BJR

https://doi.org/10.1259/bjr.20180067

Received: 11 January 2018 | Revised: 18 March 2018 | Accepted: 02 May 2018

Cite this article as:
Waduud MA, Drozd M, Linton E, Wood B, Manning J, Bailey MA, et al. Influences of clinical experience in the quantification of morphometric sarcopaenia: a cohort study. *Br J Radiol* 2018; **091**: 20180067.

FULL PAPER

Influences of clinical experience in the quantification of morphometric sarcopaenia: a cohort study

[1]MOHAMMED ABDUL WADUUD, MRCS, MSc, [2]MICHAEL DROZD, MBChB, BSc, [1]EMMA LINTON, MBChB, [1]BENJAMIN WOOD, MBChB, BSc, [2]JAMES MANNING, [2]MARC A BAILEY, MRCS, PhD, [1]CHRISTOPHER HAMMOND, FRCR, MRCS, MA, [1]JULIAN A SCOTT, FRCS, MD and On behalf of the Vascular Surgeons and Interventional Radiologists at the Leeds Vascular Institute

[1]Leeds Vascular Institute, Leeds General Infirmary, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[2]Leeds Institute for Cardiovascular and Metabolic Medicine, University of Leeds, LIGHT Laboratories, Leeds, UK

Address correspondence to: Dr Mohammed Abdul Waduud
E-mail: *m.a.waduud@leeds.ac.uk*

Objective: The measurement of total psoas muscle area (TPMA) on CT imaging is commonly made using either manual tracing or a semi-automated technique. We examined whether clinical experience influenced measurement of TPMA when utilising these two commonly used methods and describe the relationship between techniques.
Methods: Pre-operative cross-sectional CT imaging of 114 consecutive patients undergoing elective endovascular aneurysm repair were analysed. Retrospective measurements of the TPMA were performed by four independent investigators with a range of clinical experience (medical student to specialist surgical registrar) using either technique. Intra- and inter-observer differences were assessed.
Results: There was no significant intra- or inter-observer differences when measuring the TPMA. Clinical experience also did not influence TPMA measurements recorded. Significant differences were observed between techniques when measuring TPMA (mean −65.8, 239.3 SD, p = 0.004). Measurement differences between techniques were highly correlated and modelled using linear regression.
Conclusion: Both manual tracing and semi-automated technique quantification methods of measuring TPMA are highly reproducible and independent of assessor bias and clinical experience.
Advances in knowledge: Either of the commonly used techniques to measure TPMA may be reliably used by an individual with appropriate training. We describe a relationship to facilitate comparison between these methods by which sarcopaenia may be quantified in patients with routine CT imaging.

INTRODUCTION

Sarcopaenia, defined by the loss of muscle mass and strength, is estimated to affect around 9–18% of individuals over the age of 65 years old.[1] To date several studies have demonstrated adverse outcomes following abdominal surgery to be associated with sarcopaenia indicating a potential role in patient risk stratification.[2–7] The two dominant methods used to radiologically quantify sarcopaenia include the measurement of the total psoas muscle area (TPMA) or total abdominal muscle area, which typically includes the: psoas, erector spinae, quadratus lumborum, transversus abdominis, external and internal oblique's, and rectus abdominis muscles. Measurements are universally made from single slice imaging at the level of the third lumbar vertebrae or umbilicus.[2–7]

Manual tracing and the semi-automated technique are two commonly used methods to measure TPMA.[2–7] The manual technique involves the observer to draw freehand around the psoas to calculate its axial area. Alternatively, the semi-automated technique is based on Hounsfield unit (HU) thresholds that can be utilised to highlight skeletal muscle automatically.[2,3,7] Both these methods described require a limited knowledge of anatomy to identify the psoas muscle and delineate it from the quadratus lumborum and crura.

Despite the reported use of both image analysis methods extending over 30 years, there is limited evidence to suggest whether observer clinical experience has any influence on measurement accuracy and reproducibility when utilising either method. It is on this background we aimed to investigate whether clinical experience influences observer measurement of TPMA. Secondly, we investigated whether the TPMA measurements were influenced by the measurement technique used.

METHODS AND MATERIALS

We analysed anonymised pre-operative cross-sectional abdominal CT scans in patients who had undergone elective endovascular aneurysm repair for an abdominal aortic aneurysm. Pre-operative CT imaging of the abdomen is routinely performed on all patients undergoing planned elective intervention, in the supine position with a breath-hold to minimise motion artefact. Consecutive patients were identified from a prospectively maintained database, the Health Quality Improvement Partnership National Vascular Registry), from January 2008 to December 2014.[8] Inclusion into the study required the patient to have a cross-sectional CT image available at the level of the third lumbar vertebrae on the local picture archiving and communications system.[9] Patients were excluded if the procedure was an open surgical repair, acute endovascular aneurysm repair, regional cases (where images were not available) or if they had incomplete imaging.

This study followed the Strengthening the Reporting of Observational Studies in epidemiology guideline for reported retrospective cohort studies. Ethical approval was granted by the local radiology research authorisation group and Health Research Authority (IRAS project ID, 228484).

Observer selection

Four independent observers were selected based solely on clinical experience from a selection of employees and students working within the academic department of the Leeds Vascular Institute at the Leeds Teaching Hospitals NHS Trust. These included two doctors with 2 years postgraduate clinical experience (Observers 1 and 2), surgical registrar with 5 years postgraduate clinical experience (Observer 3) and a second-year medical student (Observer 4). Observers 1 and 2 were selected so that differences in measurements between methods could be assessed independent of bias of clinical experience. Observer 3 was selected to ensure anatomical familiarity. Observer 4 was selected as they were in the early years of training to ensure knowledge relating to anatomy was limited.

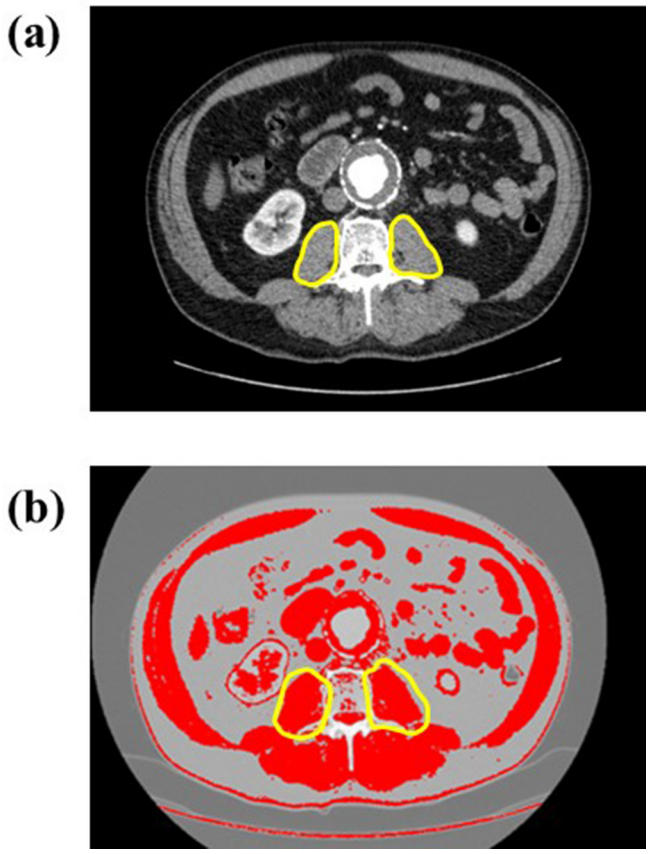
Data collection

Demographic data collected included age, gender, height and weight. All images were assessed for inclusion by a single investigator. Transverse cross-sectional images at the level of the transverse processes were identified by counting up from the sacrum to the third lumbar vertebrae on the sagittal view of the abdomen. Images were subsequently downloaded in the digital imaging and communications in medicine format with preservation of actual dimensions to avoid magnification indices and scales.

Image analysis

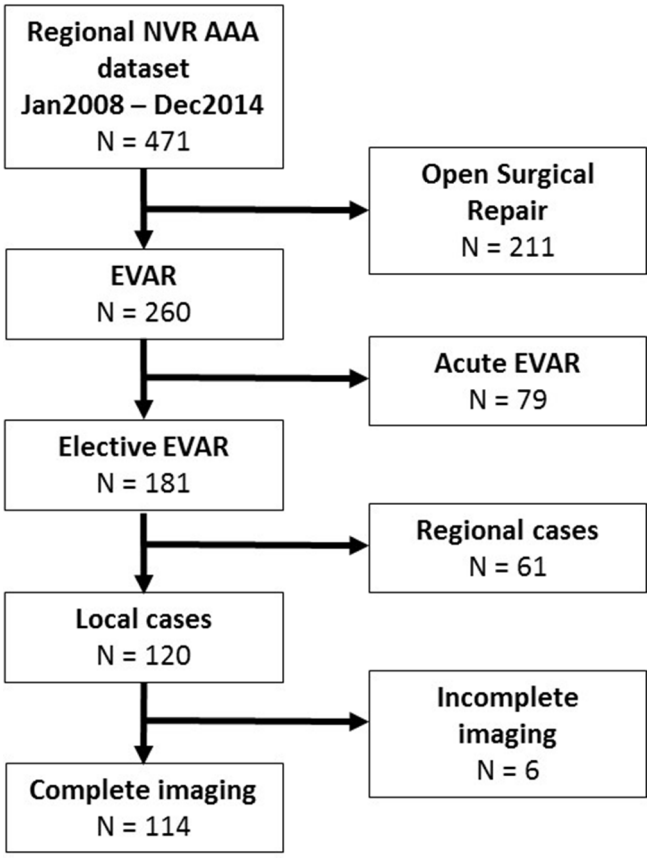
Individual observers were trained by an experienced investigator (consultant radiologist for the manual tracing method and surgical registrar for the semi-automated technique method). Manual tracing was performed using a free version of OsiriX 6.0 (Pixmeo SARL, Bernex, Switzerland) by identifying the psoas muscle and drawing around the presumed border (Figure 1a).[10] The semi-automated technique was performed using ImageJ (National Institutes of Health, Bethesda, MD) by circling around a region within which the psoas muscle was evident after setting the HU range between −30 and 130 (Figure 1b).[11] ImageJ is a free open platform image processing software developed by the National Institutes of Health (https://imagej.nih.gov/ij/) which may be used to process digital imaging and communications in medicine images. We acknowledge that there are a variety of other software packages which may have been used for this purpose.

Figure 1. Example of transverse cross-sectional digital image of the same patient assessed using (a) the manual training method and (b) the semi-automated technique.



Images were analysed by measuring both the left psoas muscle area (LPMA) and right psoas muscle area (RPMA), and subsequently deriving the TPMA. Observers were trained to identify and exclude the quadratus lumborum and crura. All images were individually analysed using either the manual tracing or semi-automated technique method by Observers 1 and 2 respectively. 50 images were subsequently randomly selected using a random number generator (https://www.random.org). The two initial observers (Observers 1 and 2) reanalysed the images blinded to the results obtained from their initial analyses so that any intra-observer differences could be identified. We felt it to be unlikely that observers would be able to recall the measurements of all 100 measurements (50 LPMA and 50 RPMA) they had previously recorded. The images were then analysed with manual tracing by Observer 3 and then using semi-automated technique by Observer 4. This facilitated the assessment of inter-observer

Figure 2. Flow chart illustrating the inclusion and exclusion of local patients identified from the NVR. AAA, abdominal aortic aneurysm; EVAR, endovascular aneurysm repair; NVR, National Vascular Registry.



differences in measurements based on clinical experience. All observers were blinded to the results of other observers.

Statistical analysis

All measurements were calculated as an area (square millimetre). Continuous variables were reported as a mean alongside a 95% confidence interval (95% CI) or standard deviation (SD). Intra-observer and inter-observer differences were evaluated for measurements of TPMA using Bland–Altman plots and differences in measurements evaluated using Student's *t*-test. The limits of agreement were calculated from the SD of differences calculated between observer measurements. Prior to comparing both methods all measurements of TPMA obtained for each technique were averaged to minimise observer error. The relationship between the two assessment modalities was similarly assessed using a Bland–Altman plot and trends validated using Pearson's correlation and linear regression. Statistical significance was defined as $p < 0.05$. Statistical analyses were performed using Minitab 17 (Minitab Inc, Coventry, Pennsylvania).[12]

RESULTS

Overall 471 patients were identified initially from the National Vascular Registry. A total of 114 patients (24%) were identified to be eligible for measurement of the TPMA (Figure 2). Corresponding CT imaging were performed between January 2008 and December 2014. In this cohort 97 patients were males (85.1%) and the mean age was 76 years [95% CI (74.8–77.6)]. The mean height was 171 cm [95% CI (169.3–174.1)] and mean weight 81.3 kg [95% CI (77.0–85.5)].

Methods of image analysis

The mean TPMA measured by Observer 1 using manual tracing was 2055.6 mm^2 (SD 537.2). Repeat assessment by Observer 1 did not highlight any significant intra-observer differences in measurements (mean difference −8.9 mm^2, SD 94.4, *p*-value = 0.508) (Figure 3a). Mean TPMA measurements by Observer 3 using manual tracing were 2070.1 mm^2 (SD 567.6). No significant inter-observer differences were evident when using manual tracing (mean difference −14.5 mm^2, SD 120.1, *p*-value 0.399) (Figure 3b). When utilising the semi-automated technique mean TPMA measurements by Observer 2 were 2137.4 mm^2 (SD 569.8). Again repeat assessment by Observer 2 did not highlight any significant intra-observer differences (mean difference −6.6 mm^2, SD 65.0, *p*-value = 0.479) (Figure 3c). Mean measurement of TPMA using semi-automated technique by Observer 4 was 2142.1 mm^2 (SD 577.6). No significant inter-observer differences were observed when using the semi-automated technique (mean difference −4.7, SD 60.3, *p*-value = 0.585) (Figure 3d).

Measurement of the TPMA was significantly lower with manual tracing (mean difference −65.8, SD 239.3 mm^2, $p < 0.0005$) (Table 1). This may be due to the measurement of the LPMA being significantly lower with manual tracing compared to semi-automated assessment (mean difference 46.5 mm^2, SD 133.9 mm^2, $p < 0.0005$). Furthermore comparative measurements of the RPMA with manual tracing had a slight negative skew in relation to measurements with the semi-automated technique (skewness −0.07) implying a trend towards lower measurements of the RPMA with MT. Despite these significant differences in measurements between both techniques, measurements were significantly correlated on Pearson's correlation which could be modelled using linear regression (Table 1 and Figure 4).

Intra-observer and inter-observer differences

The range of clinical experience of observers is shown in Figure 5. Comparison of differences in measurements by Observers 1 and 3 were not significant as described above. Similarly, there were no significant differences in measurements by Observers 2 and 4. Comparison of corrected semi-automated technique measurements by Observer 4, using the linear regression equation calculated (Figure 4), there were significant differences between measurements by Observers 3 and 4 (mean difference −2 mm^2, SD 199.7, *p*-value = 0.985). Clinical experience did not appear to make any significant difference when measuring the TPMA using either of the two dominant techniques.

DISCUSSION

This study, to our knowledge, is the first to explore the impact of observer clinical experience on the measurement of TPMA. Furthermore, we add to the existing evidence that both manual tracing and semi-automated technique are robust and results reproducible following basic training from an experienced investigator.

Figure 3. Bland–Altman plots to assess intra- and inter-observer differences when using manual tracing and semi-automated technique to measure TPMA. TPMA, total psoas muscle area.

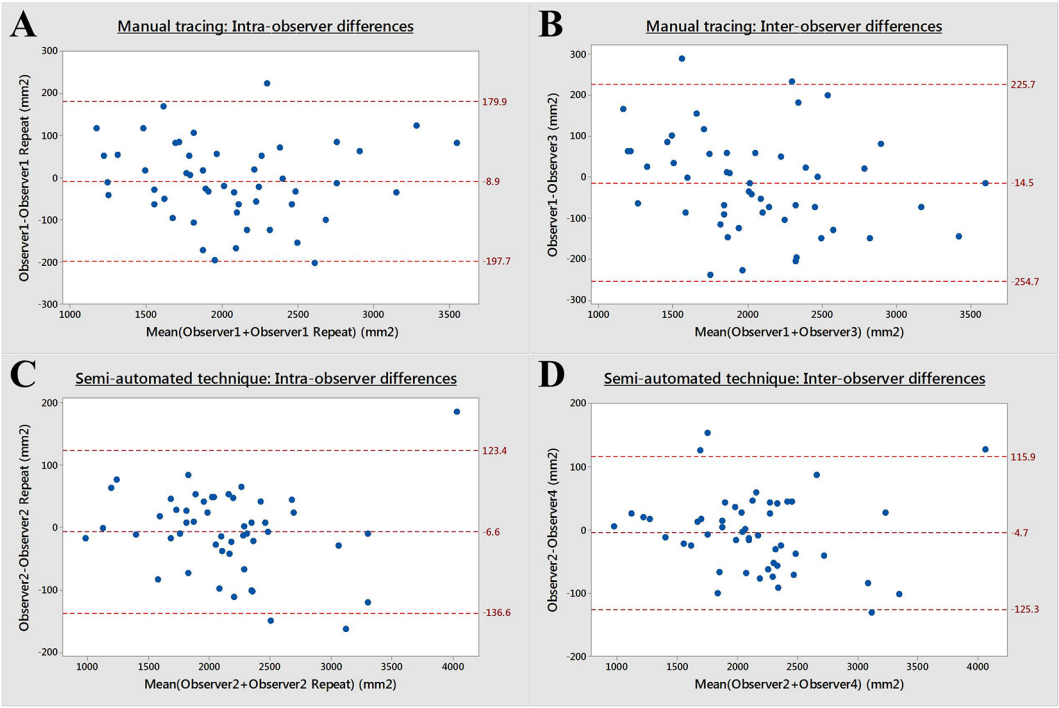


Table 1. Statistical comparison of psoas muscle measurements by observers comparing manual tracing and the semi-automated technique

Psoas muscle area	Manual tracing (mean, SD)	Semi-automated technique (mean, SD)	Inter-method difference (mean difference, SD)	*p*-value (*t*-test)	Pearson's correlation	Correlation *p*-value	*R*2 (adjusted)
LPMA (mm^2)	1009.6 (272.4)	1056.1 (258.1)	−46.5 (133.9)	<0.005	0.874	<0.0005	76.2%
RPMA (mm^2)	1015.2 (275.1)	1034.5 (271.4)	−19.3 (148.4)	0.168	0.853	<0.0005	72.5%
TPMA (mm^2)	2024.8 (527.1)	2090.6 (511.7)	−65.8 (239.3)	0.004	0.894	<0.0005	79.8%

LPMA, left psoas muscle area; RPMA, right psoas muscle area; TPMA, total psoas muscle area.

Risk stratification is a key component when planning for elective intervention or treatment in a multi-disciplinary setting.[13,14] Decisions are often based on patient co-morbidity, investigation results, patient preference and clinical suitability. The additional assessment of risk based on sarcopaenia would require the accurate and reliable measurement of the TPMA. Our data suggest

Figure 4. Bland–Altman plot and linear regression analysis assessing inter-method measurement differences in the measurement of TPMA with manual tracing and semi-automated technique. TPMA, total psoas muscle area.

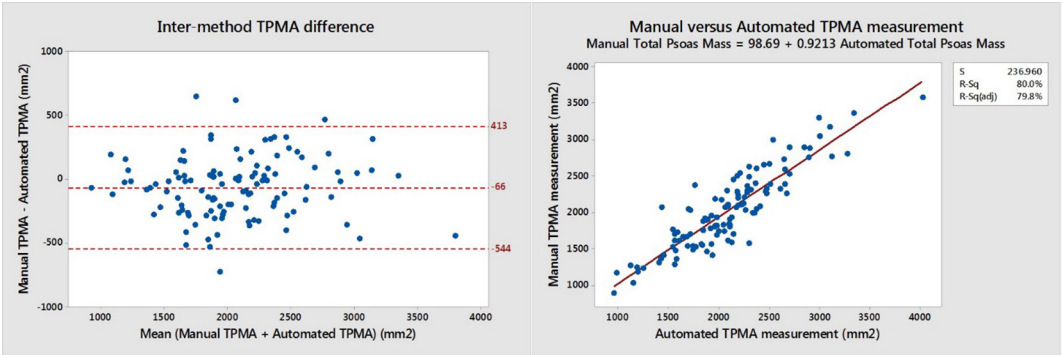
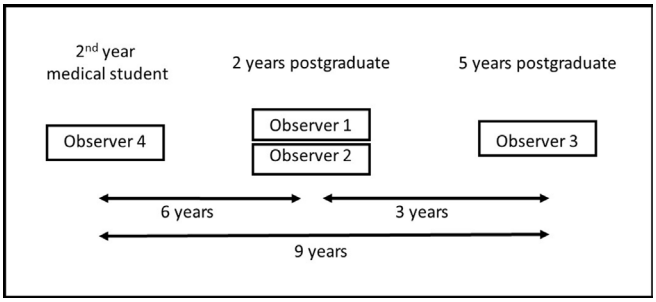


Figure 5. Illustration of differences in years of clinical experience between observers.



that such measurements could be accurately obtained by non-expert staff following a brief episode of training. Studies which have previously measured TPMA have either used a trained observer, as described in our study, however, they often failed to describe the clinical experience or background of the observer.[4] Jones et al described utilising a trained consultant radiologist to manual trace around the TPMA to assess for sarcopaenia without using HU to highlight regions.[15] It is not possible to delineate from the existing literature whether there is any advantage in using an experienced clinician, therefore our study helps address this important issue.

Our study design enables us to assess whether there was any additional advantage in measurements being carried out by individuals with greater clinical experience as: Observer 3 was 9 years senior to Observer 4 and 3 years senior to Observers 1 and 2. During undergraduate training medical students acquire increased clinical exposure in hospitals as they progress through the years. The interpretation of abdominal CT imaging is not routinely expected especially during early years of undergraduate learning or the first few years of postgraduate clinical training.[16,17] The knowledge and ability to perform simple interpretation of routine imaging is acquired through increased in-hospital exposure and clinical experience. We demonstrate additional postgraduate clinical experience did not result in significantly different results. Therefore, it may be possible that other members of the clinical team may be suitable to perform the measurements provided they are suitably trained. It may therefore be possible in the future to implement the sarcopaenia risk stratification tool within the multi-disciplinary setting without the need of specialist image interpretation.

Both dominant methods, manual tracing and semi-automated technique, require some understanding of the physical location of the psoas muscle within the cross-sectional view so that it may be identified and measured. We thought initially it may be easier to identify the psoas muscle using the HU-based semi-automated technique. However, it was evident that this was not the case from the data we present. Measurement of the LPMA and the derived TPMA was significantly lower with manual tracing. This may have been partly due to difficulties in identifying the boundaries of the psoas muscle resulting in the encroachment into the psoas muscle area when tracing around the boundaries. These inaccuracies were also evident even when the observer had greater clinical experience. This limitation was not demonstrated when utilising the semi-automated technique as the psoas muscle boundaries were often clearly visible and the observer had to only define the area of interest and not the exact boundaries of the psoas muscle.

It is also important to acknowledge that many muscle groups undergo fat atrophy in sarcopaenia, therefore using HUs to delineate density from size is useful. However the psoas muscle is unique as it retains a consistent muscle density and does not become enlarged with adipose tissue. Both these techniques also have limited utilisation in patients who have pathology that may have a direct effect on the psoas muscle size. These include patients who have had previous hip surgery, severe hip osteoarthritis or severe scoliosis. Artefact from prosthesis may also limit the utilisation of these techniques.

CONCLUSIONS

The measurement of TPMA is commonly used in clinical research as a method of risk stratification to quantify sarcopaenia. The data we present demonstrate there are no significant differences in measurements of TPMA in relation to clinician experience. The results from our study would facilitate the utilisation of a sarcopaenia-based risk stratification model derived with either method of measuring TPMA in the routine clinical practice.

ACKNOWLEDGEMENTS

We would like to thank Mr David Watson in help us retrieve data from the national vascular registry.

FUNDING

No external funding sources were required to conduct this work. MAW is an academic clinical fellow supported by the National Institute for Health Research. MAB is an intermediate clinical fellow supported by the British Heart Foundation.

REFERENCES

1. Patel HP, Syddall HE, Jameson K, Robinson S, Denison H, Roberts HC, et al. Prevalence of sarcopenia in community-dwelling older people in the UK using the European Working Group on Sarcopenia in Older People (EWGSOP) definition: findings from the Hertfordshire Cohort Study (HCS). *Age Ageing* 2013; **42**: 378–84. doi: https://doi.org/10.1093/ageing/afs197
2. Miyamoto Y, Baba Y, Sakamoto Y, Ohuchi M, Tokunaga R, Kurashige J, et al. Sarcopenia is a negative prognostic factor after curative resection of colorectal cancer. *Ann Surg Oncol* 2015; **22**: 2663–8. doi: https://doi.org/10.1245/s10434-014-4281-6
3. Montano-Loza AJ, Meza-Junco J, Baracos VE, Prado CM, Ma M, Meeberg G, et al. Severe muscle depletion predicts

postoperative length of stay but is not associated with survival after liver transplantation. *Liver Transpl* 2014; **20**: 640–8. doi: https://doi.org/10.1002/lt.23863
4. Peng P, Hyder O, Firoozmand A, Kneuertz P, Schulick RD, Huang D, et al. Impact of sarcopenia on outcomes following resection of pancreatic adenocarcinoma. *J Gastrointest Surg* 2012; **16**: 1478–86. doi: https://doi.org/10.1007/s11605-012-1923-5
5. Peyton CC, Heavner MG, Rague JT, Krane LS, Hemal AK. Does sarcopenia impact complications and overall survival in patients undergoing radical nephrectomy for stage III and IV kidney cancer? *J Endourol* 2016; **30**: 229–36. doi: https://doi.org/10.1089/end.2015.0492
6. Smith AB, Deal AM, Yu H, Boyd B, Matthews J, Wallen EM, et al. Sarcopenia as a predictor of complications and survival following radical cystectomy. *J Urol* 2014; **191**: 1714–20. doi: https://doi.org/10.1016/j.juro.2013.12.047
7. Tamandl D, Paireder M, Asari R, Baltzer PA, Schoppmann SF, Ba-Ssalamah A. Markers of sarcopenia quantified by computed tomography predict adverse long-term outcome in patients with resected oesophageal or gastro-oesophageal junction cancer. *Eur Radiol* 2016; **26**: 1359–67. doi: https://doi.org/10.1007/s00330-015-3963-1
8. Waton S, Johal A, Heikkilä K, Cromwell D, Loftus I. *National vascular registry: annual report*. London, UK: HQIP; 2015.
9. Cooke Jr RE, Gaeta MG, Kaufman DM, Henrici JG. Picture archiving and communication system. *Google Patents* 2003.
10. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* 2004; **17**: 205–16. doi: https://doi.org/10.1007/s10278-004-1014-6
11. Abr moff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. *Biophotonics international* 2004; **11**: 36–42.
12. Minitab I. *MINITAB release 17: statistical software for windows*. USA: Minitab Inc; 2014.
13. Stephens MR, Lewis WG, Brewster AE, Lord I, Blackshaw GR, Hodzovic I, et al. Multidisciplinary team management is associated with improved outcomes after surgery for esophageal cancer. *Dis Esophagus* 2006; **19**: 164–71. doi: https://doi.org/10.1111/j.1442-2050.2006.00559.x
14. Cima R, Dankbar E, Lovely J, Pendlimari R, Aronhalt K, Nehring S, et al. Colorectal surgery surgical site infection reduction program: a national surgical quality improvement program-driven multidisciplinary single-institution experience. *J Am Coll Surg* 2013; **216**: 23–33. doi: https://doi.org/10.1016/j.jamcollsurg.2012.09.009
15. Jones KI, Doleman B, Scott S, Lund JN, Williams JP. Simple psoas cross-sectional area measurement is a quick and easy method to assess sarcopenia and predicts major surgical complications. *Colorectal Dis* 2015; **17**: O20–O26. doi: https://doi.org/10.1111/codi.12805
16. Committee GMCE. *Tomorrow's doctors: recommendations on undergraduate medical education: General Medical Council London*; 1993.
17. Miles KA. Diagnostic imaging in undergraduate medical education: an expanding role. *Clin Radiol* 2005; **60**: 742–5. doi: https://doi.org/10.1016/j.crad.2005.02.011