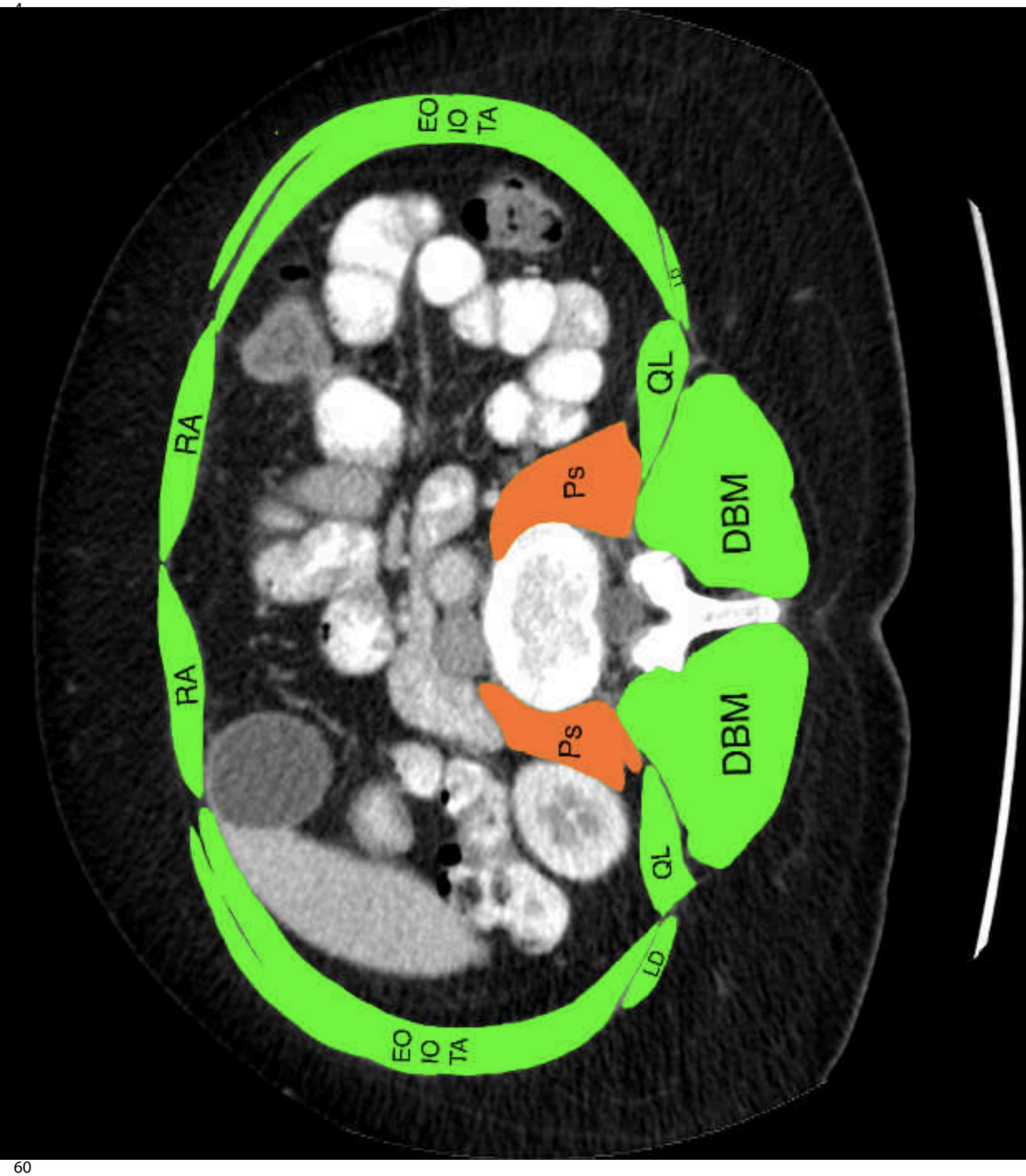

Article:
Dale MacLaine, T, Baker, O, Omura, M et al. (3 more authors) (2022) Prospective comparison of two methods for assessing sarcopenia and interobserver agreement on retrospective CT images. Postgraduate Medical Journal. ISSN 0032-5473

https://doi.org/10.1136/postmj/postgradmedj-2021-141301

Postgraduate Medical Journal

A prospective comparison of two methods for assessing sarcopenia and interobserver agreement on retrospective CT images

Journal:	*Postgraduate Medical Journal*
Manuscript ID	postgradmedj-2021-141301.R2
Article Type:	Original research
Date Submitted by the Author:	n/a
Complete List of Authors:	Dale MacLaine, Thomas; University of Leeds, Medical School Baker, Oliver; Leeds Teaching Hospitals NHS Trust Omura, Miyuki; Leeds Teaching Hospitals NHS Trust Clarke, Christopher; Nottingham University Hospitals NHS Trust Howell, Simon; Leeds Teaching Hospitals NHS Trust Burke, Dermot; Leeds Teaching Hospitals NHS Trust
Keywords:	Colorectal surgery < SURGERY, Diagnostic radiology < RADIOLOGY & IMAGING, GERIATRIC MEDICINE


EO
IO
TA
LD
QL
DBM
Ps
RA
RA
Ps
DBM
QL
LD
EO
IO
TA

Title

A prospective comparison of two methods for assessing sarcopenia and interobserver agreement on retrospective CT images in a colorectal surgical population

Authors:

Thomas Dale MacLaine[1] – Corresponding author

Oliver Baker[2]

Miyuki Omura[2]

Christopher Clarke[3]

Simon J. Howell[2]

Dermot Burke[2]

1- University of Leeds, Leeds, UK
2- Leeds Teaching Hospitals NHS Trust, Leeds, UK
3- Nottingham University Hospitals NHS Trust, Nottingham, UK

Corresponding author address: Dr Thomas Dale MacLaine, Medical School, University of Warwick Warwick Medical School, Coventry, Coventry, UK; thomas.dale-maclaine@warwick.ac.uk

Keywords: MeSH Terms

Sarcopenia
Colorectal surgery
X-Ray Computed Tomography
Radiologists

Word count: 2997/3000
Tables/illustrations: 6/6

Twitter: Thomas Dale MacLaine, @DrTaDM

Main Messages

- CT-determined sarcopenia can be identified by junior clinicians, those with anatomical understanding, and radiologists.
- Our data indicated poor association between sarcopenia and adverse surgical outcomes in our colorectal population.
- Currently available cut-offs require refinement for potential confounding factors, to provide more valuable clinical information.

Research questions arising from this work:

1. Is there intra-rator reliability in clinically trained assessors of CT-derived sarcopenia?
2. What is the clinical value of pre-operative sarcopenia identification in colorectal surgical patients?
3. What are the appropriate cut-off points for CT-derived sarcopenia in available methodologies for colorectal surgical patients?
4. Do CT-derived sarcopenia analyses for colorectal surgical patients need to be adjusted for confounding variables, beyond that of patient sex?

Key Messages

One of the strengths of this study is its comparison of the use of two radiological assessments for sarcopenia, where much work suggests the use of these measures

A further strength is that it examines whether these assessments can reliably predict surgical outcomes among colorectal cancer patients

Additionally, this study takes a novel look at the agreement between non-specialist assessors in identifying sarcopenia on CT scans, which would have implications for its routine clinical application.

The limitations of this study include its retrospective nature and the fact that it is a single centre study.

Abbreviations

CT – Computed Tomography

EWGSOP - European Working Group of Sarcopenia in Older Persons

OR – Odds Ratio

PA – Psoas Area

PACS - Picture Archiving and Communication System

ROC – Receiver Operating Characteristic

SD – Standard Deviation

TCSA – Total Cross-sectional Area

Statements

Author contributions:

Design of the study – TDM, SH, DB

Data collection – TDM, OB, MO, CC

Analysis – TDM, OB, MO

Write up - TDM

Editing – TDM, OB, MO, CC, SH, DB

Supervision – SH, DB

Funding:

This study was completed as part of a doctorate research programme, and as such received no funding.

Competing interests:

Authors have no duality of interests to declare.

Acknowledgements:

The authors would like to acknowledge the St James' University Hospital Colorectal Research team and the St James' University Hospital radiology department for the use of their equipment to undertake this study.

Abstract

Purpose of the study

To compare the relationships between two computed tomography (CT) derived sarcopenia assessment methods, and compare their relationship with inter- and intra-rater validations and colorectal surgical outcomes.

Study design

157 CT scans were identified across Leeds Teaching Hospitals NHS Trust for patients undergoing colorectal cancer surgery. 107 had body mass index data available, required to determine sarcopenia status. This work explores the relationship between sarcopenia, as measured by both total cross sectional-area (TCSA) and psoas-area (PA), and surgical outcomes. All images were assessed for inter- and intra-rater variability for both TCSA and PA methods of sarcopenia identification. The raters included a radiologist, an anatomist, and two medical students.

Results

Prevalence of sarcopenia was different when measured by PA (12.2%-22.4%) in comparison to TCSA (60.8%-70.1%). Strong correlation exists between muscle areas in both TCSA and PA measures, however there were significant differences between methods after the application of method-specific cut-offs. There was substantial agreement for both intra-rater and inter-rater comparisons for both TCSA and PA sarcopenia measures. Outcome data were available for 99/107 patients. Both TCSA and PA have poor association with adverse outcomes following colorectal surgery.

Conclusions

CT-determined sarcopenia can be identified by junior clinicians, those with anatomical understanding, and radiologists. Our study identified sarcopenia to have a poor association with adverse surgical outcomes in a colorectal population. Published methods of identifying sarcopenia are not translatable to all clinical

populations. Currently available cut-offs require refinement for potential confounding factors, to provide more valuable clinical information.

Abstract word count: 248/250

Introduction

The number of elderly patients diagnosed with cancer has increased in recent years, with significantly more being considered for surgical management. Poor agreement in geriatric syndromes, increased complexity of patients and the heterogeneity of the elderly population have led to subsequent increased interest into the mechanisms behind biological age, such as sarcopenia[1,2].

Sarcopenia is the progressive loss of skeletal muscle mass and strength, associated with increasing age and frailty[3,4]. Previous studies indicate that sarcopenia is a good predictor of post-operative outcomes and mortality[5,6]. Abdominal computed tomography (CT) analysis has been validated in its use for identifying sarcopenia and shown to provide details regarding muscle composition[6,7]. Consequently, CT derived sarcopenia assessments in surgical populations have been widely adopted[8].

Variations exist in approaches to quantifying muscle mass using CT scans. These often involve the selection of an axial CT image, measuring muscle surface area, then adjusting the value using the patient's height or BMI[9]. However, controversy exists between studies as to which muscles to include when assessing sarcopenia[10]. There are two widely used methods of assessing sarcopenia on abdominal CT: total transverse cross-sectional area (TCSA)[11], and total psoas area (PA)[6]. This study aims to identify whether PA or TCSA is a more reliable method of calculating sarcopenia in a colorectal surgical population.

Some works have explored inter-rater and intra-rater variation in identifying sarcopenia[12]. However, few studies have looked at this among colorectal cancer patients and extended this to include non-specialists assessing CT scans[12]. There may be variations in how measurements are taken for PA and TCSA measurements. This study explores the inter- and intra-rater agreement of sarcopenia assessments, with both radiologist and non-radiologist raters.

Most colorectal surgical patients and specifically colorectal cancer patients routinely receive pre-operative CT scans[13]. This provides an opportunity to investigate for sarcopenia and assess risk of complications, without incurring any additional costs, or radiation exposure to the patient. Whilst evidence suggests sarcopenia is associated with surgical outcomes[5,14], this varies across surgical populations[15] and there are no clear systematic reviews in this field to date. We explore the differences between PA and TCSA methods of identifying sarcopenia and their relationship with adverse surgical outcomes.

Methodology

This single-centre retrospective study assessed CT scans of adult elective colorectal surgical patients discussed during multi-disciplinary team (MDT) meetings at Leeds Teaching Hospitals NHS Trust from July to December 2015. University of Leeds Medicine and Health Research Ethics Committee granted ethical approval (MREC16-099). CT scans of adult elective colorectal cancer patients were used. Patients were excluded if their CT scan was taken over six months prior to the surgical procedure. Images were analysed using axial CT images on Trust-calibrated Picture Archiving and Communication System (PACS) monitors.

CT technique and image analysis

Measurements were taken from the CT-scan slice closest to the most inferior aspect of the L3 vertebral body, either using a corresponding coronal view or measured from the sacrum moving superiorly after assessing for lumbar-sacral vertebral fusion variations[16]. Axial images were chosen where the vertebral body was predominantly cortical bone, with minimal intervertebral disc present. All muscle boundaries in this image were hand-drawn on PACS, guided by their density. Cross-sectional areas were individually measured, resulting in a computed value for total cross-sectional area. An example of sarcopenia measurements can be seen in figure 1.

TCSA measurement of sarcopenia:

Based on that of Baracos and colleagues[11], assessors traced outlines of all lumbar muscle groups on the CT image. Muscle areas were calculated and summed, giving total muscle cross sectional muscle (figure 1). This value was then normalised, using the patient's height: [TCSA (cm^2) / 100] / height^2 (m^2). This normalised value was subjected to sex-specific cut-offs to distinguish between sarcopenic and non-sarcopenic individuals, with the cut off for males as 55.4 cm^2/m^2 and females as 38.9 cm^2/m^2.

PA measurement of sarcopenia:

On the same images, paired psoas muscle areas were outlined (figure 1) as described by Jones and colleagues[6]. The cross-sectional areas were summed together, then normalised using the patient's height: PA (mm^2)/ height^2 (m^2). This normalised value was subjected to sex-specific cut-offs to determine the presence of sarcopenia, with the cut off for males as 542 mm^2/m^2 and females as 385 mm^2/m^2.

Inter and intra-rater agreements

CT scans were analysed independently by a gastrointestinal radiology fellow, an anatomist and two final-year medical students. All received identical training on identifying TCSA and PA on CT and followed a standard operating procedure for assessing the CT scans. Assessors were blinded to each other's results, as well as patient outcomes. The anatomist re-analysed CT scans to investigate intra-rater agreement, after a 6-month window separating the initial and second readings, and without access to the results of the first measures.

Surgical outcomes

Surgical outcome data were collected from patient electronic records, operation notes and the MDT database. These data were collected by research nurses who were blinded to the sarcopenia measurements. Outcomes included length of stay, the post-operative level of care required, re-admission rates and whether the patient experienced complications. Complications were grouped and graded in severity using the Clavien-Dindo classification[17].

Statistical analyses

There is minimal information available in the literature on the inter-rater agreement in sarcopenia measures, so estimate values were adopted. Two-tailed sample size calculations were performed in Stata[18]. The power of the rater variability was set to 0.9, with a significance level of 0.01. The standard deviation within groups was set at the maximal value of one, and an estimation of a standard deviation of 0.2 was used for the variance between groups. For the rater variability, the sample size was calculated to be 27. Literature available around the hazards of sarcopenia for post-operative outcomes is unclear[5,8,19], so a hazards ratio of 4.5 was used as an estimate. For the association with post-operative outcomes, the sample size was calculated to be 44. This study aimed to collect a minimum of 70 patients to meet the requirements for both rater variability and to assess for relationships with post-operative outcomes.

Statistical plan for the comparison of radiological assessments

Analyses were performed using SPSS[20]. Associations between TCSA and PA cross-sectional areas, before normalisation, were examined using Pearson's correlation statistic. These comparisons were drawn to aid in determining whether differences found between the two methods occur because of the measurements or from the cut-offs involved in the methods.

The Chi-squared statistic was used to analyse the differences in prevalence of sarcopenia between the two methods. Paired t-tests were used to analyse differences in left and right paired muscles. The agreement between TCSA and PA methods were analysed using Cohen's Kappa statistic. Pairwise Kappa statistics were used to determine the agreement of sarcopenia identification between the four assessors.

Sensitivity, specificity, negative and positive predictive values and likelihood ratios were used to analyse the relationship between sarcopenia measures and surgical outcomes. Missing data was handled through pair-wise deletion.

Results

From the 110 eligible patients identified from the MDT database, CT scans were available for 107. There were 49/107 (45.8%) females, the mean(SD) age was 64.2(13.6) with 56.4% being aged 65 or older, and the mean BMI was 27.5kg/m^2.

The total number of patients identified as sarcopenic per assessor and approach is detailed in Table 1. Prevalence of sarcopenia was 12.2%-22.4% when assessed by the PA approach, and 60.8%-70.1% when assessed by the TCSA approach. There was a significant difference identified by the Chi squared statistic between PA and TCSA for the anatomist (A) (x^2=10.5, p=0.001), medical students (MS) one and two (x^2=9.9, p=0.002; x^2=9.5, p=0.004 respectively), and the radiologist (R) (x^2=16.0, p<0.001). There was slight agreement between PS and TCSA as measured by A (K=0.18, p=0.001), MS1 (K=0.17, p=0.002) and MS2 (K=0.15, p=0.004), with fair agreement measured by R (K=0.28, p<0.001). Males were more sarcopenic than females according to the TCSA assessment (MS1 p=0.001, MS2 p=0.014, A p=0.002, R p=0.007), however no difference was observed with the PA methods (MS1 p=0.859, MS2 p=0.226, A p=0.677, R p=0.646).

Table 1: Prevalence of CT-derived sarcopenia as measured by the four assessors involved in this study using two different methods.

	Healthy (%)	Sarcopenic (%)
TCSA method		
Medical Student 1	39 (36.5)	68 (44.8)
Medical Student 2	37 (34.6)	70 (65.4)
Radiologist	42 (39.3)	65 (60.7)
Anatomist	32 (29.9)	75 (70.1)
Psoas method		
Medical Student 1	94 (87.0)	16 (13.0)
Medical Student 2	91 (85.0)	16 (15.0)
Radiologist	83 (77.6)	24 (22.4)
Anatomist	87 (81.3)	20 (18.7)

TCSA – Total Cross-Sectional Area.

The mean(SD) muscle mass for TCSA and PA ranged from 12344.45(36.87)-12884.67(36.74)mm^2 and 1756.10(676.37)-1964.84(943.88)mm^2 respectively (Table 2). All assessors identified highly positive and significant correlations between the sum of psoas and total cross-sectional areas ($p<0.001$). This remained the case after the sum of psoas value was normalised by patients' heights ($p<0.001$).

Table 2: Average muscle mass for total cross-sectional area and psoas area methods of identifying sarcopenic patients before normalisation.

	Cross sectional area												
Methods	Ps left	Ps right	ES left	ES right	QL left	QL right	LD left	LD right	OT left	OT right	RA left	RA right	Sum
TCSA (SD)	892.3 (386.4)	877.1 (345.9)	2051.6 (569.79)	2002.5 (567.4)	473.2 (223.4)	434.3 (184.7)	85.0 (91.8)	81.9 (84.5)	2224.3 (824.3)	2280.2 (796.5)	496.7 (221.2)	486.0 (232.8)	12384.9 (4528.6)
PA (SD)	892.3 (386.4)	877.1 (345.9)											1769.4 (732.3)
T test P value (between left and right muscles)		0.6	0.4		0.4		0.7		0.5		0.4		

TCSA: total cross-sectional area, PA: psoas area, SD: standard deviation, Ps: Psoas, ES: Erector Spinae, QL: Quadratus Lumborum, LD: Latissimus Dorsi, OT: Obliques and Transverse muscles, RA: Rectus Abdominis, SD: Standard Deviation.

Results regarding the analysis of agreement can be seen in table 3. There was substantial to almost perfect agreement with sarcopenia identification using the TCSA assessment, across the four assessors (all analyses p<0.001). The PA assessment had slightly less agreement between raters, ranging from moderate to substantial agreement (all analyses p<0.001).

Table 3: Kappa grid, exploring the agreement between medical students, a radiologist and an anatomist, for two methods of identifying sarcopenia. Values highlighted represent the kappa statistic for the agreement between the two individuals- such as medical student 1 and medical student 2 agreeing almost perfectly K=0.808 for total cross-sectional area, and substantially K=0.751 for psoas-alone sarcopenia. A colour code key for the kappa analyses is provided with the table.

		Total cross-sectional area method of sarcopenia identification				Psoas area method of sarcopenia identification			
		Significance x 10^3				Significance x 10^3			
		MS1	MS2	R	A	MS1	MS2	R	A
Cohen's Kappa Agreement x 10^3	MS1		<001	<001	<001		<001	<001	<001
	MS2	808		<001	<001	751		<001	<001
	R	757	841		<001	615	468		<001
	A[1]	784	878	826		788	705	640	

Key for value of Kappa x10^3	Negative Disagree	000 – 200 None	201 – 400 Fair	401 – 600 Moderate	601 – 800 Substantial	801+ Almost perfect

Outcome data were available for 99/107 patients. The median(inter-quartile range) length of stay was 7(0-11.8) days, with the longest hospital stay lasting 156 days. Pre-operatively, 96 patients were predicted to require ward-based care, and 3 high-dependency care. There were 18 patients that required a change in care level post-operatively; 78 patients required ward-based care, 13 required high-dependency care and 8 required intensive care in the immediate post-operative period. Thirty-four patients who suffered post-operative complications. Complication severity was graded using Clavien-Dindo: 4 were Grade one, 15 were Grade two, 3 were Grade three, 8 were Grade four, and 4 were Grade five (death). Six patients were

1

readmitted within 30-days of discharge, with four requiring antibiotics for post-operative infections, one reattended with a dysfunctional colostomy and one was admitted for post-operative pain management. There were limited instances of significant differences between sarcopenia state and adverse outcome being identified by any assessor (table 4).

Table 4: Analysis of the differences between sarcopenia status and the presence of an adverse outcome, determined using Mann-Whitney U analyses. P values for analyses presented

		Extended Length of Stay	Bed Change	Readmission	Complication	Death
TCSA	A	0.018	0.141	0.913	0.229	0.180
	MS1	0.065	0.769	0.874	0.141	0.133
	MS2	0.012	0.309	0.794	0.079	0.117
	R	0.004	0.069	0.680	0.081	0.094
PA	A	0.02	0.949	0.202	0.227	0.002
	MS1	0.004	0.149	0.025	0.338	0.027
	MS2	0.034	0.309	0.034	0.031	0.012
	R	0.100	0.454	0.411	0.553	0.153

Mann-Whitney U tests, p values

Contingency analyses highlighted that risk ratios associated with each outcome, for both TCSA and PA, were inconsistent (table 5). Risk ratios for all outcomes were not significant in TCSA and PA methods in both anatomist and medical student 1's data. Medical student 2 identified significant risk ratios for post-operative complications in both TCSA (RR 1.73, 95% CI 1.06-2.76) and PA (RR 2.07, 95% CI 1.26-3.41). The radiologist identified that the TCSA has statistically significant risk ratios for patients requiring an increased post-operative level of care (RR 2.47, 95% CI 1.01-6.08) and post-operative complications (RR 1.70, 95% CI 1.07-2.57). Risk ratios for mortality was unavailable for all assessors as there were no deaths seen in healthy patients scored by the TCSA as measured by all assessors. The likelihood ratios for TCSA and PA were close to unity for all outcome measures.

Table 5a, b, c, d: Comparison of total cross-sectional area and psoas area methods and their relations with outcomes. The four sub-tables tables a, b, c and d represent the data from the anatomist, medical student 1, medical student 2 and radiologist respectively.

5a - anatomist

Change of the in-hospital level of care required post-operatively							
Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	Likelihood ratio
TCSA	0.14	2.28 (0.81-6.43)	34.57	83.33	90.32	22.06	1.16
PA	0.95	1.04 (0.34-3.20)	83.905	16.67	81.93	18.75	1.01
Died							
Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.18	Unavailable	31.58	100	100	5.80	1.06
PA	0.002	14.47 (0.75-280.10)	85.26	75.00	98.78	17.65	1.20
Post-operative complications							
Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.23	1.48 (0.91-2.40)	35.38	76.47	74.19	38.24	1.20
PA	0.23	1.48 (0.85-2.59)	86.15	23.53	68.29	47.05	1.29
Re-admission within 30 days							
Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.91	0.91 (0.14-5.90)	31.18	66.67	93.55	5.88	0.99
PA	0.20	2.80 (0.48-16.22)	86.02	33.33	95.24	13.33	1.10

TCSA; Total cross-sectional area sarcopenia, PA; Psoas-alone sarcopenia, RR: Risk Ratio, CI: 95% Confidence Interval, PPV: Positive Predictive Value, NPV: Negative Predictive Value, LR: Likelihood Ratio.

5b - Medical student 1

Change of the in-hospital level of care required post-operatively

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	Likelihood ratio
TCSA	0.77	1.14 (0.52-2.50)	37.04	66.67	83.33	19.05	1.03
PA	0.15	2.07 (0.83-5.19)	90.12	22.22	83.91	33.33	1.26

Died

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.13	Unavailable	36.84	100	100	6.25	1.07
PA	0.03	6.62 (0.71-61.65)	88.42	50.00	97.67	15.38	1.15

Post-operative complications

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.14	1.59 (0.99-2.53)	41.54	73.53	75.00	39.68	1.24
PA	0.34	1.42 (0.75-2.67)	89.23	17.65	67.44	46.15	1.25

Re-admission within 30 days

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.87	1.14 (0.18-7.44)	36.56	66.67	94.44	6.35	1.01
PA	0.07	4.00 (0.72-22.37)	90.32	33.33	95.45	18.18	1.17

TCSA; Total cross-sectional area sarcopenia, PA; Psoas-alone sarcopenia, RR: Risk Ratio, CI: 95% Confidence Interval, PPV: Positive Predictive Value, NPV: Negative Predictive Value, LR: Likelihood Ratio.

5c - Medical student 2

Change of the in-hospital level of care required post-operatively

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	Likelihood ratio
TCSA	0.31	1.62 (0.71-3.69)	40.74	72.22	86.84	21.31	1.10
PA	0.31	1.78 (0.62-5.11)	91.36	16.67	83.15	30.00	1.19

Died

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.11	Unavailable	38.95	100	100	6.45	1.07
PA	0.01	8.00 (0.88-73.02)	90.53	50.00	97.73	18.18	1.19

Post-operative complications

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.08	1.73 (1.08-2.76)	44.62	73.53	76.32	40.98	1.29
PA	0.03	2.07 (1.26-3.41)	93.85	20.59	69.32	63.64	1.91

Re-admission within 30 days

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.79	1.25 (0.19-8.13)	38.71	66.67	94.74	6.56	1.01
PA	0.03	5.00 (0.92-27.11)	92.47	33.33	95.56	22.22	1.23

TCSA; Total cross-sectional area sarcopenia, PA; Psoas-alone sarcopenia, RR: Risk Ratio, CI: 95% Confidence Interval, PPV: Positive Predictive Value, NPV: Negative Predictive Value, LR: Likelihood Ratio.

5d - Radiologist

Change of the in-hospital level of care required post-operatively

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	Likelihood ratio
TCSA	0.07	2.47 (1.01-6.08)	45.68	77.78	90.24	24.14	1.19
PA	0.45	1.43 (0.59-3.48)	80.25	27.78	83.33	23.81	1.09

Died

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.09	Unavailable	42.11	100	100	6.78	1.07
PA	0.15	3.71 (0.38-36.07)	80.00	50.00	97.43	9.52	1.08

Post-operative complications

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.08	1.70 (1.07-2.67)	47.69	70.59	75.61	41.38	1.29
PA	0.55	1.22 (0.68-2.19)	81.54	23.53	67.09	40.00	1.19

Re-admission within 30 days

Method	Chi Square P values	RR (CI)	PPV	NPV	Sensitivity	Specificity	LR
TCSA	0.68	1.41 (0.22-9.25)	41.94	66.67	95.12	6.90	1.02
PA	0.41	1.98 (0.33-11.71)	80.65	33.33	94.94	10.00	1.06

TCSA; Total cross-sectional area sarcopenia, PA; Psoas-alone sarcopenia, RR: Risk Ratio, CI: 95% Confidence Interval, PPV: Positive Predictive Value, NPV: Negative Predictive Value, LR: Likelihood Ratio.

Discussion

Our data highlight that prevalence of sarcopenia is highly dependent on the method used; with between 12.2% and 70.1% of patients presenting for colorectal surgery classified as sarcopenic. This raises concerns regarding the use of CT-derived sarcopenia diagnoses. There is an increased risk of error in measuring TCSA, with 2

additional muscle bulks to measure, in comparison to the PA method. This may account for a small amount of the difference in prevalence between TCSA and PA. However, given the strong agreement in measurements across all assessors, concerns regarding precision of the TCSA measure are secondary.

There was a strong correlation between the cross-sectional areas measured by the both the PA and the TCSA methods. However, there were significant differences between the two methods in the identification of sarcopenia as identified by all four investigators. This strongly suggests that it is the cut-off levels for each method that cause the disparity.

Inter-rater agreement and variations

This study identified substantial to almost perfect agreement between the four investigators when assessing sarcopenia through the TCSA method, and moderate to substantial agreement when assessing sarcopenia using the PA method. These data are promising, suggesting that CT-derived sarcopenia measures can accurately be measured by junior clinical colleagues, which might prove valuable if such measures were to be integrated into clinical practice. Our data also suggest CT-derived sarcopenia measurements are reliably repeatable, adding to their construct validity.

Assessments in the literature use similar but not identical methods. Muscle cross-sectional area measurement is predominantly reported at the third lumbar level[9]. Whilst this appears to give a fixed anatomical landmark, there are differences in what structures are present at the third lumbar level. This may partially stem from which section of the vertebrae the measurement is taken from; the third lumbar vertebra has a mean vertebral body height of between 29mm and 30mm for females and 30mm to 31mm in males[21]. There is limited information available in the literature regarding how best to ensure consistency in the axial cross section analysed, regardless of which vertebrae is identified[22]. Methods may use L3 to identify the axial cross section, however with a vertebral height of approximately 29mm and images being taken between 1-5mm thickness[23], there could be between 6-29 different axial

images on which to identify muscle mass measurements. This variation in axial images may have not been identified in this study, as there was moderate to substantial agreement between assessors. However, this agreement may be improved further by ensuring the same axial image is identified through this methodology. This potential disparity needs consensus before the results can be universally understood.

Agreement studies are of particular interest in radiological studies – in two years, 280 studies were published with 81% investigating agreement with ≤2 raters[24]. A strength of this study is the breadth of raters involved in the assessment of sarcopenia through CT analyses. Our study explores intra-rater agreement of the anatomist, inter-rater agreement across three skill levels and evaluates the difference between two separate individuals of the same skill level. Farzin and colleagues [24] concluded that available literature shows a scarcity of agreement studies, with minimal raters and under-explored data. We provide valuable agreement data from four raters, with three different backgrounds, at one site with images from one population. It would be of further benefit to explore agreement across other institutions, differing clinical populations and additional intra-disciplinary assessors in order to validate these findings.

Association with adverse surgical outcomes

Both methods of identification of sarcopenia gave a high number of false positives, as many patients were identified as sarcopenic, but did not suffer post-operative complications. This resulted in a poor predictive association between sarcopenia and adverse outcomes. This may be related to the cut-off levels that were used, or to the limited number of adverse outcomes recorded, in this study. Further work is needed to identify whether cut-off levels identified in similar studies apply to individual populations or if they can be used more widely. For example, body anthropometrics differ between ethnicities[25], and thereby between populations. Neither our study, nor many similar studies, adjust for ethnicity. Such adjustments could be important when comparing levels of sarcopenia between different populations, as has been previously suggested for BMI[25].

Cross-sectional area is arguably not a measure of muscle mass. The cross-sectional areas identified by only CT analyses do not take into consideration the composition of the muscle within the areas identified. It may be that some muscles are non-uniformly infiltrated with fat deposits, which might impact the quality of the muscle itself[26]. Quality of muscle function, and thereby ability to recover from surgery, might be better assessed by traditional measures such as grip strength and gait speed. The results of the CT-derived sarcopenia assessments would benefit from being compared to the complete EWGSOP criteria, to determine how valid these assessments are as an abbreviated measure of EWGSOP sarcopenia.

With regards to the validation of sarcopenia measurements, several publications have reported an association between of CT-derived measures of sarcopenia and post-operative outcomes[5,14,27]. Data from this study do not show the same association. In this study the likelihood ratios for all outcome measures, measured by all investigators, were close to unity. Some studies in sarcopenia adopt a cut-off determined by the lowest quartile of skeletal muscle area or by internally derived cut-offs modelled against outcomes[8,27,28]. Using internally derived cut-offs, modelled to fit observed data in individual studies, limits the realistic inferences that can be drawn and increases the likelihood of identifying a relationship with an outcome measure[29]. A key point made is that the over-estimation seen in this approach to deriving cut-offs is further amplified in small studies[30]. However, refining the cut-offs in either methodological approach may result in accurate prognostic value highlighting patients' risk for adverse outcomes, seen in several studies[5,14,27,28,31]. A prospective, multi-centre study with clear eligibility criteria that resulted in regression modelling would provide a good starting point. Larger, controlled studies could then explore potential confounding factors.

Limitations

This study is a retrospective analysis of CT scans for sarcopenia, from a single centred study with sample sizes limited by time-period. Single-centred trials lack external validity and are not an appropriate foundation to implement changes of

care[30]. The results of our study highlight the limitations of the TCSA and PA methods of identifying sarcopenia and work towards external validation.

The retrospective nature of this study is a potential limitation. Retrospective studies are often limited by 'missing data'[32]. This study is limited by the potential missing data and by using an MDT database to identify eligible patients. Retrospective studies are more prone to selection bias; the MDT database may not represent all relevant colorectal surgical patients but a significant proportion. This may reduce the applicability of the results of this study to inform clinical practice. Prospective cohort studies are needed to provide further validation to support the applications to a clinical population. Our study may be limited in terms of a small sample size[33] however the lack of effect suggests that a larger sample size would give similar results.

Conclusion

Sarcopenia is considered a geriatric disease that is reported to have an association with post-operative outcomes in a colorectal surgical setting. Literature recommends the adoption of CT derived sarcopenia assessments in the clinical setting. Our data suggest this technique is both replicable and allows assessment by clinicians with different expertise. However, our data also suggest that these assessments may not be as predictive of adverse outcomes for colorectal surgical patients as desired, in their current format. Higher-powered studies are needed to refine these CT-alone sarcopenia definitions to ensure there is consistency and clinical value to the results. Understanding the prevalence of sarcopenia by using more functional measures and a comparison of this with CT-derived sarcopenia measures would be useful to understand whether cross-sectional area could be used as a reliable short-cut method.

References

1. Dodds, R. M., Roberts, H. C., Cooper, C. & Sayer, A. A. The Epidemiology of Sarcopenia. *J. Clin. Densitom.* **18**, 461–466 (2015).
2. Dale Maclaine, T., Baker, O., Burke, D. & Howell, S. J. Prevalence of frailty and reliability of established frailty instruments in adult elective colorectal surgical

patients: A prospective cohort study. *Postgrad. Med. J.* 1–5 (2021). doi:10.1136/postgradmedj-2020-139417

3. Cruz-Jentoft, A. J. *et al.* Sarcopenia: Revised European consensus on definition and diagnosis. *Age Ageing* **48**, 16–31 (2019).
4. Morley, J. E., Anker, S. D. & von Haehling, S. Prevalence, incidence, and clinical impact of sarcopenia: facts, numbers, and epidemiology—update 2014. *J. Cachexia. Sarcopenia Muscle* **5**, 253–259 (2014).
5. Lieffers, J. R., Bathe, O. F., Fassbender, K., Winget, M. & Baracos, V. E. Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery. *Br. J. Cancer* **107**, 931–936 (2012).
6. Jones, K. I., Doleman, B., Scott, S., Lund, J. N. & Williams, J. P. Simple psoas cross-sectional area measurement is a quick and easy method to assess sarcopenia and predicts major surgical complications. *Color. Dis.* **17**, O20–O26 (2015).
7. Gibson, D. J., Burden, S. T., Strauss, B. J., Todd, C. & Lal, S. The role of computed tomography in evaluating body composition and the influence of reduced muscle mass on clinical outcome in abdominal malignancy: A systematic review. *European Journal of Clinical Nutrition* **69**, 1079–1086 (2015).
8. Englesbe, M. J. *et al.* Sarcopenia and post-liver transplant mortality. *J. Surg. Res.* **211**, 271–278 (2010).
9. Shen, Y., Hao, Q., Zhou, J. & Dong, B. The impact of frailty and sarcopenia on postoperative outcomes in older patients undergoing gastrectomy surgery: A systematic review and meta-analysis. *BMC Geriatr.* **17**, 1–8 (2017).
10. Vergara-Fernandez, O., Trejo-Avila, M. & Salgado-Nesme, N. Sarcopenia in patients with colorectal cancer: A comprehensive review. *World J. Clin. Cases* **8**, 1188–1202 (2020).
11. Baracos, V. Body composition in patients with non–small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am. J. Clin. Nutr.* **91**, 1133–1137 (2010).
12. Park, J. *et al.* Reliable and robust method for abdominal muscle mass quantification using CT/MRI: An explorative study in healthy subjects. *PLoS One* **14**, e0222042 (2019).
13. NICE Guidelines. Colorectal cancer: the diagnosis and management of

colorectal cancer. *NICE Clin. Guidel.* **131**, 1–37 (2011).

14. Wang, S. L. *et al.* Sarcopenia Adversely Impacts Postoperative Clinical Outcomes Following Gastrectomy in Patients with Gastric Cancer: A Prospective Study. *Ann. Surg. Oncol.* **23**, 556–564 (2016).
15. Tegels, J. J. W. *et al.* Sarcopenia is highly prevalent in patients undergoing surgery for gastric cancer but not associated with worse outcomes. *J. Surg. Oncol.* **112**, 403–407 (2015).
16. Vasuki, A. K. M. *et al.* Anatomical variations of the sacrum and its clinical significance. *Int. J. Anat. Res.* **4**, 1859–1863 (2016).
17. Clavien, P. a *et al.* The Clavien-Dindo classification of surgical complications: five-year experience. *Ann. Surg.* **250**, 187–196 (2009).
18. StataCorp. Stata Statistical Software. (2013).
19. Zhuang, C.-L. *et al.* Sarcopenia is an Independent Predictor of Severe Postoperative Complications and Long-Term Survival After Radical Gastrectomy for Gastric Cancer. *Medicine (Baltimore).* **95**, e3164 (2016).
20. SPSS Inc. IBM SPSS Statistics. (2016).
21. Zhou, S. H., McCarthy, I. D., McGregor, A. H., Coombs, R. R. H. & Hughes, S. P. F. Geometrical dimensions of the lower lumbar vertebrae - analysis of data from digitised CT images. *Eur. Spine J.* **9**, 242–248 (2000).
22. Cruz-Jentoft, A. J. *et al.* Sarcopenia: Revised European consensus on definition and diagnosis. *Age Ageing* **48**, 16–31 (2019).
23. Rubin, G. & Rofsky, N. *CT and MR Angiography: Comprehensive Vascular Assessment*. *CT and MR Angiography: Comprehensive Vascular Assessment* (2009).
24. Farzin, B. *et al.* Agreement studies in Radiology research. *Diagn. Interv. Imaging* **98**, 227–233 (2017).
25. Deurenberg, P., Yap, M. & Van Staveren, W. A. Body mass index and percent body fat: A meta analysis among different ethnic groups. *Int. J. Obes.* **22**, 1164–1171 (1998).
26. Hamaguchi, Y., Kaido, T., Okumura, S. & Kobayashi, A. Muscle Steatosis is an Independent Predictor of Postoperative Complications in Patients with Hepatocellular Carcinoma. *World J. Surg.* **40**, 1959–1968 (2016).
27. Abbass, T., Tsz Ho, Y. T., Horgan, P. G., Dolan, R. D. & McMillan, D. C. The

relationship between computed tomography derived skeletal muscle index, psoas muscle index and clinical outcomes in patients with operable colorectal cancer. *Clin. Nutr. ESPEN* **39**, 104–113 (2020).

28. Peng, P. *et al.* Impact of Sarcopenia on Outcomes Following Resection of Pancreatic Adenocarcinoma. *J. Gastrointest. Surg.* **16**, 1478–1486 (2012).
29. Potgieter, D. *et al.* N-terminal pro-B-type natriuretic peptides' prognostic utility is overestimated in meta-analyses using studyspecific optimal diagnostic thresholds. *Anesthesiology* **123**, 264–271 (2015).
30. Bellomo, R., Warrillow, S. & Reade, M. Why we should be wary of single-center trials. *Crit Care Med* **37**, 3114–9 (2009).
31. Peng, P. D. *et al.* Sarcopenia negatively impacts short-term outcomes in patients undergoing hepatic resection for colorectal liver metastasis. *Hpb* **13**, 439–446 (2011).
32. Anthonisen, N. R. Retrospective studies. *Can. Respir. J.* **16**, 117–8 (2009).
33. Trejo-Avila, M. *et al.* Sarcopenia predicts worse postoperative outcomes and decreased survival rates in patients with colorectal cancer: a systematic review and meta-analysis. *Int J Colorectal Dis.* **36**, 1077-1096 (2021).

List of Legends for Figures, Tables and Graphs

Figures

Figure 1: L3 computed tomography cross section highlighting all abdominal muscles involved in CT-derived sarcopenia methodology. RA: rectus abdominis. EO: external obliques. IO: internal obliques. TA: Transversus abdominis. LD: latissimus dorsi. QL: quadratus lumborum. Ps: psoas muscle group. DBM: deep back muscle group. The orange psoas muscles indicate those muscles involved in Psoas-area measurements of sarcopenia. All blue and orange muscles are involved in the total cross-sectional area measurements of sarcopenia. Image adapted, courtesy of Dr Andrew Dixon, Radiopedia.org rID:36677

Tables

Table 1: Prevalence of CT-derived sarcopenia as measured by the four assessors involved in this study using two different methods.

Table 2: Average muscle mass for total cross-sectional area and psoas area methods of identifying sarcopenic patients before normalisation.

Table 3: Kappa grid, exploring the agreement between medical students, a radiologist and an anatomist, for two methods of identifying sarcopenia. Values highlighted represent the kappa statistic for the agreement between the two individuals- such as medical student 1 and medical student 2 agreeing almost perfectly K=0.808 for total cross-sectional area, and substantially K=0.751 for psoas-alone sarcopenia. A colour code key for the kappa analyses is provided with the table

Table 4: Analysis of the differences between sarcopenia status and the presence of an adverse outcome, determined using Mann-Whitney U analyses. P values for analyses presented

Table 5a, b, c, d: Comparison of total cross-sectional area and psoas area methods and their relations with outcomes. The four sub-tables tables a, b, c and d represent the data from the anatomist, medical student 1, medical student 2 and radiologist respectively.