HHS Public Access

Author manuscript

AJR Am J Roentgenol. Author manuscript; available in PMC 2022 April 22.

Published in final edited form as:

AJR Am J Roentgenol. 2022 January ; 218(1): 124–131. doi:10.2214/AJR.21.26486.

Fully Automated Deep Learning Tool for Sarcopenia Assessment on CT: L1 Versus L3 Vertebral Level Muscle Measurements for Opportunistic Prediction of Adverse Clinical Outcomes

Perry J. Pickhardt, MD[1], **Alberto A. Perez, MD**[1], **John W. Garrett, PhD**[1], **Peter M. Graffy, BS, MPH**[1], **Ryan Zea, MS**[1], **Ronald M. Summers, MD, PhD**[2]

[1]Department of Radiology, The University of Wisconsin School of Medicine and Public Health, E3/311 Clinical Science Center, 600 Highland Ave, Madison, WI 53792-3252.

[2]Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD.

Abstract

BACKGROUND.—Sarcopenia is associated with adverse clinical outcomes. CT-based skeletal muscle measurements for sarcopenia assessment are most commonly performed at the L3 vertebral level.

OBJECTIVE.—The purpose of this article is to compare the utility of fully automated deep learning CT-based muscle quantitation at the L1 versus L3 level for predicting future hip fractures and death.

METHODS.—This retrospective study included 9223 asymptomatic adults (mean age, 57 ± 8 [SD] years; 4071 men, 5152 women) who underwent unenhanced low-dose abdominal CT. A previously validated fully automated deep learning tool was used to assess muscle for myosteatosis (by mean attenuation) and myopenia (by cross-sectional area) at the L1 and L3 levels. Performance for predicting hip fractures and death was compared between L1 and L3 measures. Performance for predicting hip fractures and death was also evaluated using the established clinical risk scores from the fracture risk assessment tool (FRAX) and Framingham risk score (FRS), respectively.

Address correspondence to P. J. Pickhardt (ppickhardt2@uwhealth.org).

P. J. Pickhardt has served as an advisor to Bracco, Zebra, and GE Healthcare and is a shareholder in SHINE, Elucent Medical, and Cellectar Biosciences. J. W. Garrett is a shareholder in Nvidia. R. M. Summers receives royalties from iCAD, PingAn, Philips Healthcare, and ScanMed and research support from PingAn and Nvidia. The remaining authors declare that they have no disclosures relevant to the subject matter of this article.

The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as representing the views of the National Institutes of Health Clinical Center.

ARRS is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to provide continuing medical education activities for physicians. The ARRS designates this journal-based CME activity for a maximum of 1.00 AMA PRA Category 1 Credits™ and 1.00 American Board of Radiology©, MOC Part II, Self-Assessment CME (SA-CME). Physicians should claim only the credit commensurate with the extent of their participation in the activity. To access the article for credit, follow the prompts associated with the online version of this article.

An electronic supplement is available online at doi.org/10.2214/AJR.21.26486.

RESULTS.—Median clinical follow-up interval after CT was 8.8 years (interquartile range, 5.1–11.6 years), yielding hip fractures and death in 219 (2.4%) and 549 (6.0%) patients, respectively. L1-level and L3-level muscle attenuation measurements were not different in 2-, 5-, or 10-year AUC for hip fracture (p = .18–.98) or death (p = .19–.95). For hip fracture, 5-year AUCs for L1-level muscle attenuation, L3-level muscle attenuation, and FRAX score were 0.717, 0.709, and 0.708, respectively. For death, 5-year AUCs for L1-level muscle attenuation, L3-level muscle attenuation, and FRS were 0.737, 0.721, and 0.688, respectively. Lowest quartile hazard ratios (HRs) for hip fracture were 2.20 (L1 attenuation), 2.45 (L3 attenuation), and 2.53 (FRAX score), and for death were 3.25 (L1 attenuation), 3.58 (L3 attenuation), and 2.82 (FRS). CT-based muscle cross-sectional area measurements at L1 and L3 were less predictive for hip fracture and death (5-year AUC ≤ 0.571; HR ≤ 1.56).

CONCLUSION.—Automated CT-based measurements of muscle attenuation for myosteatosis at the L1 level compare favorably with previously established L3-level measurements and clinical risk scores for predicting hip fracture and death. Assessment for myopenia was less predictive of outcomes at both levels.

CLINICAL IMPACT.—Alternative use of the L1 rather than L3 level for CT-based muscle measurements allows sarcopenia assessment using both chest and abdominal CT scans, greatly increasing the potential yield of opportunistic CT screening.

Keywords

CT; myosteatosis; opportunistic screening; outcomes; sarcopenia

Sarcopenia can refer to loss of skeletal muscle mass, function, or quality and is associated with many adverse clinical outcomes, including hip fractures and death [1, 2]. Objective muscle data from body CT scans can be obtained as an opportunistic assessment for sarcopenia, regardless of the clinical indication for imaging [2, 3]. CT can assess for loss of muscle mass and quality, indicating myopenia and myosteatosis, respectively [1, 4, 5]. A wide variety of CT-based approaches for muscle assessment have been reported, ranging from manual to fully automated deep learning techniques and using multiple different anatomic levels [4]. A recent systematic review of 388 published studies of CT muscle measurements found that the most frequently used technique was assessment of skeletal muscle at the L3 vertebral level [4]. More recent studies using fully automated deep learning techniques have also favored the L3 level for CT-based muscle assessment [6-8]. However, if CT-based sarcopenia assessment at the L1 level was shown to be comparable to assessment at the L3 level, then both chest and abdominal CT scan data could be used, greatly increasing the overall yield of opportunistic sarcopenia screening.

The purpose of our study was to compare the utility of fully automated deep learning CT-based muscle quantitation at the L1 versus L3 vertebral levels for predicting future hip fracture and death. The performance of the CT-based muscle measurements is also compared with two validated clinical risk scores, the fracture risk assessment tool (FRAX) score and the Framingham risk score (FRS).

Methods

Study Cohort and Adverse Clinical Outcomes

This HIPAA-compliant investigation was approved by the institutional review board at the University of Wisconsin School of Medicine and Public Health and the Office of Human Subjects Research Protection at the National Institutes of Health Clinical Center. The requirement for written informed consent was waived for this retrospective assessment. The study's patient group represents a cohort of patients with documented clinical outcomes that has been previously used for external validation of a variety of automated CT-based body composition tools and clinical outcomes [6-12]. The initially identified sample included 9399 consecutive adult outpatients who were asymptomatic and who underwent low-dose unenhanced abdominopelvic CT for routine colorectal cancer screening between 2004 and 2016 at a single medical center. Excluded were 63 patients with symptoms that prompted the CT colonography evaluation, 82 who had less than 1 year of clinical follow-up after the CT evaluation in the absence of an earlier defined adverse event, three with unavailable or corrupted CT DICOM data, and 28 with muscle tool failure. These exclusions resulted in a final included study cohort of 9223 asymptomatic adults (mean age, 57 ± 8 [SD] years; 5152 women, 4071 men).

A dedicated search script of the electronic health record (EHR) was used to identify two potential clinical outcomes for all patients after the CT scanning: osteoporotic hip fractures and all-cause mortality. Manual per-patient data abstraction was not required.

CT Scanning Protocol

The low-dose unenhanced supine MDCT scans used for this investigation were performed using the same general protocol on scanners from a single vendor (GE Healthcare), with 120 kVp and modulated tube current to achieve a noise index of 50, typically resulting in an effective dose of 2–3 mSv. The low-dose prone series was not used. The original supine series were all reconstructed to axial 3-mm slice thickness using a standard soft-tissue kernel for the automated body composition tools. The specific additional colonography-related techniques for bowel preparation and colonic distention have been previously described [13, 14].

Deep Learning Quantitative Visualization Tool for Muscle Measurements

The muscle segmentation algorithm used in this study represents a fully automated deep learning system developed, trained, and tested using CT images from separate cohorts [15, 16]. The initial processing step was a fully automated spine segmentation leveling tool that determines the endplates of the vertebral levels and labels the spine from T12 to L5 [15, 16]. A U-Net neural network model was then used for rapid automated detection and segmentation of muscles (psoas, paraspinal, and abdominal wall musculature) at each vertebral level [17]. Prior training of the tool involved manual muscle segmentation using images distinct from the CT dataset used in this study. Cross-entropy was used for the loss function, and optimization was performed using the adaptive moment estimation (Adam) method [18]. Manual muscle segmentations at two axial levels for each lumbar vertebra

were used to train the model for the deep learning system. During testing, the trained model was deployed for muscle segmentations at multiple lumbar vertebral levels.

The Dice coefficient relative to manual skeletal muscle segmentation for this tool was previously shown to be highest at the L3 level (0.930–0.939), lowest at the L5 level (0.821–0.850), and intermediate for the L1 level (0.879–0.890) [16]. The neural network model was subsequently upgraded to a 3D U-Net variant that provided a smoother segmentation in the longitudinal direction [19]. The upgrade was performed by running the 2D U-Net muscle segmenter on a dataset of 176 contrast-enhanced body CT scans that included the abdomen and their synthetic unenhanced CT counterparts [20]. The output of the 2D muscle segmenter on these cases was then used to train the 3D U-Net. The tool allows visual confirmation of the automated segmentation via single-slice display. Figure 1 shows an example of the tool's output.

This automated deep learning tool provides an analysis of both muscle density and bulk at the desired segmented level. Muscle density is measured by the mean of the CT attenuation values (in HU) of the segmented voxels. Muscle bulk is measured by the cross-sectional area of the segmented voxels (i.e., the ratio between the volume and slice thickness). The segmentation generally included intramuscular adipose tissue, which is anticipated to favor the performance of muscle attenuation over muscle cross-sectional area.

The same deep learning tool was applied to all CT scans in the patient cohort. The muscle tissue measurement algorithm used a Core i7 processor (Intel) with 4× Titan X GPUs (Nvidia) and required less than 1 minute to process the entire T12–L5 region. This research tool is currently not commercially available.

Clinical Risk Scores

The utility of the CT-based muscle measurements was compared with two multivariable clinical risk scores: FRAX and FRS [21, 22]. The FRAX score provides a 10-year risk estimate for hip fracture and a separate risk estimate for any fragility fracture. The FRS provides a 10-year risk estimate for major cardiovascular events. Consistent with prior research [7], we used the FRS as a proxy for risk of death. We constructed an EHR search script to extract all relevant available clinical data for each patient, using values closest in time to the CT scan. For FRAX score, this included patient age, sex, height, weight, smoking status, previous fracture, and femoral neck T-score from dual-energy x-ray absorptiometry, along with additional clinical variables [22]. For FRS, this included patient age, sex, blood pressure, cholesterol, lipids, diabetic status, and smoking status. The actual 10-year risk estimates for FRAX score and FRS according to these values were then derived by a research assistant with 6 years of experience in medical research (P.M.G.).

Statistical Analysis

Summary statistics were compiled and compared between patients with and without subsequent hip fracture and between patients who survived versus those who died during the clinical follow-up interval using the Wilcoxon-Mann-Whitney test. To investigate the association between predictive measures and downstream adverse events, we used both logistic regression analysis to compute ROC curves and event-free survival analysis. For

ROC curve analysis, datasets were restricted to defined follow-up time intervals (2, 5, and 10 years) to ensure fair comparisons between event and nonevent groups. AUCs with 95% CIs were calculated. For the time-to-event survival analysis, Kaplan-Meier curves were generated by splitting predictor variables into quartiles. Cox proportional hazards models were also used to derive concordance values and individual risk predictions. In addition, hazard ratios (HRs) with 95% CI were computed for each predictor, comparing the highest-risk quartile against the other three quartiles. AUCs and Cox model concordance were compared between L1 and L3 measurements using the DeLong method. Separate thresholds were derived for the attenuation measures to achieve 80% sensitivity or 80% specificity for hip fracture and for death at 5 years. The time-to-event plots were visually assessed for differences between measures. R (version 3.6, R Project for Statistical Computing) was used for statistical analyses. Results for L1-level CT muscle markers have not been previously reported; the results for L3-level CT muscle markers, FRAX, and FRS have been previously reported [6, 7].

Results

Over a median clinical follow-up interval after CT of 8.8 years (interquartile range, 5.1–11.6 years), a total of 219 (2.4%) patients had a hip fracture and 549 (6.0%) died. Table 1 summarizes the CT markers and clinical risk scores stratified by the two patient outcomes. Patients with subsequent hip fracture compared with those without hip fracture exhibited significantly lower CT-based muscle attenuation at the L1 level (17.6 ± 15.3 vs 22.9 ± 15.3 HU; $p < .001$) and L3 level (24.2 ± 12.7 vs 29.0 ± 12.1 HU; $p < .001$), significantly lower CT-based muscle area at the L3 level (145.8 ± 41.2 vs 154.6 ± 46.0 cm^2; $p = .006$), and a significantly higher FRAX score (1.8% ± 3.3% vs 0.8% ± 1.8%; $p < .001$), though there was no significant difference in CT-based muscle area at the L1 level (107.5 ± 36.4 vs 110.6 ± 37.9 cm^2; $p = .22$). Patients who died during the study interval compared with those who survived exhibited significantly lower CT-based muscle attenuation at the L1 level (12.9 ± 16.7 vs 23.4 ± 15.0 HU; $p < .001$) and L3 level (20.8 ± 15.3 vs 29.4 ± 11.7 HU; $p < .001$), significantly lower CT-based muscle area at the L1 level (110.2 ± 41.5 vs 114.4 ± 37.6 cm^2; $p = .005$), and significantly higher FRS (6.7% ± 5.9% vs 3.4% ± 4.0%; $p < .001$), though there was no significant difference in CT-based muscle area at the L3 level (150.9 ± 46.5 vs 154.6 ± 45.9 cm^2; $p = .13$).

Table 2 presents the 2-, 5-, and 10-year AUCs and Cox model concordance coefficients for the two study endpoints for the various study measures. For hip fracture, the 5-year AUC was 0.717 (95% CI, 0.657–0.777) for L1-level muscle attenuation, 0.709 (95% CI, 0.639–0.778) for L3-level muscle attenuation, 0.521 (95% CI, 0.449–0.594) for L1-level muscle area, 0.571 (95% CI, 0.495–0.647) for L3-level muscle area, and 0.708 (95% CI, 0.629–0.787) for the FRAX score. For overall survival, the 5-year AUC was 0.737 (95% CI, 0.703–0.770) for L1-level muscle attenuation, 0.721 (95% CI, 0.683–0.759) for L3-level muscle attenuation, 0.536 (95% CI, 0.496–0.576) for L1-level muscle area, 0.554 (95% CI, 0.513–0.596) for L3-level muscle area, and 0.688 (95% CI, 0.650–0.727) for the FRS. There was no significant difference (all $p > .05$) between the L1 and L3 measures for muscle attenuation for hip fracture ($p = .18$–$.98$) or for either muscle attenuation or muscle area for death ($p = .19$–$.95$) for 2-, 5-, or 10-year AUC. There was also no significant difference

between the L1 and L3 measures for muscle area for hip fracture (p = .08–.54) for 2- or 5-year AUC. Figure 2 shows ROC curves for the CT-based measurements using the 5-year outcomes.

For predicting hip fracture, the L1-level muscle attenuation at a threshold of 24.7 HU achieved sensitivity of 80% (45/56) and specificity of 55% (3738/6812) at 5 years, whereas a threshold of 15.5 HU achieved 80% (5459/6812) specificity and 52% (29/56) sensitivity. The L3-level muscle attenuation at a threshold of 31.8 HU achieved sensitivity of 80% (45/56) and specificity of 49% (3314/6812), whereas a threshold of 21.4 HU achieved 80% specificity (5439/6812) and 54% (30/56) sensitivity. For predicting death, the L1-level muscle attenuation at a threshold of 26.2 HU achieved sensitivity of 80% (173/216) and specificity of 52% (3454/6675), whereas a threshold of 12.9 HU achieved 80% (5340/6675) specificity and 53% (115/216) sensitivity. The L3-level muscle attenuation at a threshold of 32.2 HU achieved sensitivity of 80% (172/216) and specificity of 48% (3195/6675), whereas a threshold of 21.8 HU achieved 80% (5340/6675) specificity and 53% (115/216) sensitivity.

The HR for future hip fracture comparing the highest-risk quartile with the remaining three quartiles was 2.20 (95% CI, 1.68–2.90) for L1-level muscle attenuation, 2.45 (95% CI, 1.87–3.02) for L3-level muscle attenuation, 1.15 (95% CI, 0.86–1.54) for L1-level muscle area, 1.56 (95% CI, 1.18–2.07) for L3-level muscle area, and 2.53 (95% CI, 1.94–3.32) for the FRAX score. The HR for death comparing the highest-risk quartile with the remaining three quartiles was 3.25 (95% CI, 2.75–3.84) for L1-level muscle attenuation, 3.58 (95% CI, 3.02–4.23) for L3-level muscle attenuation, 1.35 (95% CI, 1.12–1.62) for L1-level muscle area, 1.26 (95% CI, 1.05–1.51) for L3-level muscle area, and 2.82 (95% CI, 2.36–3.37) for the FRS score. Figures 3 and S1 (Fig. S1 can be viewed in the *AJR* electronic supplement to this article at www.ajronline.org) show time-to-fracture and time-to-death plots by quartile for L1 and L3 muscle attenuation and area measurements. These plots visually show better separation among quartiles using muscle attenuation than muscle area measurements, though similar separation is seen between L1 and L3 measurements.

Discussion

In this study, we show similar performance in predicting key clinical outcomes of hip fracture and death for CT-based muscle assessment between the L1 level and the standard L3 level. Further, CT muscle attenuation assessment at both the L1 and L3 levels compared favorably to the FRAX and FRS clinical risk scores. The practical implication of these findings is that L1-level muscle assessment would allow sarcopenia assessment using chest CT scans in addition to abdominal CT scans. Conversely, L1- and L3-level muscle cross-sectional area measurements were not found to be useful markers for predicting hip fracture or death. We also show the utility of a fully automated deep learning tool for CT-based muscle assessment, which provides objective results that can be applied to large patient cohorts.

Sarcopenia can refer to loss of muscle mass, function, or quality and is closely associated with cachexia (involuntary wasting), frailty (physiologic decline), and aging in general [1, 5]. A wide range of clinical and imaging-based criteria for sarcopenia have been

applied, and a consensus definition has been proposed [23]; nonetheless, clear overarching definitions remain lacking [5]. A variety of manual, semiautomated, and fully automated approaches exist for CT-based skeletal muscle assessment [2, 5, 8]. The large volume of body CT scans performed for a wide variety of clinical indications may be leveraged to provide opportunities for muscle evaluation [2, 3]. A systematic review by Amini et al. [4] of 388 published studies that used CT to assess for sarcopenia found that measurements were most commonly performed at the L3 level, followed by the L4–5 level and the thigh. Relatively few studies outside of cohorts of patients with cancer have used the L1 level for muscle assessment or have used muscle attenuation rather than area [24, 25]. More recent works have used automated CT-based muscle assessment by a deep learning tool, though they have also primarily used the L3 level [6-8]. Lenchik et al. [26] found that automated CT-based muscle assessment at T12 correlated with survival in men undergoing lung cancer screening. Boutin et al. [27] showed that manual CT-based muscle assessment was slightly better at the T12 level than the L4 level in predicting mortality after hip fracture.

Muscle quality can be assessed at CT using mean attenuation values obtained by either a manual ROI or by a more comprehensive cross-sectional segmentation process. Some tools separate intermuscular adipose tissue from muscle. Muscle bulk or quantity is often reflected by the cross-sectional area of segmented psoas, paraspinal, and body wall musculature at a given level, which can be normalized according to patient height into a skeletal muscle index. The relative importance of muscle quantity (myopenia) versus quality (myosteatosis) is controversial [1]. Prior studies have generally focused on muscle area, commonly evaluated using the skeletal muscle index. Although the relative performance of these measures may depend in part on the segmentation methodology, CT-based myosteatosis appears to be an overall more powerful predictor of adverse outcomes [28]. A recent publication found that CT-based muscle attenuation measurement was the best predictor of mortality among patients with hip fracture compared with many other clinical and CT variables, whereas skeletal muscle index was not a significant predictor [29]. Our muscle tool generally includes within the segmented volume areas of intermuscular adipose tissue within the paraspinal and body wall musculature; such fat would also generally be included within a manually placed ROI. This inclusion of intermuscular fat lowers attenuation values but generally does not alter cross-sectional area. Consequently, muscle attenuation measured using this CT approach more rapidly decreases with increasing patient age than does muscle area [8]. Furthermore, muscle attenuation has been previously shown to be a useful marker for predicting adverse outcomes such as osteoporotic fractures and death [6, 7]. Other developed CT-based muscle tools specifically exclude these areas of fatty deposition, which may hinder the predictive ability of CT-based muscle attenuation measurements [5].

We acknowledge limitations to this study. We investigated a cohort of generally healthy outpatients without symptoms who were undergoing a screening examination. As such, fewer downstream adverse outcomes might be expected; nonetheless, our prolonged follow-up interval, which was often greater than a decade, uncovered a considerable number of hip fractures and deaths within the cohort. We cannot exclude the possibility that some adverse outcomes may have been missed by the EHR search. The CT technique consisted of only unenhanced imaging, and small corrections may be needed to reflect the impact of IV contrast material on muscle and other body composition measures [30, 31]. Also,

we did not use thoracic CT scan data in our study. Although in our experience the L1 level is generally included in the FOV of chest CT scans, it is possible that some chest CT protocols do not always include the L1 level. Our results should also be confirmed in different practice settings to show generalizability, including variation in CT scanners and technique, and in more diverse patient cohorts. Also, although we only evaluated hip fractures and death, CT-based muscle measures have proven useful for predicting any osteoporotic fracture [6], cardiovascular events [7], and postoperative complications and cancer frailty [32-34]. Finally, we conducted a focused comparison of univariable L1- and L3-level muscle measurements. However, predictive ability may be improved by combining these muscle measurements with other CT-based body composition measures (e.g., of bone and fat) [35].

In summary, an automated deep learning tool for quantitative visualization of body wall musculature at the L1 vertebral level compares favorably with the previously established L3 level and clinical risk scores for predicting hip fractures and death. Assessment of muscle attenuation (myosteatosis) performed better than assessment of muscle area (myopenia), likely in part because of the automated tool's inclusion of intramuscular fat within the segmented areas. Because L1-level muscle measures can generally be derived from both chest and abdominal CT scans, use of the L1 level rather than the L3 level increases the overall potential yield of opportunistic CT screening for sarcopenia.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported in part by the Intramural Research Program of the National Institutes of Health (NIH) Clinical Center and the high performance computing capabilities of the NIH Biowulf cluster.

References

1. Boutin RD, Yao L, Canter RJ, Lenchik L. Sarcopenia: current concepts and imaging implications. AJR 2015; 205:[web]W255–W266 [PubMed: 26102307]
2. Lenchik L, Boutin RD. Sarcopenia: beyond muscle atrophy and into the new frontiers of opportunistic imaging, precision medicine, and machine learning. Semin Musculoskelet Radiol 2018; 22:307–322 [PubMed: 29791959]
3. Pickhardt PJ, Graffy PM, Perez AA, Lubner MG, Elton DC, Summers RM. Opportunistic screening at abdominal CT: use of automated body composition biomarkers for added cardiometabolic value. RadioGraphics 2021; 41:524–542 [PubMed: 33646902]
4. Amini B, Boyle SP, Boutin RD, Lenchik L. Approaches to assessment of muscle mass and myosteatosis on computed tomography: a systematic review. J Gerontol A Biol Sci Med Sci 2019; 74:1671–1678 [PubMed: 30726878]
5. Boutin RD, Lenchik L. Value-added opportunistic CT: insights into osteoporosis and sarcopenia. AJR 2020; 215:582–594 [PubMed: 32755187]
6. Pickhardt PJ, Graffy PM, Zea R, et al. Automated abdominal CT imaging biomarkers for opportunistic prediction of future major osteoporotic fractures in asymptomatic adults. Radiology 2020; 297:64–72 [PubMed: 32780005]

7. Pickhardt PJ, Graffy PM, Zea R, et al. Automated CT biomarkers for opportunistic prediction of future cardiovascular events and mortality in an asymptomatic screening population: a retrospective cohort study. Lancet Digit Health 2020; 2:e192–e200 [PubMed: 32864598]
8. Graffy PM, Liu J, Pickhardt PJ, Burns JE, Yao J, Summers RM. Deep learning-based muscle segmentation and quantification at abdominal CT: application to a longitudinal adult screening cohort for sarcopenia assessment. Br J Radiol 2019; 92:20190327 [PubMed: 31199670]
9. Graffy PM, Liu J, O'Connor S, Summers RM, Pickhardt PJ. Automated segmentation and quantification of aortic calcification at abdominal CT: application of a deep learning-based algorithm to a longitudinal screening cohort. Abdom Radiol (NY) 2019; 44:2921–2928 [PubMed: 30976827]
10. Graffy PM, Sandfort V, Summers RM, Pickhardt PJ. Automated liver fat quantification at nonenhanced abdominal CT for population-based steatosis assessment. Radiology 2019; 293:334–342 [PubMed: 31526254]
11. Lee SJ, Liu J, Yao J, Kanarek A, Summers RM, Pickhardt PJ. Fully automated segmentation and quantification of visceral and subcutaneous fat at abdominal CT: application to a longitudinal adult screening cohort. Br J Radiol 2018; 91:20170968 [PubMed: 29557216]
12. Pickhardt PJ, Lee SJ, Liu J, et al. Population-based opportunistic osteoporosis screening: validation of a fully automated CT tool for assessing longitudinal BMD changes. Br J Radiol 2019; 92:20180726 [PubMed: 30433815]
13. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003; 349:2191–2200 [PubMed: 14657426]
14. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403–1412 [PubMed: 17914041]
15. Jianhua Y, Connor SDO, Summers RM. Automated spinal column extraction and partitioning. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006. 2006:390–393
16. Burns JE, Yao J, Chalhoub D, Chen JJ, Summers RM. A machine learning algorithm to estimate sarcopenia on abdominal CT. Acad Radiol 2020; 27:311–320 [PubMed: 31126808]
17. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi A, eds. Medical image computing and computer-assisted intervention: MICCAI 2015. Springer International Publishing, 2015:234–241
18. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv website. arxiv.org/abs/1412.6980. Published December 22, 2014. Revised January 30, 2017. Accessed October 26, 2020
19. Sandfort V, Yan K, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. Sci Rep 2019; 9:16884 [PubMed: 31729403]
20. Roth HR, Lu L, Seff A, et al. A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. In: Golland P, Hata N, Barillot C, Hornegger J, Howe R, eds. Medical image computing and computer-assisted intervention: MICCAI 2014. Springer International Publishing, 2014:520–527
21. D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008; 117:743–753 [PubMed: 18212285]
22. Kanis JA, Oden A, Johansson H, Borgström F, Ström O, McCloskey E. FRAX and its applications to clinical practice. Bone 2009; 44:734–743 [PubMed: 19195497]
23. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. ; Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), and the Extended Group for EWGSOP2. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019; 48:16–31 [PubMed: 30312372]
24. Kang DO, Park SY, Choi BG, et al. Prognostic impact of low skeletal muscle mass on major adverse cardiovascular events in coronary artery disease: a propensity score-matched analysis of a single center all-comer cohort. J Clin Med 2019; 8:712
25. Han JW, Song H, Kim SH. The association between L1 skeletal muscle index derived from routine CT and in-hospital mortality in CAP patients in the ED. Am J Emerg Med 2021; 42:49–54 [PubMed: 33450707]

26. Lenchik L, Barnard R, Boutin RD, et al. Automated muscle measurement on chest CT predicts all-cause mortality in older adults from the national lung screening trial. J Gerontol A Biol Sci Med Sci 2021; 76:277–285 [PubMed: 32504466]
27. Boutin RD, Bamrungchart S, Bateni CP, et al. CT of patients with hip fracture: muscle size and attenuation help predict mortality. AJR 2017; 208:[web]W208–W215 [PubMed: 28267356]
28. Aleixo GFP, Shachar SS, Nyrop KA, Muss HB, Malpica L, Williams GR. Myosteatosis and prognosis in cancer: systematic review and meta-analysis. Crit Rev Oncol Hematol 2020; 145:102839 [PubMed: 31877534]
29. Bae SJ, Lee SH. Computed tomographic measurements of the psoas muscle as a predictor of mortality in hip fracture patients: muscle attenuation helps predict mortality in hip fracture patients. Injury 2021; 52:1456–1461 [PubMed: 33279172]
30. Boutin RD, Kaptuch JM, Bateni CP, Chalfant JS, Yao L. Influence of IV contrast administration on CT measures of muscle and bone attenuation: implications for sarcopenia and osteoporosis evaluation. AJR 2016; 207:1046–1054 [PubMed: 27556335]
31. Perez AA, Pickhardt PJ, Elton DC, Sandfort V, Summers RM. Fully automated CT imaging biomarkers of bone, muscle, and fat: correcting for the effect of intravenous contrast. Abdom Radiol (NY) 2021; 46:1229–1235 [PubMed: 32948910]
32. Martin L, Birdsell L, Macdonald N, 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 2013; 31:1539–1547 [PubMed: 23530101]
33. Prado CMM, Lieffers JR, McCargar LJ, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. Lancet Oncol 2008; 9:629–635 [PubMed: 18539529]
34. Simonsen C, de Heer P, Bjerre ED, et al. Sarcopenia and postoperative complication risk in gastrointestinal surgical oncology: a meta-analysis. Ann Surg 2018; 268:58–69 [PubMed: 29373365]
35. Ormsbee MJ, Prado CM, Ilich JZ, et al. Osteosarcopenic obesity: the role of bone, muscle, and fat on health. J Cachexia Sarcopenia Muscle 2014; 5:183–19 [PubMed: 24740742]

HIGHLIGHTS

Key Finding

- Among 9223 asymptomatic adults who underwent abdominal CT, muscle attenuation measurements at the L1 and L3 levels obtained using a fully automated deep learning tool showed similar utility in predicting subsequent hip fracture and death; muscle attenuation measurements at both levels showed performance comparable with established clinical risk scores.

Importance

- Use of L1 (typically included on both chest and abdominal CT examinations) rather than L3 level expands the reach of opportunistic CT screening for sarcopenia.

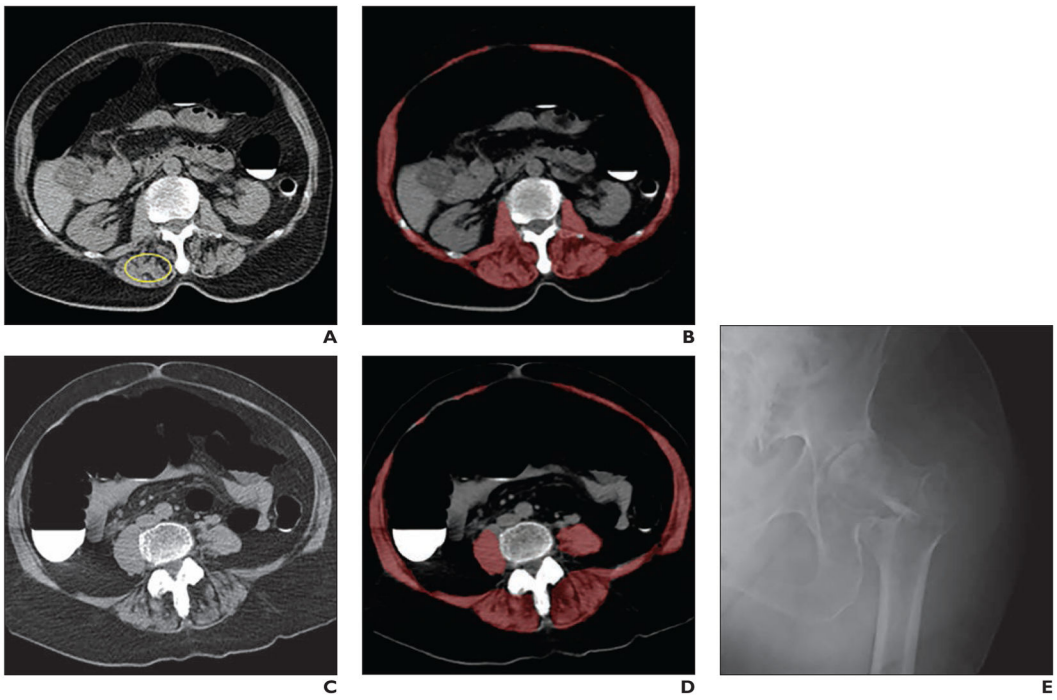

Fig. 1—. Sarcopenia (myosteatosis) at screening CT colonoscopy in 79-year-old woman with subsequent hip fracture.

A and **B,** CT images show L1 vertebral level without (**A**) and with (**B**) overlay of automated skeletal muscle segmentation (**B**, *red*). Intramuscular fat is present within paraspinal muscles (*circle*, **A**). Mean muscle attenuation is similar for manually placed ROI (1.8 HU) and automated tool (3.9 HU) and is markedly decreased for both approaches. Muscle cross-sectional area is relatively preserved when intramuscular fat is included.

C and **D,** CT images at L3 level without (**C**) and with (**D**) muscle segmentation show similar findings.

E, Hip radiograph shows left intertrochanteric femoral fracture that occurred when patient fell 2 years after initial CT. Patient died 4 years after fall.

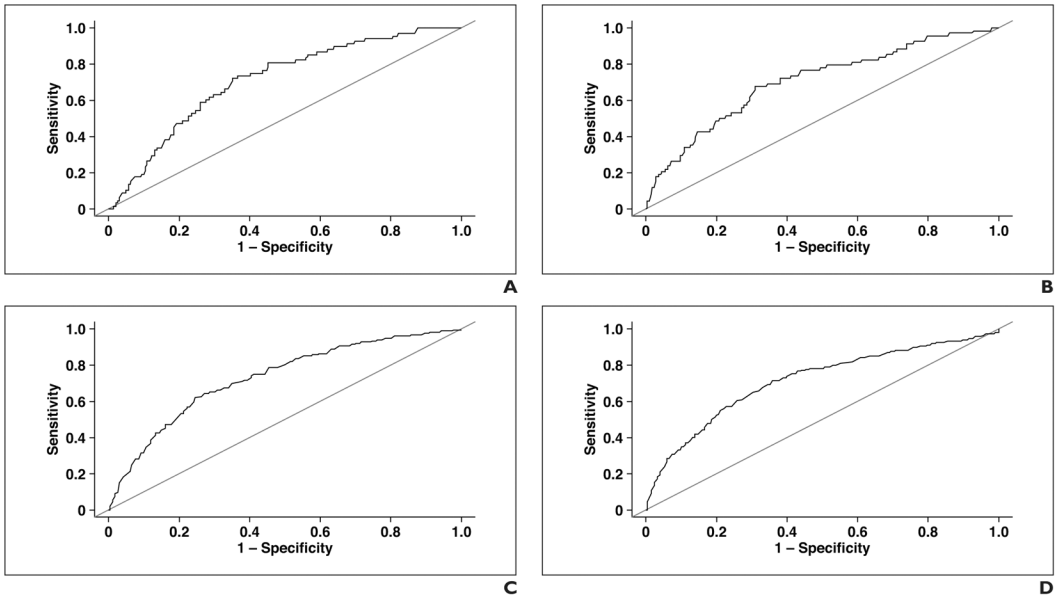

Fig. 2—. Five-year ROC curves for prediction of hip fracture and death.

A and **B,** ROC curves show prediction of hip fracture with automated muscle attenuation measurements at L1 (**A**) and L3 (**B**) levels, with corresponding AUCs of 0.717 and 0.709, respectively.

C and **D,** ROC curves show prediction of death with automated muscle attenuation measurements at L1 (**C**) and L3 (**D**) levels, with corresponding AUCs of 0.737 and 0.721, respectively.

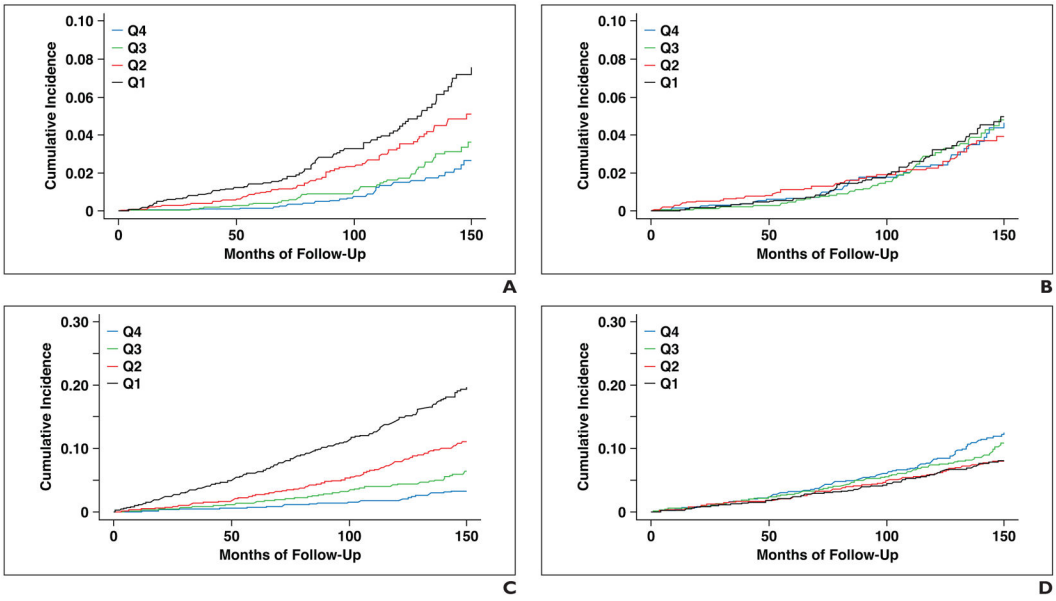

Fig. 3—. Time-to-event plots by quartile (Q) using muscle measurements at L1 level.

A and **B,** Time-to–hip fracture plots for muscle attenuation (**A**) and muscle area (**B**) divide these CT-based measures into patient quartiles, which are then followed over time for adverse outcome. Quartile separation for predicting hip fracture is visually improved for attenuation compared with area.

C and **D,** Time-to-death plots for muscle attenuation (**C**) and muscle area (**D**) also show visually improved quartile separation for attenuation compared with area.

TABLE 1:

Summary Data for CT Muscle Measures and Clinical Risk Scores

CT-Based Measure	Total Cohort (*n* = 9223)	Hip Fracture		Death	
		Yes (*n* = 219)	No (*n* = 9004)	Yes (*n* = 549)	No (*n* = 8674)
Muscle attenuation (HU)					
L1 level					
Mean ± SD	22.8 ± 15.3	17.6 ± 15.3	22.9 ± 15.3	12.9 ± 16.7	23.4 ± 15.0
Median (IQR)	25 (14–34)	19 (7–30)	26 (14–34)	15 (3–25)	26 (15–35)
L3 level					
Mean ± SD	28.9 ± 12.1	24.2 ± 12.7	29.0 ± 12.1	20.8 ± 15.3	29.4 ± 11.7
Median (IQR)	31 (22–38)	25 (17–34)	31 (22–38)	22 (12–31)	31 (23–38)
Muscle area (cm^2)					
L1 level					
Mean ± SD	110.5 ± 37.8	107.5 ± 36.4	110.6 ± 37.9	110.2 ± 41.5	114.4 ± 37.6
Median (IQR)	107 (82–134)	105 (80–130)	107 (82–134)	107 (82–134)	113 (85–139)
L3 level					
Mean ± SD	154.4 ± 45.9	145.8 ± 41.2	154.6 ± 46.0	150.9 ± 46.5	154.6 ± 45.9
Median (IQR)	148 (122–186)	142 (117–171)	148 (122–186)	149 (118–186)	147 (122–186)
FRAX score (%)[a,b]					
Mean ± SD	0.8 ± 1.8	1.8 ± 3.3	0.8 ± 1.8	NA	NA
Median (IQR)	0.3 (0.2–0.7)	0.6 (0.2–1.7)	0.3 (0.2–0.7)	NA	NA
Framingham risk score (%)[a]					
Mean ± SD	3.6 ± 4.2	NA	NA	6.7 ± 5.9	3.4 ± 4.0
Median (IQR)	2 (1–6)	NA	NA	5 (2–10)	2 (1–5)

Note—IQR = interquartile range, FRAX = fracture risk assessment tool, NA = not applicable.

[a]Some of the values were previously reported [6, 7].

[b]FRAX score represents 10-year risk estimate for hip fracture.

TABLE 2:

AUCs and Cox Model Concordances of Study Measures for Predicting Hip Fracture and Death

Outcome, CT-Based Measure	AUC			Cox Model Concordance
	2-Year	5-Year	10-Year	
Hip fracture				
Muscle attenuation				
L1 level	0.713 (0.607–0.802)	0.717 (0.657–0.777)	0.648 (0.605–0.691)	0.647
L3 level[a]	0.746 (0.630–0.863)	0.709 (0.639–0.778)	0.649 (0.604–0.694)	0.651
p (L1 vs L3)	.18	.68	.98	
Muscle area				
L1 level	0.527 (0.415–0.639)	0.521 (0.449–0.594)	0.512 (0.466–0.558)	0.525
L3 level	0.460 (0.330–0.591)	0.571 (0.495–0.647)	0.575 (0.530–0.612)	0.562
p (L1 vs L3)	.54	.08	< .01	
FRAX score[a,b]	0.732 (0.602–0.863)	0.708 (0.629–0.787)	0.621 (0.572–0.671)	0.630
Death				
Muscle attenuation				
L1 level	0.731 (0.677–0.786)	0.737 (0.703–0.770)	0.725 (0.701–0.749)	0.706
L3 level[a]	0.736 (0.676–0.796)	0.721 (0.683–0.759)	0.717 (0.691–0.743)	0.700
p (L1 vs L3)	.95	.19	.23	
Muscle area				
L1 level	0.504 (0.444–0.565)	0.536 (0.496–0.576)	0.551 (0.522–0.581)	0.539
L3 level	0.569 (0.506–0.631)	0.554 (0.513–0.596)	0.543 (0.513–0.573)	0.535
p (L1 vs L3)	.22	.58	.73	
Framingham risk score[a]	0.700 (0.638–0.762)	0.688 (0.650–0.727)	0.693 (0.665–0.720)	0.681

Note—Numbers in parentheses are 95% CI. FRAX = fracture risk assessment tool.

[a]Some of the values were reported previously [6, 7].

[b]FRAX score represents 10-year risk estimate for hip fracture.