect the C3 vertebra and segment the corresponding C3 musculature, respectively. We show that our approach can faithfully generate segmentations comparable to ground-truth human-generated segmentations. By fully automating the sarcopenia determination workflow, we can ensure rapid, reproducible, and accurate measurements for use in clinical decision-making.

Materials and methods

Patient and imaging data

495 patients from the head and neck squamous cell carcinoma (HNSCC) publicly available dataset collection on The Cancer Imaging Archive (TCIA) (27–29) were retrospectively collected in 2021. All patients had a histopathologically-proven diagnosis of squamous cell carcinoma of the oropharynx and were treated with curative-intent intensity-modulated radiotherapy. DICOM- formatted contrast-enhanced CT scans were acquired from the TCIA databases (27–29). Of the 495 patients available in the HNSCC collection, 396 were selected due to their inclusion of the C3 vertebrae on imaging. Subsequently, 6 patients were removed due to image reconstruction errors (n=1), image processing errors (n=1), or oblique image orientations (n=4), leading to a final set of 390 patients used in this analysis. The clinical and demographic characteristics of these patients are shown in Table 1. The majority of patients were male (86.6%) with base of tongue tumors (51.6%). SM (paraspinal and sternocleidomastoid muscles) was manually segmented for each CT image in one slice (2D image) at the level of the C3 vertebra. The segmentations were performed using sliceOmatic, version 5.0 (Tomovision) using previously published Hounsfield unit thresholds to define muscle and fat (12, 30);

---

**Abbreviations:** DSC, Dice similarity coefficient; HNC, head and neck cancer; ROI, region of interest; SM, skeletal muscle; SMI, skeletal muscle index.
TABLE 1  Clinical demographics of patients whose data were used in this study.

| Characteristic          | Count  |
|-------------------------|--------|
| Age (median, range)     | 57 (28–87) |
| Sex                     |        |
| Male                    | 337    |
| Female                  | 52     |
| Tumor subsite           |        |
| Base of tongue          | 201    |
| Glossopharyngeal sulcus | 9      |
| Soft palate             | 6      |
| Tonsil                  | 157    |
| Not otherwise specified | 16     |
| HPV status              |        |
| Negative                | 36     |
| Positive                | 215    |
| Unknown                 | 138    |
| T-category              |        |
| T1                      | 77     |
| T2                      | 166    |
| T3                      | 91     |
| T4                      | 55     |
| N-category              |        |
| N0                      | 36     |
| N1                      | 44     |
| N2                      | 301    |
| N3                      | 8      |
| AJCC stage (7th ed)     |        |
| I                       | 3      |
| II                      | 12     |
| III                     | 57     |
| IV                      | 317    |

AJCC, American Joint Committee on Cancer. One patient did not have clinical information from The Cancer Imaging Archive so was not included in this table.

Table: Clinical demographics of patients whose data were used in this study.

Specifically, a range of -29 to +150 Hounsfield units was used to initially define SM followed by manual corrections. No pathological tissue was located in the segmented SM. The single-slice 2D CT images selected for segmentation and the corresponding SM segmentation masks were exported as DICOM files and tag files, respectively. Segmentations are made publicly available on Figshare (doi: 10.6084/m9.figshare.18480917); additional information on the dataset used in this analysis can be found in the corresponding data descriptor (31).

Image processing

The DICOM 3D volumetric and single-slice 2D CT images were converted to Neuroimaging Informatics Technology Initiative (NIIfTI) format using the DICOM processing toolkit DICOMRTTool v. 0.3.21 (32). The SM segmentation. tag files were converted to NIIfTI format using an in-house Python script. The NIIfTI files for the single-slice 2D CT images and SM segmentation were used to train the 2D segmentation model (described below). The 2D CT slice location in the C3 vertebra was extracted from the DICOM file, which was then used to generate the ground-truth segmentation mask for the C3 section, defined as a volume 33 mm in thickness centered at the location of the 2D CT slice. The tissue regions in the 3D CT images were distinguished from the background by thresholding the images using a value of greater than -500 Hounsfield units with any air gaps within the tissue region filled to generate a binary mask for the external boundaries. The generated external boundary masks and the locations of the 2D CT slices were used to create the ground-truth C3 section segmentations to train the 3D model (described below). As we have described elsewhere (33), all the images and masks were resampled to a fixed image resolution of 1 mm across all dimensions. The CT intensities were truncated in the range of [-250, 250] Hounsfield units to increase soft tissue contrast and then normalized to the range of [-1, 1] scale (Figures 1A, B). We used the Medical Open Network for AI (MONAI) (34) software transformation functions to rescale and normalize images.

Segmentation model

We used a multi-stage deep learning convolutional neural network approach for SM segmentation. Our approach was based on the U-Net architecture with residual connections (ResUNet) included in the MONAI software package, as we have described in previous publications (33, 35). In the first stage of our approach (Figure 1C), a 3D ResUNet model auto-segmented the C3 vertebra section (33 mm), which was then followed by auto-selection of the middle slice of the section. In the second stage of our approach (Figure 1D), a 2D ResUNet model auto-segmented the SM on the auto-selected slice of the C3 section. Additional details of our architecture are described in Appendix A.

Model implementation

We randomly split the data into 300 patients for training and 90 patients for testing. For training, we used a 5-fold cross-validation approach where the 300 patients from the training data were divided into five non-overlapping sets. Each set (60 patients) was used for model validation while the 240 patients in the remaining sets were used for training, i.e., each set was used once for testing and four times for training, leading to five sub-models. The processed CT and corresponding masks for 3D ResUNet model and 2D ResUNet models (C3 section and SM, respectively) were randomly cropped to four random fixed-sized regions (patches) of size (96, 96, 96) and (96, 96) per patch per patient, respectively. Additional details on the model implementation are described in Appendix A.
implementation are described in Appendix A. We implemented additional data augmentation to both image and mask patches to minimize overfitting, including random horizontal flips of 50% and random affine transformations with an axial rotation range of 12 degrees and a scale range of 10%. We used the Adam optimizer for computing the parameter updates and the soft Dice loss function. The models were trained for 300 iterations with a learning rate of $2 \times 10^{-4}$ for the first 250 iterations and $1 \times 10^{-4}$ for the remaining 50 iterations. The values for the Adam optimizer coefficients $b_1$ and $b_2$ were 0.9 and 0.999, respectively. Data augmentation and loss functions were provided by the MONAI framework (34). The final segmentations on the test set for both models were obtained by a majority vote on a pixel-by-pixel basis for all predicted segmentation masks by the 5-fold cross-validation sub-models (model ensemble), as described in a previous study (33).

Model validation

For both the 3D ResUNet and 2D ResUNet models, we evaluated the performance on the corresponding cross-validation sets as well as the final ensemble segmentation on the test set using the Dice similarity coefficient (DSC) (36). Specific to the 3D model, we also evaluated the accuracy of the C3 section segmentation by quantifying the absolute difference between the slice locations of the mid-section of the C3 section predicted by the 3D model and the 2D CT ground-truth image (in mm). Specific to the 2D model, we compared the SM cross-sectional areas obtained using the SM ground-truth segmentation with 1. the 2D model predicted SM segmentations on the same ground-truth CT image (Pred_GT) and 2. the 2D model predicted SM segmentations on the slices auto-selected by the 3D model (Pred_C3). We evaluated the correlation between the SM cross-sectional areas using the Pearson correlation coefficient; we also used a two-sided Wilcoxon signed-rank test to determine if these SM values were significantly different. Additionally, to derive the SMI, we normalized the SM cross-sectional areas (in cm²) with the patients’ heights (in m²). We then examined the correlation between the SMI values produced by the ground-truth and deep learning segmentations using the Pearson correlation coefficient; we also used a two-sided Wilcoxon signed-rank test to determine if these SMI values were significantly different. Based on previous work by Swartz et al. (15) and van Rijn-Dekker et al. (6), we used Equation 1 to calculate the cross-sectional area.
(CSA) at the L3 lumbar level based on the CSA at the C3 cervical level and subsequently Equation 2 to calculate the lumbar SMI:

\[
\text{CSA at L3 (cm}^2\text{)} = 27.304 + 1.363 \times \text{CSA at C3 (cm}^2\text{)} - 0.671 \times \text{age (years)} + 0.640 \times \text{weight (kg)} + 26.422 \times \text{sex (sex = 1 for female, 2 for male)} + 0.640 \times \text{weight (kg)} + 2.6422 \times \text{sex (sex = 1 for female, 2 for male)}
\]

(Eq.1)

\[
\text{Lumbar SMI (cm}^2\text{m}^{-2}\text{)} = \frac{\text{CSA at L3 (cm}^2\text{)}}{\text{height}^2\text{(m}^2\text{)}}
\]

(Eq.2)

Based on previous work by Prado et al. (30), SMI thresholds of 52.4 cm²/m² (males) and 38.5 cm²/m² (females) were applied to lumbar SMI derived from SM ground-truth and deep learning segmentations to stratify patients by sarcopenia status (‘normal’ and ‘depleted’ muscle); body composition related measurements in the training and testing sets are shown in Appendix B. These stratifications were then used for Kaplan-Meier analysis to determine associations between sarcopenia status and overall survival probabilities. To determine the sarcopenia status for the whole data set (i.e., 390 patients), we implemented Kaplan-Meier analysis on the 5-fold cross-validation data and the test data. We aggregated the SMI estimated for each cross-validation fold (i.e., 60 patients per fold) using the corresponding trained 3D and 2D models in addition to the SMI for the test data using the average predictions of the five cross-validation models.

**Results**

**3D ResUNet model performance: C3 section auto-segmentation**

The performance of the 3D ResUNet model for segmenting the C3 section of the neck is summarized in Figure 2A. When assessing the performance of each individual sub-model from our 5-fold cross-validation, the DSCs calculated between the predicted region segmentations and the ground-truth region segmentations were high and consistent between all training folds, with a mean (± standard deviation) DSC of 0.95 ± 0.01. When combining the cross-validation fold predictions using our ensemble approach, the performance on the test set increased to 0.96 ± 0.06. The middle slices of the predicted 3D regional segmentations for the test set were mostly within 4 mm of the ground-truth segmentation slice locations, with the greatest number of patients being within 1 mm (Figure 2B); the maximum outlier was at a distance of 10 mm.
Examples of test set predictions for cases with low, medium, and high performance compared to the mean DSC are shown in Figures 2C–E. As can be visually confirmed, the low-performance case still generated a segmentation such that the middle slice was contained in the C3 region.

2D ResUNet model performance: SM auto-segmentation

The performance of our 2D ResUNet model for segmenting the C3 vertebra SM is summarized in Figure 3A. The DSCs calculated between the model-predicted segmentations and the ground-truth segmentations were high and consistent between all training folds, with a mean DSC of 0.95 ± 0.002. When combining the cross-validation fold predictions using our ensemble approach, the mean DSC performance on the test set remained consistent at 0.95 ± 0.02.

The cross-sectional areas derived from the 2D model predictions using both the ground-truth slice locations and auto-selected slice locations from the 3D ResUNet model were highly correlated to the cross-sectional areas derived from the ground-truth segmentations (Figure 3B). The predicted areas using the ground-truth slice locations had a Pearson r=0.98 (p < 0.0001) and nonsignificant Wilcoxon test (p=0.43). Similarly, the predicted areas using the auto-selected slice locations had a Pearson r=0.98 (p < 0.0001) and nonsignificant Wilcoxon test (p=0.22). Examples of test set predictions for cases with low, medium, and high performance compared to the mean DSC for predictions using ground-truth slice location are shown in Figures 3C–E. As can be visually confirmed, the low-performance case successfully generated a segmentation for musculature that was not included in the ground-truth segmentation. Moreover, the predictions using the auto-selected slice location from the 3D ResUNet model yielded virtually indistinguishable results for the low-performance and medium-performance cases.
performance cases (Figures 3F, G) and identical results for the high-performance case (Figure 3E). Finally, when investigating the percentage difference in cross-sectional areas between the model-generated and ground-truth segmentations, there was no significant difference when using the ground-truth slice location or the auto-selected slice location (p=0.37) (Figure 3H).

SMI measurement comparisons

We compared SMI values for test set patients calculated using ground-truth SM segmentations with predicted SMI values calculated using SM segmentations generated from our 2D ResUNet models using the ground-truth slice location (Figure 4A) or auto-selected slice location (Figure 4B). Both model SM segmentations led to predicted SMI values that were highly correlated to the ground-truth SMI values. The predicted SMI values using the ground-truth slice location for males and females both had a Pearson r=0.98 (p < 0.0001) and nonsignificant Wilcoxon signed-rank tests (p=0.17 and p=0.43, respectively) compared to ground-truth SMI values. Similarly, the predicted SMI values using the auto-selected slice location for males and females had Pearson r values of 0.97 and 0.96, respectively (both p < 0.0001) and nonsignificant Wilcoxon signed-rank tests (p=0.19 and p=0.98, respectively) compared to the ground-truth SMI values.

Survival analysis

The results of the overall survival analysis based on sarcopenia thresholds are shown in Figure 5. Independent of the method of SMI calculation (GT, Pred_GT, or Pred_C3), there were significant differences in overall survival of males between those with normal and depleted muscle tissue (Figures 5A-C), while females exhibited no significant differences (Figures 5D-F). Hazard ratios (95% confidence intervals) in males for GT, Pred_GT, and Pred_C3 were 1.82 (1.1-3.0), 1.95 (1.18-3.22), and 1.97 (1.19-3.25), respectively. Hazard ratios (95% confidence intervals) in females for GT, Pred_GT, and Pred_C3 were 2.76 (0.59-13.02), 3.4 (0.73-15.83), and 3.72 (0.8-17.31), respectively.

Discussion

In this study, we utilized a multi-stage deep learning approach to segment the C3 region of the head and neck, auto-select a single representative slice, and auto-segment the corresponding SM. Our approach determined slice location and segmented SM with high accuracy when compared to ground-truth segmentations. By fully automating this workflow, we have enabled more rapid testing and application of sarcopenia-related clinical decision-making. To our knowledge, this is the first study to fully automate sarcopenia prediction based on non-abdominal HNC imaging.

We utilized both 2D and 3D ResUNet models in our approach. By decomposing the C3 detection and SM segmentation problem into two separate tasks, we ensure that accurate representations of patient anatomy are identified by the models (C3 region) and subsequently maximize performance for SM segmentation. While previous SM auto-segmentation studies often required specific slices as model inputs (21, 26) or utilized separate pre-processing software (23, 25), multi-stage deep learning methods have recently been adapted in this domain as well (22, 24). Both the 2D and 3D ResUNet models that make up our segmentation pipeline had high performance, with mean DSC values in the test set above 0.95. Importantly, the performance of our C3 SM segmentation model is comparable to that of previous L3 SM deep learning segmentation models, which also demonstrate test set DSCs of ~0.95 (21–26). Moreover, for cases with relatively low performance, we visually confirmed that results were reasonable, i.e., the auto-selected slice was still contained within the C3 region for the 3D model, and the correct musculature was...
segmented on the 2D model. Importantly, we also showed minimal differences in the 2D SM segmentation model regardless of how the slice location was determined, indicating the model is robust to the specific C3 slice location. Consistent with quantitative measures of segmentation performance, using our deep learning segmentations to calculate SMI demonstrated a high correlation with ground-truth SMI independent of gender stratification.

A recent meta-analysis by Surov et al. calculated the cumulative prevalence of sarcopenia in HNC patients at 42% (5), highlighting the clinical need for accurate quantification of sarcopenia. Several previous studies have demonstrated that SMI values can be used to stratify patients into sarcopenia subgroups that are strongly associated with prognostic outcomes (5-7). Using lumbar SMI conversion equations previously derived by Swartz et al. (15) and van Rijn-Dekker et al. (6) combined with validated SMI thresholds (12), we demonstrated that calculations based on our deep learning segmentations predict similar overall survival outcomes as calculations based on ground-truth segmentations. Moreover, p-values for all methods were significant for males but not females. These results are consistent with recent literature by Olson et al. (16) which emphasized that sarcopenia was associated with poor survival outcomes in males but not in females. Our results suggest that our automated methods are dependable for use in prognostic outcome prediction.

While our study presents encouraging results towards full automation of sarcopenia-related clinical decision-making for HNC patients, there were some limitations. First, we only tested our method on pre-therapy images. Importantly, some studies have suggested prognostic evidence for sarcopenia measurements based on alternative or additional timepoints (e.g., body composition changes) (7, 37). Therefore, further confirmatory work is needed to ensure our methods can be used accurately and reproducibly for intra-therapy and post-therapy imaging. Additionally, when defining sarcopenia using SMI cutoffs, we have relied on historically accepted thresholds in literature, but several recent developments in standardizing SMI values, e.g., through body mass index (38), have been proposed that warrant further exploration. We must also note that while no universal consensus on sarcopenia definitions currently exists, European consensus guidelines (39) emphasize the importance of evaluating muscle performance and strength in addition to muscle mass; therefore, by European consensus guidelines we have only investigated “presarcopenia” in this analysis. Moreover, we have limited our analysis to CT images as CT is the most common imaging modality for HNC radiotherapy treatment planning. However, the use of additional modalities for SM segmentation, i.e., MRI, as has been utilized in other studies (40), may warrant additional auto-segmentation investigations. Finally, while we believe current model performance is satisfactory for clinical applications as demonstrated by comparisons with ground-truth segmentations and SMI measures, different architectural choices or ensemble approaches could be further explored to improve performance.

Conclusions

In summary, using open-source toolkits and public data, we applied 3D and 2D deep learning approaches to head and neck CT images to develop an end-to-end automated workflow for SM segmentation at the C3 vertebral level. When evaluated on independent test data, our fully-automated approach yielded mean DSCs of up to 0.96 for segmenting the C3 vertebra region and 0.95 for segmenting the corresponding SM. Cross-sectional areas and calculated SMI values derived from our approach were highly correlated to ground-truth (r>0.95), indicating their potential clinical
acceptability. Moreover, our methods can be reliably combined with validated SMI thresholds for use in prognostic stratification. Our study is an essential first step towards fully-automated workflows for sarcopenia-related clinical-decision making. Future studies should consider incorporating additional imaging timepoints and modalities for automated sarcopenia prediction.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Segmentations generated in this project are available on Figshare, doi: 10.6084/m9.figshare.18480917. The original unprocessed images used in this project can be found on The Cancer Imaging Archive: https://wiki.cancerimagingarchive.net/display/DOI/Radiomics+outcome+prediction+in+Ororopharyngeal+cancer.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Study concepts: MN, KW, AG, BO, RJ, AM, KK, and CF; Study design: MN, KW, AG, BO, RJ, and AM; Data acquisition: AG, BO, RJ, DE-H, CD, VS, and MA; Quality control of data and algorithms: MN, KW, RH, JJ, JS, and KK; Data analysis and interpretation: MN, KW, AG, BO, RH, JJ, and JS; Manuscript editing: MN, KW, AG, BO, RJ, AM, and KK. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Institutes of Health (NIH)/National Cancer Institute (NCI) through a Cancer Center Support Grant (CCSG; P30CA016672-44). MN is supported by an NIH grant (R01DE028290-01). KW is supported by the Dr. John J. Kopchick Fellowship through The University of Texas MD Anderson UTHealth Graduate School of Biomedical Sciences, the American Legion Auxiliary Fellowship in Cancer Research, and an NIH/National Institute for Dental and Craniofacial Research (NIDCR) F31 fellowship (1 F31DE031502-01). AG received funding from the National Cancer Institute (K08 245188, R01 CA264133) and the American Association for Cancer Research/Mark Foundation “Science of the Patient” Award (20-60-51-MARK). BO received funding from the Radiologic Society of North America Research Medical Student Grant (RMS2026). VS received funding from The University of Texas, Graduate School of Biomedical Sciences Graduate research assistantship. CF received funding from the NIH/NIDCR (1R01DE025248-01/R56DE025248); an NIH/NIDCR Academic-Industrial Partnership Award (R01DE028290); the National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679); the NIH Big Data to Knowledge (BD2K) Program of the NCI Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01CA214825); the NCI Early Phase Clinical Trials in Imaging and Image-Guided Interventions Program (1R01CA218148); an NIH/NCI Pilot Research Program Award from the UT MD Anderson CCGS Radiation Oncology and Cancer Imaging Program (P30CA016672); an NIH/NCI Head and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50CA097007); and the National Institute of Biomedical Imaging and Bioengineering (NIBIB) Research Education Program (R25EB025787).

Acknowledgments

The authors acknowledge Sunita Patterson (Research Medical Library, MD Anderson Cancer Center) for editorial assistance.

Conflict of interest

CF has received direct industry grant support, speaking honoraria, and travel funding from Elekta AB.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.930432/full#supplementary-material
38. Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol Off J Am Soc Clin Oncol United States* (2013) 31(12):1539–47. doi: 10.1200/JCO.2012.45.2722

39. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, et al. Sarcopenia: European consensus on definition and diagnosis: Report of the European working group on sarcopenia in older people. *Age Ageing* (2010) 39 (4):412–23. doi: 10.1093/ageing/afq034

40. Zwart AT, Becker J-N, Lamers MJ, Dierckx RA, de Bock GH, Halmos GR, et al. Skeletal muscle mass and sarcopenia can be determined with 1.5-T and 3-T neck MRI scans, in the event that no neck CT scan is performed. *Eur Radiol* Springer (2021) 31(6):4053–62. doi: 10.1007/s00330-020-07440-1