ORIGINAL ARTICLE

Journal of Cachexia, Sarcopenia and Muscle 2022; **13**: 2974–2984
Published online 2 September 2022 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/jcsm.13080

Body composition from single versus multi-slice abdominal computed tomography: Concordance and associations with colorectal cancer survival

Ijeamaka Anyene[1], Bette Caan[1], Grant R. Williams[2,3], Karteek Popuri[4], Leon Lenchik[5], Smith Giri[2,3], Vincent Chow[4], Mirza Faisal Beg[4] & Elizabeth M. Cespedes Feliciano[1]*

[1]*Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA;* [2]*Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, AL, USA;* [3]*Division of Hematology/Oncology, School of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA;* [4]*School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada;* [5]*Wake Forest School of Medicine, Winston-Salem, NC, USA*

Abstract

Background Computed tomography (CT) scans are routinely obtained in oncology and provide measures of muscle and adipose tissue predictive of morbidity and mortality. Automated segmentation of CT has advanced past single slices to multi-slice measurements, but the concordance of these approaches and their associations with mortality after cancer diagnosis have not been compared.

Methods A total of 2871 patients with colorectal cancer diagnosed during 2012–2017 at Kaiser Permanente Northern California underwent abdominal CT scans as part of routine clinical care from which mid-L3 cross-sectional areas and multi-slice T12–L5 volumes of skeletal muscle (SKM), subcutaneous adipose (SAT), visceral adipose (VAT) and intermuscular adipose (IMAT) tissues were assessed using Data Analysis Facilitation Suite, an automated multi-slice segmentation platform. To facilitate comparison between single-slice and multi-slice measurements, sex-specific z-scores were calculated. Pearson correlation coefficients and Bland–Altman analysis were used to quantify agreement. Cox models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for death adjusting for age, sex, race/ethnicity, height, and tumour site and stage.

Results Single-slice area and multi-slice abdominal volumes were highly correlated for all tissues (SKM $R = 0.92$, $P < 0.001$; SAT $R = 0.97$, $P < 0.001$; VAT $R = 0.98$, $P < 0.001$; IMAT $R = 0.89$, $P < 0.001$). Bland–Altman plots had a bias of 0 (SE: 0.00), indicating high average agreement between measures. The limits of agreement were narrowest for VAT (± 0.42 SD) and SAT (± 0.44 SD), and widest for SKM (± 0.78 SD) and IMAT (± 0.92 SD). The HRs had overlapping CIs, and similar magnitudes and direction of effects; for example, a 1-SD increase in SKM area was associated with an 18% decreased risk of death (HR = 0.82; 95% CI: 0.72–0.92), versus 15% for volume from T12 to L5 (HR = 0.85; 95% CI: 0.75–0.96).

Conclusions Single-slice L3 areas and multi-slice T12–L5 abdominal volumes of SKM, VAT, SAT and IMAT are highly correlated. Associations between area and volume measures with all-cause mortality were similar, suggesting that they are equivalent tools for population studies if body composition is assessed at a single timepoint. Future research should examine longitudinal changes in multi-slice tissues to improve individual risk prediction.

Keywords adipose tissue; automated segmentation; body composition; colorectal cancer; computed tomography; muscle

Received: 21 April 2022; Revised: 6 July 2022; Accepted: 14 August 2022
**Correspondence to: Elizabeth M. Cespedes Feliciano, ScD, Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA.*
Email: elizabeth.m.cespedes@kp.org

Introduction

Computed tomography (CT) scans are routinely obtained in oncology patients for diagnosis, surgical planning and surveillance. Measures of muscle and adipose tissue derived from CT images are predictive of surgical complications,[1–3] treatment toxicity,[4] and morbidity and mortality[5] after a cancer diagnosis,[1,6,7] demonstrating the importance of body composition to cancer outcomes. As automated tools to segment muscle and adipose tissue on CT images become more efficient, accurate and available,[8–13] these data have the potential to provide more than just research insights: Future applications might include improving patient care by providing personalized cancer treatments and tailoring lifestyle interventions for cancer survivors' individual body composition.

Although most oncology research has used manually segmented muscle and adipose tissue area from a single-slice axial CT image (typically the mid-slice at the third lumbar vertebra [L3][14]), use of artificial intelligence approaches for automated vertebral landmarking and segmentation has allowed body composition research to advance past single-slice tissue areas[15] to multi-organ, multi-slice, volumetric quantifications across larger fields of view.[10] Such advances in automation allow rapid quantification of body composition in large cohort studies. Additionally, by segmenting tissue at vertebral landmarks beyond the L3, automated multi-slice segmentation may better characterize tissues that are variable across vertebral levels (e.g., intermuscular or visceral adipose tissue) and enable muscle and adipose distribution across larger regions of the body to be studied. Finally, it is hypothesized that multi-slice tissue volumes may be more sensitive or accurate than single-slice areas for longitudinal evaluation of within-person changes in body composition. However, multi-slice segmentation also introduces new challenges including variable fields of view across patients and the need for standardized approaches to model these new volumetric exposures.

An essential first step towards integration of multi-slice measurement into body composition research is to understand the relationship between single-slice and multi-slice measurement muscle and adipose tissue measurements and their associations with clinical outcomes. Towards that end, this analysis of 2871 patients with non-metastatic colorectal cancer had two objectives: first, to examine the concordance of abdominal body composition assessed using single-slice areas at the mid-L3 to that assessed using multi-slice volumes from the twelfth thoracic to fifth lumbar (T12–L5) vertebrae, the most common field of view for gastrointestinal cancer patients, and, second, to compare the magnitude and direction of associations with mortality after colorectal cancer diagnosis between these two measurement approaches.

Methods

Study population and setting

We identified patients diagnosed with stage I–III first primary, invasive colorectal cancer at Kaiser Permanente Northern California (KPNC) between 2012 and 2017. KPNC maintains a Virtual Data Warehouse (VDW) in which electronic medical record[16] (EMR) data are curated for research purposes. Individuals were eligible for the cohort if they were at least 18 years of age and $<$80 at time of diagnosis, underwent surgical resection, received an abdominal or pelvic CT scan performed within 4 months of colorectal cancer diagnosis, but before receipt of any chemotherapy or radiation, and had a body mass index (BMI) available within 6 months of CT date, but prior to adjuvant chemotherapy (N = 3309). Further, the individual's CT scan had to include T12 to L5 vertebral bodies (N = 2996). In addition, individuals were excluded due to inadequate CT image abnormalities concerns such as severe truncation artefacts (N = 125) as discussed in greater detail later in the methods as well as in the results, bringing our final analytic cohort to 2871 patients. A large majority of scans, 89.1% (N = 2559), were taken prior to surgical resection.

Automated body composition measurements

To obtain body composition measurements, the Data Analysis Facilitation Suite (DAFS) by Voronoi Health Analytics, Inc. (https://www.voronoihealthanalytics.com/) was utilized. In brief, DAFS provides an automated and accurate end-to-end pipeline for rapid multi-organ segmentation and extraction of body composition measurements from CT images. DAFS takes a PACS (picture archiving and communication system)-exported folder from a scanning instance and first curates the axial DICOM (Digital Imaging and Communications in Medicine standard) images into individual 3D scans.[13] Next, the curated scans are run through non-linear image processing algorithms that provide two key derivations from the CT scan: (1) multi-slice segmentation covering the entirety of the field of view into the multiple organs and tissues present and (2) annotation of each axial slice with the label of the vertebral bone contained in the slice. DAFS also generates automated visualizations, enabling manual assessment of segmentation accuracy and CT imaging abnormalities. In prior validation against manual analysis, the average dice similarity coefficients (spatial overlap index used for validation in image segmentation that quantifies overlap at the pixel or voxel level) were 0.97 for SKM, 0.99 for SAT, 0.96 for VAT[13] and over 0.90 for most of the organs for the DAFS software, and errors in annotation of slices based on vertebral levels were close to 0.

All images and segmentations were reviewed by a single trained research assistant to exclude or flag imaging abnormalities based on predefined criteria. Scans in which the patient had a metal implant, muscle cut-off, bright truncation artefact (beam hardening when the patient's body is outside the field of view) or major segmentation error were removed from analysis (N = 125; see *Figure S1* for sample images). In addition, scans were flagged, but not excluded, if they contained features that could have impacted the tissue quantifications or the radiation attenuation values on the scan. These flags included pannus, an extremity (arm and/or hand) in view, minor to moderate subcutaneous adipose tissue cut-off and minor truncation artefacts (N = 159; see *Figure S1* for sample images).

Body composition measurements

There were four body composition tissues of interest analysed in this study: skeletal muscle (SKM), subcutaneous adipose tissue (SAT), visceral adipose tissue (VAT) and intermuscular adipose tissue (IMAT). For all four tissues, single-slice measurements were calculated at the mid-L3, and multi-slice measurements were calculated across multi-slice scans covering the T12–L5 vertebra. The directly computed measurements were therefore abdominal tissues quantified as cross-sectional area in cm^2 and volume in cm^3 (*Figure 1*). Secondarily, for the SKM region, the mean radiation attenuation value or radiodensity in Hounsfield units (HU) was recorded. We included this as it is commonly used as a measure of myosteatosis or ectopic fat infiltration into skeletal muscle. The radiodensities of adipose tissues were not examined as they are less commonly used in the oncology literature. Of note, the DAFS software default does not impose predefined ranges on the HU values for specific tissues; although this can be customized, we used the default settings as the tissue-specific HU ranges applied in manual analysis protocols vary. The observed means, SD and ranges for each tissue are reported in *Table S3*.

Covariate selection

Demographic characteristics including sex (female, male) and self-reported race/ethnicity (Non-Hispanic White, Hispanic, Black or African American, Asian or Pacific Islander, Other/Missing) were collected from the KPNC EMR via the VDW Demographics file. Body measurements including BMI at scan (kg/m^2) and height at scan (metres [m]) were measured at clinical visits and then extracted from the KPNC VDW vital signs file. Colorectal cancer diagnosis characteristics including cancer stage at diagnosis, primary site of tumour and receipt of adjuvant chemotherapy were obtained from the KPNC Cancer Registry.[16]

Outcomes

All-cause mortality was ascertained from KPNC's VDW mortality files[16] through the last available update (31 December 2019), which incorporates internal data from the KPNC health system, and external linkages with mortality information from the State of California, the Social Security Administration and the National Death Index.

Statistical analysis

Sex-specific z-scores with mean of zero and standard deviation of one for mid-L3 area in cm^2 and T12–L5 volume in cm^3 of SKM, SAT, VAT and IMAT were calculated to allow for unit-free comparison of single-slice and multi-slice measurements. Single-slice area and multi-slice volume z-scores for SKM, SAT, VAT and IMAT were then categorized into three-level ordinal variable for each tissue type depending on whether the z-scores were within 1 SD of the sex-specific mean, >1 SD higher or >1 SD lower. Sex-specific z-scores and categories were also calculated for SKM radiodensity.

The linear correlation for SKM, SAT, VAT and IMAT single-slice area and multi-slice volume z-scores, as well as SKM radiodensity HU (single and multi-slice) was estimated using Pearson correlation coefficients. To quantify the agreement between the single-slice and multi-slice measurements, we used the Bland–Altman method. The bias (mean of the differences) and standard error, as well as the upper and lower limits of agreement (LOA: bias ± 1.96 * SD) were calculated.[17] Individual measurements were classified as outliers if they fell outside the LOA for the respective tissue type. Cohen's kappa coefficients were estimated to assess the agreement between the single-slice area and multi-slice

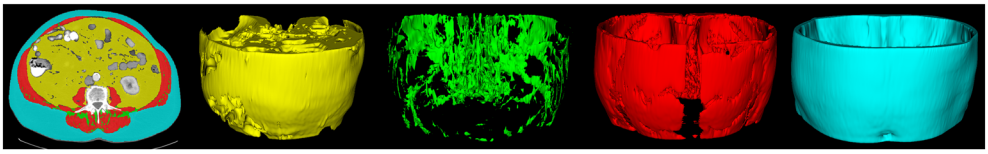

Figure 1 Body composition segmented as 2D tissue areas from the mid-L3 slice (first panel), versus multi-slice, T12–L5 volumes. Visceral adipose tissue is shown in yellow; intermuscular adipose tissue in green; skeletal muscle in red; and subcutaneous adipose tissue in blue.

volume categories described above (three-level, SD-based variable), as well as equivalent categories based on SKM radiodensity. Aligning with Cohen's definition,[18] Cohen's kappa coefficients were interpreted as follows: 0.01–0.20 as no correlation, 0.21–0.39 as minimal, 0.40–0.59 as weak, 0.60–0.79 as moderate, 0.80–0.90 as strong and above 0.90 as near-perfect correlation.

Cox proportional hazards models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) and to compare the strength of association of single slice versus multi-slice measurements with all-cause mortality. 'Single-tissue models', with separate models for each exposure, were fit for single-slice areas (mid-L3 SAT, VAT, IMAT and SKM areas in cm^2, plus SKM radiodensity in HU averaged) and multi-slice volumes (T12–L5 SAT, VAT, IMAT and SKM volumes in cm^3, plus SKM radiodensity in HU averaged). Then, using the same exposure parameterizations, two 'multi-tissue' models were fit with all relevant tissues as mutually adjusted exposures within one model: one for single-slice areas and another for multi-slice volumes. For both single-tissue and multi-tissue models, the exposures were operationalized as either the continuous z-score or three-level, SD-based categories. For SKM radiodensity, we entered the exposure in its native units for all analyses other than survival analyses given that mean single-slice and mean multi-slice radiodensity are both measured in HU. When using the three-level, SD-based category in the cox model, the highest SKM category and the lowest SAT, VAT and IMAT categories were used as the reference. All models were adjusted for the covariates of sex, race/ethnicity, height, cancer stage and primary site of tumour. Person-time was calculated as the years from the date of CT scan to the date of death. If no event occurred, participants were censored at the end of the study period, 31 December 2019.

We conducted two sensitivity analyses. First, we repeated analyses excluding scans that were flagged as having imaging abnormalities for any reason. Second, we repeated analyses after scaling multi-slice tissue volumes in cm^3 to torso length in mm; to do this, we first divided the multi-slice volume of each tissue by the height of the slab in mm from T12 to L5 and then rederived the sex-specific z-scores for these exposures.

All statistical analyses were conducted using R Version 4.0.2. Specifically, data preparation was enabled by the tidyverse, cox proportional hazard regression models were enabled by the survival package, figures were generated using ggplot2 from tidyverse and patchwork packages and tables were generated using gtsummary package.[19–23]

Table 1 Baseline characteristics of patients diagnosed with colorectal cancer at Kaiser Permanente Northern California from 2012 to 2017

Characteristic	Overall, *N* = 2871	Male, *N* = 1551	Female, *N* = 1320
	Mean (SD) or *N* (%)		
Age at diagnosis	61 (11)	60 (11)	61 (12)
Race/ethnicity			
Hispanic	165 (5.7%)	93 (6.0%)	72 (5.5%)
Black or African American	221 (7.7%)	94 (6.1%)	127 (9.6%)
Asian or Pacific Islander	506 (18%)	289 (19%)	217 (16%)
Non-Hispanic White	1947 (68%)	1059 (68%)	888 (67%)
Other	32 (1.1%)	16 (1.0%)	16 (1.2%)
Body mass index at scan (kg/m^2)	28.5 (5.9)	28.7 (5.2)	28.3 (6.6)
Height at scan (m)	1.70 (0.10)	1.76 (0.08)	1.62 (0.07)
AJCC cancer stage, 8th edition			
1	843 (29%)	440 (28%)	403 (31%)
2	876 (31%)	505 (33%)	371 (28%)
3	1152 (40%)	606 (39%)	546 (41%)
Primary site of tumour			
Colon	2088 (73%)	1080 (70%)	1008 (76%)
Rectal	783 (27%)	471 (30%)	312 (24%)
Receipt of adjuvant chemotherapy	1349 (47%)	734 (47%)	615 (47%)
Single-slice body composition at mid-L3			
Skeletal muscle (cm^2)	137 (39)	163 (31)	107 (21)
Subcutaneous adipose (cm^2)	223 (122)	194 (105)	257 (131)
Visceral adipose (cm^2)	173 (113)	216 (116)	122 (86)
Intermuscular adipose (cm^2)	12 (8)	12 (8)	13 (8)
Skeletal muscle (HU)	44 (10)	45 (9)	42 (11)
Multi-slice body composition from T12 to L5			
Skeletal muscle (cm^3)	2407 (784)	2931 (636)	1792 (402)
Subcutaneous adipose (cm^3)	4017 (2226)	3612 (2020)	4492 (2360)
Visceral adipose (cm^3)	2835 (1852)	3605 (1896)	1929 (1311)
Intermuscular adipose (cm^3)	291 (162)	296 (170)	285 (151)
Skeletal muscle (HU)	45 (10)	47 (9)	44 (11)

Note: Scans were taken prior to receipt of any chemotherapy, if received. Abbreviation: AJCC, American Joint Committee on Cancer.

Results

Study population

Baseline characteristics of the cohort are listed in *Table 1*. The mean age at diagnosis was 61 (SD: 11) years, and the cohort was equally male (54%) and female (46%), and primarily non-Hispanic White (68%). There was an almost even distribution between cancer stages, with Stage III being the largest category (40%). The cohort was primarily composed of individuals whose primary tumour site was the colon (73%), and less than half of the cohort received adjuvant chemotherapy (47%). A large majority of the CT images were taken with a 5 mm thickness (2090 [72.8%]) with IV contrast administered (2739 [95.4%]).

Concordance between single-slice area and multi-slice volume measurements

Strong, positive and statistically significant correlations were estimated between single-slice area z-score and multi-slice volume z-scores for all tissue types (*Figure 2*: SKM $R = 0.92$, $P < 0.001$; SAT $R = 0.97$, $P < 0.001$; VAT $R = 0.98$, $P < 0.001$; IMAT $R = 0.89$, $P < 0.001$). However, *Figure 2* illustrates that the correlation between single-slice area and multi-slice volume is a not a 1:1 match. Instead, there is a cloud of values that fall along the regression line.

To better understand the agreement between single-slice area and multi-slice volume z-scores, the Bland–Altman plots in *Figure 3* were assessed. For all tissues, the Bland–Altman plots had a bias of 0.00 (SE: 0.00), indicating high agreement

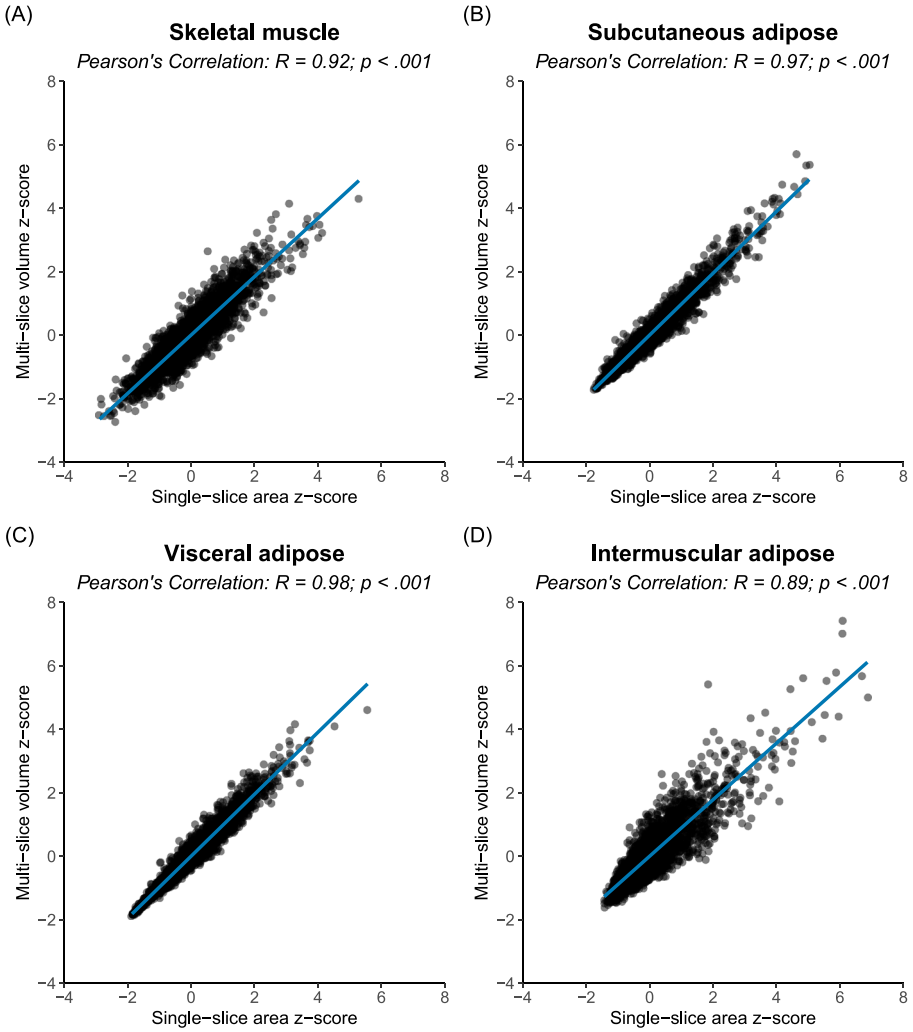

Figure 2 Scatter plots displaying the Pearson correlation coefficient comparing sex-specific z-scores for body composition assessment from single-slice, mid-L3 areas versus multi-slice, T12–L5 volumes at diagnosis of colorectal cancer.

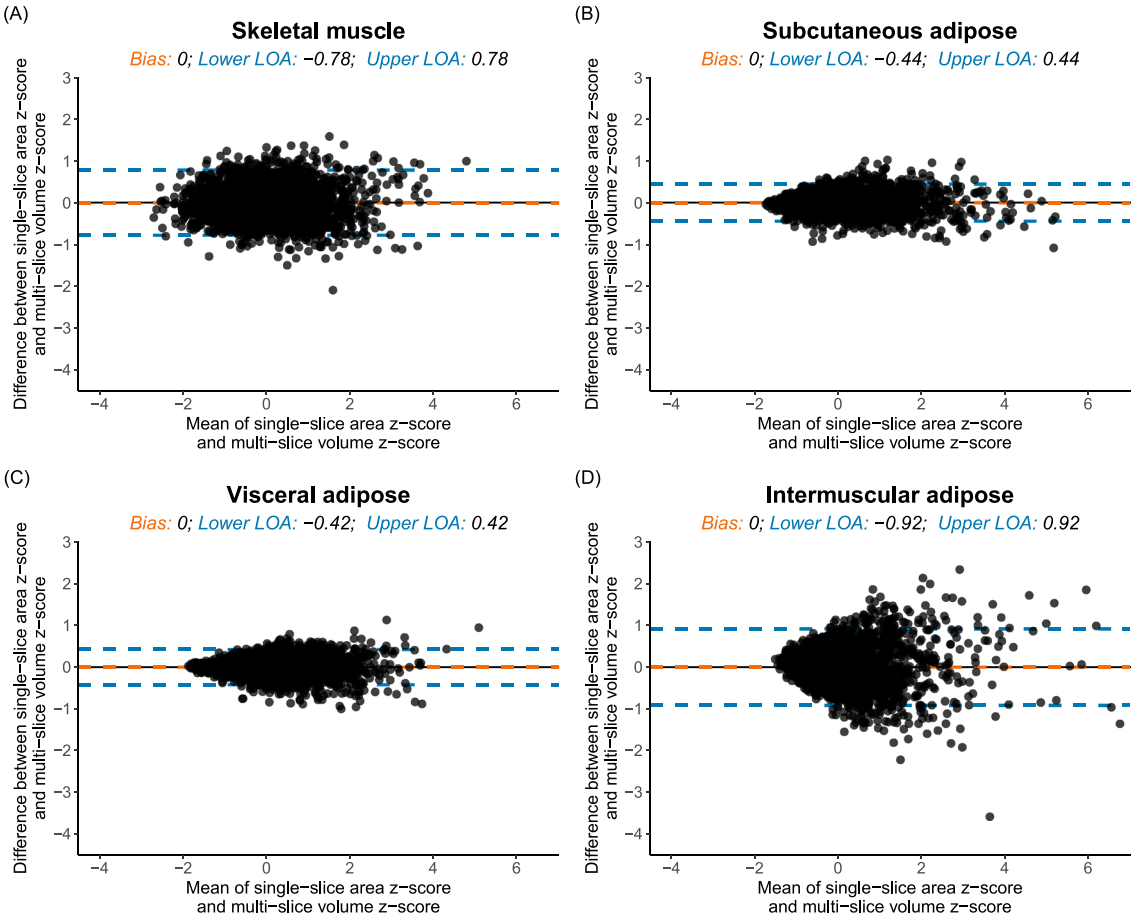

Figure 3 Bland–Altman plots displaying the agreement between sex-specific single-slice area and volume z-scores for each tissue type at diagnosis of colorectal cancer. The dashed blue line is the limits of agreement (±1.96 * SD), and the dashed orange line is the bias (the mean of the differences). (A) Skeletal muscle, (B) subcutaneous adipose, (C) visceral adipose and (D) intermuscular adipose.

between abdominal area and volume on a population level. The LOA were narrowest for VAT (± 0.42 SD) and SAT (± 0.44 SD), and widest for SKM (± 0.78 SD) and IMAT (± 0.92 SD). SKM radiodensity derived from an average across SKM single-slice area versus multi-slice volume also had a near-perfect correlation (R = 0.98, $P < 0.001$), bias of −1.26 with LOA of −5.08 to 2.55 SD (*Figures S1* and *S2*). Although most marked for IMAT (see funnel shape in *Figure 3*), in general, discrepancies between single-slice and multi-slice measurements increased with increasing area or volume of the tissue.

Using the Bland–Altman analysis, differences between single-slice and multi-slice measurements that were more extreme than the LOA were classified as outliers (5.5–6.4% of observations for each tissue). Patients classified as outliers for any tissue had a higher mean BMI (e.g., SKM: 31.5 vs. 28.4 kg/m^2; *Table 2*). This is consistent with the observation that larger patients (those with greater tissue areas or volumes) tended to have greater discrepancies between measurements. However, it is notable that patients classified as outliers for one tissue type were often not outliers for another. In *Figure 4*, the largest overlap in outliers observed is between VAT and SAT; however, this overlap only accounts for a minority of patients (only 1.3% of patients are outliers on both VAT and SAT). We additionally assessed the extent to which outliers were driven by characteristics inherent to the quality of the CT image series. Whereas few scans were flagged during quality review for any reason (n = 159), among those scans outside the LOA for a given tissue, a larger proportion were flagged as having a quality issue on the image (*Table S1*).

Cohen's K was estimated to understand whether single-slice versus multi-slice measurement ranked patients similarly; the agreement between categories based on the sex-specific SDs of single-slice and multi-slice measurements can be seen in *Table S2*. Both SKM and IMAT sex-specific SD categories had moderate agreement between multi-slice volume and single-slice area (SKM kappa coefficient = 0.67, $P < 0.001$; IMAT kappa coefficient = 0.61, $P < 0.001$). In comparison, SAT, VAT and SKM radiodensity sex-specific cat-

Table 2 Differences in characteristics among patients with and without differences between single-slice versus multi-slice sex-specific z-scores that exceed Bland–Altman limits of agreement (n = 2871 colorectal cancer patients at Kaiser Permanente Northern California at diagnosis)

	Skeletal muscle		Subcutaneous adipose		Visceral adipose		Intermuscular adipose		Skeletal muscle radiodensity	
Characteristic	Within LOA, N = 2713[a]	Outside LOA, N = 158[a]	Within LOA, N = 2699[a]	Outside LOA, N = 172[a]	Within LOA, N = 2688[a]	Outside LOA, N = 183[a]	Within LOA, N = 2698[a]	Outside LOA, N = 173[a]	Within LOA, N = 2713[a]	Outside LOA, N = 158[a]
Age at diagnosis	61 (11)	59 (11)	61 (11)	61 (11)	60 (11)	63 (10)**	60 (11)	63 (10)**	60 (11)	64 (11)**
Receipt of adjuvant chemotherapy	1267 (47%)	82 (52%)	1266 (47%)	83 (48%)**	1262 (47%)	87 (48%)	1285 (48%)	64 (37%)**	1284 (47%)	65 (41%)
Sex										
Male	1471 (54%)	80 (51%)	1478 (55%)	73 (42%)	1459 (54%)	92 (50%)	1465 (54%)	86 (50%)	1470 (54%)	81 (51%)
Female	1242 (46%)	78 (49%)**	1221 (45%)	99 (58%)**	1229 (46%)	91 (50%)**	1233 (46%)	87 (50%)**	1243 (46%)	77 (49%)
Race/ethnicity										
Hispanic	154 (5.7%)	11 (7.0%)	151 (5.6%)	14 (8.1%)	154 (5.7%)	11 (6.0%)	157 (5.8%)	8 (4.6%)	153 (5.6%)	12 (7.6%)
Black or African American	194 (7.2%)	27 (17%)	208 (7.7%)	13 (7.6%)	212 (7.9%)	9 (4.9%)	206 (7.6%)	15 (8.7%)	209 (7.7%)	12 (7.6%)
Asian or Pacific Islander	494 (18%)	12 (7.6%)	498 (18%)	8 (4.7%)	484 (18%)	22 (12%)	498 (18%)	8 (4.6%)	481 (18%)	25 (16%)
White	1840 (68%)	107 (68%)	1810 (67%)	137 (80%)	1806 (67%)	141 (77%)	1806 (67%)	141 (82%)	1840 (68%)	107 (68%)
Other	31 (1.1%)	1 (0.6%)	32 (1.2%)	0 (0%)	32 (1.2%)	0 (0%)	31 (1.1%)	1 (0.6%)	30 (1.1%)	2 (1.3%)**
AJCC cancer stage										
1	792 (29%)	51 (32%)	789 (29%)	54 (31%)	792 (29%)	51 (28%)	783 (29%)	60 (35%)	796 (29%)	47 (30%)
2	831 (31%)	45 (28%)	833 (31%)	43 (25%)	823 (31%)	53 (29%)	818 (30%)	58 (34%)	813 (30%)	63 (40%)
3	1090 (40%)	62 (39%)	1077 (40%)	75 (44%)	1073 (40%)	79 (43%)	1097 (41%)	55 (32%)	1104 (41%)	48 (30%)
Primary site of tumour										
Colon	1973 (73%)	115 (73%)	1961 (73%)	127 (74%)	1948 (72%)	140 (77%)	1956 (72%)	132 (76%)	1976 (73%)	112 (71%)
Rectal	740 (27%)	43 (27%)	738 (27%)	45 (26%)	740 (28%)	43 (23%)	742 (28%)	41 (24%)	737 (27%)	46 (29%)
Body mass index at scan (kg/m^2)	28.4 (5.7)	31.5 (7.1)**	28.2 (5.7)	33.9 (6.5)**	28.2 (5.6)	33.4 (7.2)**	28.2 (5.6)	33.5 (8.1)**	28.6 (5.8)	27.8 (7.4)
Height at scan (m)	1.70 (0.10)	1.72 (0.12)**	1.70 (0.10)	1.69 (0.11)	1.70 (0.10)	1.70 (0.12)	1.70 (0.10)	1.73 (0.11)**	1.70 (0.10)	1.69 (0.11)
5 mm thickness										
No	738 (27%)	42 (27%)	731 (27%)	49 (28%)	729 (27%)	51 (28%)	726 (27%)	54 (31%)	730 (27%)	50 (32%)
Yes	1975 (73%)	116 (73%)	1968 (73%)	123 (72%)	1959 (73%)	132 (72%)	1972 (73%)	119 (69%)**	1983 (73%)	108 (68%)
IV contrast administered										
No	80 (3.0%)	7 (4.5%)	79 (3.0%)	8 (4.7%)	78 (2.9%)	9 (5.0%)	75 (2.8%)	12 (7.1%)	81 (3.0%)	6 (3.9%)
Yes	2591 (97%)	148 (95%)	2578 (97%)	161 (95%)	2569 (97%)	170 (95%)	2582 (97%)	157 (93%)	2590 (97%)	149 (96%)
Unknown	42	3	42	3	41	4	41	4	42	3

Abbreviation: AJCC, American Joint Committee on Cancer, 8th edition.

[a]Mean (SD); n (%).

**Statistically significant difference in characteristics (P-value < 0.05). Continuous variables used Welch's two-sample t-test; categorical variables used Pearson's chi-squared test.

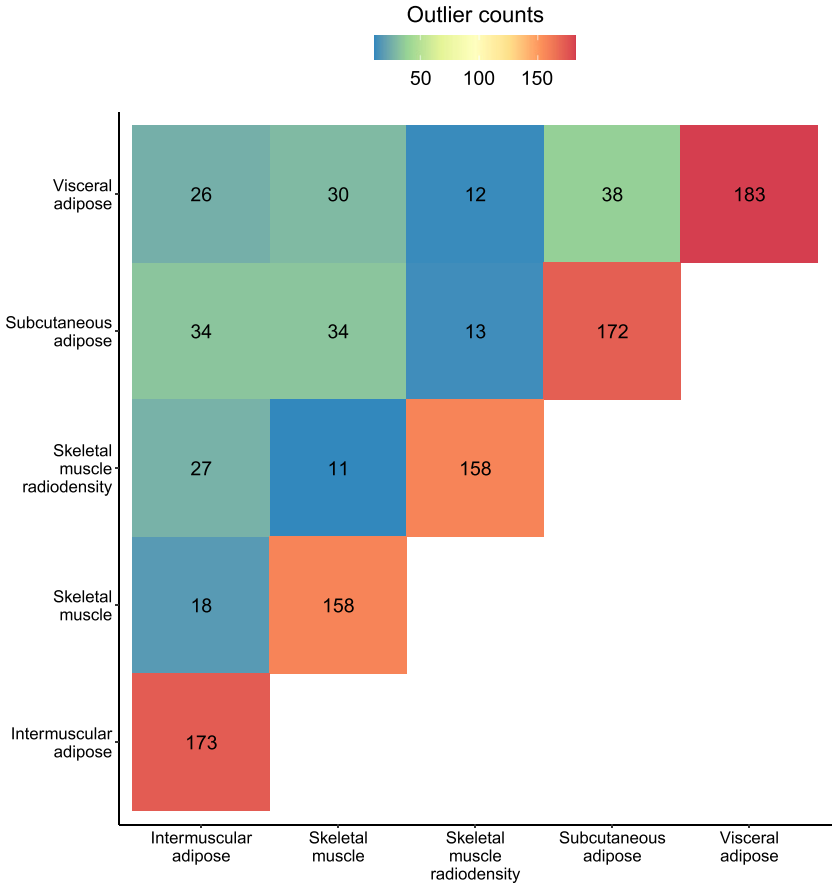

Figure 4 Heatmap of discordance between single-slice versus multi-slice quantification of body composition. The diagonal indicates the number individuals with differences in the sex-specific z-scores between the single-slice and multi-slice measurements that are more extreme than the Bland–Altman limits of agreement for that tissue type, whereas the intersection indicates the number of individuals who are outliers for both tissue types. Warmer colours indicate a greater number of individuals.

egories had strong agreement (SAT kappa coefficient = 0.82, $P < 0.001$; VAT kappa coefficient = 0.85, $P < 0.001$; SKM HU kappa coefficient = 0.85, $P < 0.001$).

Concordance between single-slice and multi-slice survival estimates

A total of 438 deaths occurred during a median follow-up period of 4.40 years (maximum 8.00 years).

Table 3 compares the associations with all-cause mortality for each tissue type between single-slice and multi-slice continuous z-scores adjusting for age, stage, tumour site, sex, race/ethnicity and height. The HRs between the single-slice and multi-slice continuous z-scores had overlapping CIs and similar magnitude and direction of effects regardless of the tissue examined. As expected, higher skeletal muscle quantity and radiodensity were associated with lower mortality risk (the HR was 0.82; 95% CI: 0.72–0.92 per SD SKM and 0.70; 95% CI: 0.64–0.78 SKM radiodensity when examining single-slice areas, compared with 0.85; 95% CI: 0.75–0.96 per SD SKM and 0.71; 95% CI: 0.64–0.78 when examining multi-slice volumes), whereas higher IMAT was associated with higher mortality risk (the HR was 1.16; 95% CI: 1.06–1.27 per SD when examining single-slice areas, compared with 1.13; 95% CI: 1.03–1.25 when examining multi-slice volumes). The magnitude and direction of associations remained similar comparing the single-slice and multi-slice body composition measurements regardless of whether models included each tissue separately, mutually adjusted all tissues or when exposures were operationalized categorically (data not shown).

Sensitivity analyses

When we further excluded scans that were flagged during quality review for any reason (n = 159), we found that corre-

Table 3 Associations of body composition measurements and all-cause mortality after colorectal cancer: hazard rations (HRs) and 95% confidence intervals (CIs) from single-tissue models fit separately for each sex-specific z-score for each single-slice area (cm^2) and multi-slice volume (cm^3) for each tissue (2871 colorectal cancer patients diagnosed 2012–2017 at Kaiser Permanente Northern California at diagnosis)

	Single-tissue models			
	Single-slice mid-L3 area		Multi-slice T12–L5 volume	
Exposure per SD	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Area and volume measurements				
Skeletal muscle	0.82 (0.72–0.92)	<0.05	0.85 (0.75–0.96)	<0.05
Subcutaneous adipose	0.95 (0.86–1.05)	0.352	0.96 (0.87–1.07)	0.497
Visceral adipose	1.04 (0.95–1.15)	0.373	1.03 (0.94–1.13)	0.537
Intermuscular adipose	1.16 (1.06–1.27)	<0.05	1.13 (1.03–1.25)	<0.05
Radiodensity measurements				
Skeletal muscle	0.70 (0.64–0.78)	<0.001	0.71 (0.64–0.78)	<0.001

Note: Models adjusted for the following covariates: age at diagnosis cancer stage, primary tumour site, sex, race/ethnicity and height at scan. All models fit separately for each tissue type exposure without mutual adjustment. Abbreviation: SD, sex-specific standard deviation units.

lations, LOA from Bland–Altman analysis and relative risks were nearly identical to our main analyses; for example, for SKM and IMAT, the correlations between single-slice and multi-slice measurements remained at R = 0.92 and R = 0.89, respectively. The only notable change was the narrowing of the LOA for IMAT from ± 0.92 to −0.86 to 0.88 SD. Scaling multi-slice tissue volumes to torso length by dividing by the height of the slab from T12 to L5 made a greater difference, improving correlations slightly to R = 0.93 for SKM and R = 0.91 for IMAT. Consistent with this, scaling removed some variability in the multi-slice volumetric measurements, resulting in narrower LOA for all tissues: The upper limits changed to ± 0.40 SD for VAT, to ± 0.40 SD for SAT, ± 0.75 SD for SKM and ± 0.83 SD for IMAT. No difference in mortality associations was noted for either sensitivity analysis.

Discussion

This is the first study to compare abdominal body composition assessed using single-slice areas at the mid-L3 to multi-slice volumes from T12 to L5; we found that these two approaches result in highly correlated measures with similar magnitude and direction of associations with mortality after diagnosis of colorectal cancer. However, there is not perfect agreement between the approaches, particularly for IMAT and SKM where for some individuals, the single-slice and multi-slice measurements rank them very differently relative to the population mean. Thus, although single-slice and multi-slice segmentation of body composition may yield similar results in population studies, there is heterogeneity in tissue quantities even across a limited field of view. Our sensitivity analyses indicate that scaling multi-slice measurements to the height of the slab may be an important step to standardize these metrics for analysis.

DAFS is the first platform to be commercially available that provides fully automated, multi-slice, multi-tissue segmentation of CT images, enabling the comparison of single-slice area measurements to multi-slice volumetric measurements. To date, the concordance of single-slice L3 and abdominal multi-slice T12–L5 measurement approaches or their associations with clinical outcomes have not been compared. Single-slice CT has been used for nearly two decades as a reference method for the measurement of body composition. Yet, single-slice CT measurements, while highly correlated with whole-body volumes from MRI, have high intra-subject variability across slices.[24] For population studies, this degree of measurement error is overcome with larger sample sizes,[25] but for individual estimation of tissue volumes or change over time within an individual,[26] this degree of error in estimating whole-body measurements from single-slice area measurements is less than ideal and may be better compensated for by using multi-slice segmentation across larger fields of view. Indeed, for some individuals in our study, the discordance between single-slice and multi-slice measurements was profound, and for most tissue areas, the LOA were wide.

Our study confirms that for population research assessing body composition at a single timepoint, single-slice area measures and multi-slice volumes are highly concordant and have similar associations with all-cause mortality; thus, both approaches can be used interchangeably. Indeed, we replicated the known associations of lower SKM quantity[2,27,28] and radiodensity[29–34] with lower survival using both measurement approaches. With the routine use of CT for surveillance and surgical planning, body composition measurements can be derived for many patients and are consistently associated with a host of clinical outcomes. However, given that the use of automated methods makes multi-slice volumes readily accessible, future studies should examine their utility for individual patient evaluation and for tracking longitudinal changes in tissues over time in response to intervention or disease progression. Further, these

novel metrics may also prove important for evaluating the prognostic importance of other body regions and compartments (e.g., the chest and pericardial adipose tissue) or evaluating the distribution of tissues such as muscle or IMAT across the abdomen or whole body. As an example of this emerging work, a recent study found that inclusion of measurements across multiple organs, tissues and slices derived from automated CT segmentation methods (aortic calcification, SKM radiodensity, ratio of VAT:SAT, and liver and bone mineral density) improved prediction of cardiovascular events.[12] Yet, much work remains to understand how the large number of correlated measurements derived from multi-slice CT segmentation can be summarized, standardized and integrated into statistical models to promote their use in clinical research.

Strengths and limitations

Ours is the first study to compare the concordance between single-slice area and multi-slice volume for quantifying body composition from abdominal CT scans; undertaking this effort in a large population of colorectal cancer patients using an automated body composition platform is a notable strength that is important given the increasing availability and accuracy of automated CT segmentation. Several limitations inherent to clinically acquired CT data must be noted. First, we compared single-slice areas only at the mid-L3 (a common reference method in the body composition literature) to multi-slice volumes from T12 to L5 (a commonly available field of view in clinical imaging). Thus, the comparison was a pragmatic one intended to inform the field of body composition research, and not a comprehensive study of concordance between single-slice areas at every vertebral landmark with regional or whole-body volumes, which are not commonly obtained in the clinic. In addition, there are no standardized methods for modelling single-slice area or multi-slice volume of body composition as this is an emerging field. Thus, we chose to focus on continuous z-scores to enable comparison across measures with different scales; however, this means that patients are ranked according to the sex-specific population mean of our cohort rather than an external reference standard and were not scaled to body size other than by adjusting for patient height and slab height in sensitivity analyses.

Conclusions and future directions

We found that single-slice and multi-slice body composition measurements are equivalent tools for population studies in colorectal cancer. Automated segmentation provides rapid, accurate and cost-effective methods for high-throughput analysis of body composition across multiple anatomical areas. To further advance research, we need to understand which body composition features available from multi-slice segmentation are most predictive of clinical outcomes; how these measurements can be integrated into inferential and predictive models to determine which are useful clinically and in what settings they can guide decision making; and how these measurements can be standardized across differing fields of view or relate to whole-body tissue volumes. Such investigations will not only harness the wealth of data available in these clinical images to better understand the importance of body composition to patient outcomes, but increasing automation and standardization is a prerequisite to integrating these data into clinical care.

Acknowledgements

The authors of this manuscript certify that they comply with the ethical guidelines for authorship and publishing in the *Journal of Cachexia, Sarcopenia and Muscle*.[35]

Funding

This work was supported by grants from the National Institutes of Health: R01AG065334 and K01CA226155. Dr. Lenchik's time on this project was supported by R21AG070804 and P30AG021332. Dr. Giri was supported by the Walter B Frommeyer Fellowship in Investigative Medicine at UAB and the International Myeloma Society Career Development Award.

Conflict of interest

Ijeamaka Anyene, Bette Caan, Grant R. Williams, Leon Lenchik, Smith Giri and Elizabeth M. Cespedes Feliciano declare that they have no conflict of interest.

Mirza Faisal Beg is a co-founder and actively directs Voronoi Health Analytics Incorporated, a Canadian corporation that sells commercial licences for the DAFS software.

Karteek Popuri is a co-founder and actively directs Voronoi Health Analytics Incorporated, a Canadian corporation that sells commercial licences for the DAFS software. Vincent Chow is a founding member and oversees operations in Voronoi Health Analytics Incorporat.

Online supplementary material

Additional supporting information may be found online in the Supporting Information section at the end of the article.

References

1. Dong QT, Cai HY, Zhang Z, Zou HB, Dong WX, Wang WB, et al. Influence of body composition, muscle strength, and physical performance on the postoperative complications and survival after radical gastrectomy for gastric cancer: a comprehensive analysis from a large-scale prospective study. *Clin Nutr* 2021;**40**: 3360–3369.
2. Trejo-Avila M, Bozada-Gutierrez K, Valenzuela-Salazar C, Herrera-Esquivel J, Moreno-Portillo M. Sarcopenia predicts worse postoperative outcomes and decreased survival rates in patients with colorectal cancer: a systematic review and meta-analysis. *Int J Colorectal Dis* 2021; **36**:1077–1096.
3. Xiao J, Caan BJ, Cespedes Feliciano EM, Meyerhardt JA, Peng PD, Baracos VE, et al. Association of low muscle mass and low muscle radiodensity with morbidity and mortality for colon cancer surgery. *JAMA Surg* 2020;**155**:942–949.
4. Drami I, Pring ET, Gould L, Malietzis G, Naghibi M, Athanasiou T, et al. Body composition and dose-limiting toxicity in colorectal cancer chemotherapy treatment; a systematic review of the literature. Could muscle mass be the new body surface area in chemotherapy dosing? *Clin Oncol (R Coll Radiol)* 2021;**33**:e540–e552.
5. Shachar SS, Williams GR, Muss HB, Nishijima TF. Prognostic value of sarcopenia in adults with solid tumours: a meta-analysis and systematic review. *Eur J Cancer* 2016;**57**:58–67.
6. Malietzis G, Aziz O, Bagnall NM, Johns N, Fearon KC, Jenkins JT. The role of body composition evaluation by computerized tomography in determining colorectal cancer treatment outcomes: a systematic review. *Eur J Surg Oncol* 2015;**41**: 186–196.
7. Bundred J, Kamarajah SK, Roberts KJ. Body composition assessment and sarcopenia in patients with pancreatic cancer: a systematic review and meta-analysis. *HPB (Oxford)* 2019;**21**:1603–1612.
8. Bridge CP, Rosenthal M, Wright B, Kotecha G, Fintelmann F, Troschel F, et al. Fully-automated analysis of body composition from CT in cancer patients using convolutional neural networks.
9. Dabiri S, Popuri K, Cespedes Feliciano EM, Caan BJ, Baracos VE, Beg MF. Muscle segmentation in axial computed tomography (CT) images at the lumbar (L3) and thoracic (T4) levels for body composition analysis. *Comput Med Imaging Graph* 2019;**75**: 47–55.
10. Higgins MI, Marquardt JP, Master VA, Fintelmann FJ, Psutka SP. Machine learning in body composition analysis. *Eur Urol Focus* 2021;**7**:713–716.
11. Koitka S, Kroll L, Malamutmann E, Oezcelik A, Nensa F. Fully automated body composition analysis in routine CT imaging using 3D semantic segmentation convolutional neural networks. *Eur Radiol* 2021;**31**: 1795–1804.
12. Pickhardt PJ, Graffy PM, Zea R, Lee SJ, Liu J, Sandfort V, 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.
13. Ma D, Chow V, Popuri K, Beg MF. Comprehensive validation of automated whole body skeletal muscle, adipose tissue, and bone segmentation from 3D CT images for body composition analysis: towards extended body composition. arXiv preprint arXiv:210600652; 2021.
14. Daly LE, Prado CM, Ryan AM. A window beneath the skin: how computed tomography assessment of body composition can assist in the identification of hidden wasting conditions in oncology that profoundly impact outcomes. *Proc Nutr Soc* 2018;**77**:135–151.
15. Cespedes Feliciano EM, Popuri K, Cobzas D, Baracos VE, Beg MF, Khan AD, et al. Evaluation of automated computed tomography segmentation to assess body composition and mortality associations in cancer patients. *J Cachexia Sarcopenia Muscle* 2020;**11**:1258–1269.
16. Ross TR, Ng D, Brown JS, Pardee R, Hornbrook MC, Hart G, et al. The HMO Research Network Virtual Data Warehouse: a public data model to support collaboration. *EGEMS (Wash DC)* 2014; **2**:1049.
17. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)* 2015;**25**: 141–151.
18. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;**22**:276–282.
19. Wickham H, Averick M, Bryan J, Chang W, D'Agostino McGowan L, Francois R, et al. Welcome to the Tidyverse. *J Open Source Softw* 2019;**4**:1686.
20. Therneau T. A package for survival analysis in R. R package version 3.3–1. 2022. https://CRAN.R-project.org/package=survival
21. Pedersen TL. patchwork: the composer of plots. R package version 1.1.1. 2020. https://CRAN.R-project.org/package=patchwork. 2020.
22. Zhu H, Travison T, Tsai T, Beasley W, Xie Y, Yu G, et al. kableExtra: construct complex table with 'kable' and pipe syntax. R package 1.3.4. 2021. https://CRAN.R-project.org/package=kableExtra
23. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. *R J* 2021;**13**:570–580.
24. Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge MP, Albu J, et al. Visceral adipose tissue: relations between single-slice areas and total volume. *Am J Clin Nutr* 2004;**80**:271–278.
25. Shen W, Punyanitya M, Wang Z, Gallagher D, Onge MP, Albu J, et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol (1985)* 2004;**97**:2333–2338.
26. Shen W, Chen J, Gantz M, Velasquez G, Punyanitya M, Heymsfield SB. A single MRI slice does not accurately predict visceral and subcutaneous adipose tissue changes during weight loss. *Obesity (Silver Spring)* 2012;**20**:2458–2463.
27. Caan BJ, Meyerhardt JA, Kroenke CH, Alexeeff S, Xiao J, Weltzien E, et al. Explaining the obesity paradox: the association between body composition and colorectal cancer survival (C-SCANS Study). *Cancer Epidemiol Biomarkers Prev* 2017;**26**:1008–1015.
28. Sun G, Li Y, Peng Y, Lu D, Zhang F, Cui X, et al. Can sarcopenia be a predictor of prognosis for patients with non-metastatic colorectal cancer? A systematic review and meta-analysis. *Int J Colorectal Dis* 2018;**33**:1419–1427.
29. Kroenke CH, Prado CM, Meyerhardt JA, Weltzien EK, Xiao J, Cespedes Feliciano EM, et al. Muscle radiodensity and mortality in patients with colorectal cancer. *Cancer* 2018;**124**:3008–3015.
30. Lee CM, Kang J. Prognostic impact of myosteatosis in patients with colorectal cancer: a systematic review and meta-analysis. *J Cachexia Sarcopenia Muscle* 2020; **11**:1270–1282.
31. Park IK, Yang SS, Chung E, Cho ES, Lee HS, Shin SJ, et al. Skeletal muscle gauge as a prognostic factor in patients with colorectal cancer. *Cancer Med* 2021;**10**: 8451–8461.
32. Shirdel M, Andersson F, Myte R, Axelsson J, Rutegård M, Blomqvist L, et al. Body composition measured by computed tomography is associated with colorectal cancer survival, also in early-stage disease. *Acta Oncol* 2020;**59**:799–808.
33. van Baar H, Beijer S, Bours MJL, Weijenberg MP, van Zutphen M, van Duijnhoven FJB, et al. Low radiographic muscle density is associated with lower overall and disease-free survival in early-stage colorectal cancer patients. *J Cancer Res Clin Oncol* 2018;**144**:2139–2147.
34. van Baar H, Winkels RM, Brouwer JGM, Posthuma L, Bours MJL, Weijenberg MP, et al. Associations of abdominal skeletal muscle mass, fat mass, and mortality among men and women with stage I–III colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2020;**29**:956–965.
35. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the *Journal of Cachexia, Sarcopenia and Muscle*: update 2021. *J Cachexia Sarcopenia Muscle* 2021;**12**:2259–2261.