



HAL
open science

A comparative study of two automated solutions for cross-sectional skeletal muscle measurement from abdominal computed tomography images

Katia Charrière, Quentin Boulouard, Svetlana Artemova, Antoine Vilotitch, Gilbert R Ferretti, Jean-luc Bosson, Alexandre Moreau-gaudry, Joris Gai, Eric Fontaine, Cécile Bétry

► To cite this version:

Katia Charrière, Quentin Boulouard, Svetlana Artemova, Antoine Vilotitch, Gilbert R Ferretti, et al.. A comparative study of two automated solutions for cross-sectional skeletal muscle measurement from abdominal computed tomography images. *Medical Physics*, 2023, 50 (8), pp.4973 - 4980. 10.1002/mp.16261 . hal-04303471

HAL Id: hal-04303471

<https://hal.science/hal-04303471>

Submitted on 23 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A comparative study of two automated solutions for cross-sectional skeletal muscle measurement from abdominal computed tomography images

Katia Charrière¹ | Quentin Boulouard¹ | Svetlana Artemova¹ | Antoine Vilotitch² | Gilbert R. Ferretti³ | Jean-Luc Bosson^{1,4} | Alexandre Moreau-Gaudry^{1,4} | Joris Gai^{1,4} | Eric Fontaine⁵ | Cécile Bétry⁶

¹Public Health Department, Clinical Investigation Center-Technological, Innovation, INSERM CIC1406, CHU Grenoble Alpes, Université Grenoble Alpes, Grenoble, France

²CHU Grenoble Alpes, Cellule d'ingénierie des données, Grenoble, France

³INSERM U1209, IAB, CHU Grenoble Alpes, Service de radiologie diagnostique et interventionnelle, Université Grenoble Alpes, Grenoble, France

⁴CNRS, UMR 5525, VetAgro Sup, Grenoble INP, CHU Grenoble Alpes, Public Health Department, TIMC, Université Grenoble Alpes, Grenoble, France

⁵Department of Endocrinology, Diabetology and Nutrition, INSERM U1055, LBFA, CHU Grenoble Alpes, Université Grenoble Alpes, Grenoble, France

⁶Department of Endocrinology, Diabetology and Nutrition, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, CHU Grenoble Alpes, TIMC, Université Grenoble Alpes, Grenoble, France

Correspondence

Cécile Bétry, CHU Grenoble Alpes (CHUGA), Service d'Endocrinologie Diabétologie Nutrition, Boulevard de la Chantourne, Université Grenoble Alpes, 38700 La Tronche, France.
Email: cbetry@chu-grenoble.fr

Funding information

Délégation Régionale à la Recherche Clinique du Centre Hospitalier Universitaire Grenoble Alpes

Abstract

Background: Measurement of cross-sectional muscle area (CSMA) at the mid third lumbar vertebra (L3) level from computed tomography (CT) images is becoming one of the reference methods for sarcopenia diagnosis. However, manual skeletal muscle segmentation is tedious and is thus restricted to research. Automated solutions are required for use in clinical practice.

Purpose: The aim of this study was to compare the reliability of two automated solutions for the measurement of CSMA.

Methods: We conducted a retrospective analysis of CT images in our hospital database. We included consecutive individuals hospitalized at the Grenoble University Hospital in France between January and May 2018 with abdominal CT images and sagittal reconstruction. We used two types of software to automatically segment skeletal muscle: ABACS, a module of the SliceOmatic software solution "ABACS-SliceOmatic," and a deep learning-based solution called "AutoMATICA." Manual segmentation was performed by a medical expert to generate reference data using "SliceOmatic." The Dice similarity coefficient (DSC) was used to measure overlap between the results of the manual and the automated segmentations. The DSC value for each method was compared with the Mann–Whitney U test.

Results: A total of 676 hospitalized individuals was retrospectively included (365 males [53.8%] and 312 females [46.2%]). The median DSC for SliceOmatic vs AutoMATICA (0.969 [5th percentile: 0.909]) was greater than the median DSC for SliceOmatic vs. ABACS-SliceOmatic (0.949 [5th percentile: 0.836]) ($p < 0.001$).

Conclusions: AutoMATICA, which used artificial intelligence, was more reliable than ABACS-SliceOmatic for skeletal muscle segmentation at the L3 level in a cohort of hospitalized individuals. The next step is to develop and validate a neural network that can identify L3 slices, which is currently a fastidious process.

KEYWORDS

artificial intelligence, body composition, computational neural networks, sarcopenia, skeletal muscle, software validation

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Medical Physics* published by Wiley Periodicals LLC on behalf of American Association of Physicists in Medicine.

1 | INTRODUCTION

Sarcopenia is defined as a combination of reduced muscle mass and function according to the European Working Group on Sarcopenia in Older People.¹ It is a frequent comorbidity: at least 2 out of 10 hospitalized adults of all ages have sarcopenia.² A recent meta-analysis found that sarcopenia increases the risk of mortality, independently from sarcopenia definition and population.³ It is also associated with poor clinical outcomes in several clinical situations, for example, sepsis, cancer, vascular surgery.^{4–7}

An important issue relating to sarcopenia diagnosis is the lack of a gold standard for use in clinical practice.⁸ For over a decade, a growing number of studies have examined the measurement of cross sectional muscle area (CSMA) at the mid third lumbar vertebra (L3) level using computed tomography (CT).⁹ Manual skeletal muscle segmentation from CT images remains tedious and is thus restricted to research. Automated segmentation could extend the use of CT for sarcopenia diagnosis to the clinical setting.

ABACS (Automatic Body composition Analyzer using Computed tomography image Segmentation software) is a module of SliceOmatic (ABACS-SliceOmatic).⁹ It contains an automated muscle segmentation solution, which is based on the a priori shape of skeletal muscle.^{10,11} Interest in the use of deep-learning based methods for muscle segmentation is also growing.^{12–14} Paris et al. developed an artificial intelligence (AI) solution termed Automated Muscle and Adipose Tissue Composition Analysis (AutoMATiCA), which they made freely available.¹⁴ The reliability of ABACS-SliceOmatic is controversial^{15,16} and AutoMATiCA has not been externally validated. Few researchers have addressed the question of the validity of automated segmentation tools. The present study aimed to independently compare the reliability of ABACS-SliceOmatic and AutoMATiCA in a retrospective cohort of individuals hospitalized in a French University Hospital.

2 | MATERIAL AND METHODS

This study is a part of a larger study: the Optimization of the DIAGnosis of SarcoPenia through the automated determination of SMI (ODIASP) study

2.1 | Ethics statement

The ODIASP study was performed in accordance with the recommendations of the Helsinki Declaration. Ethical approval was obtained on 18 August 2021 (CECIC Rhône-Alpes-Auvergne, Clermont-Ferrand, IRB 5891).

According to current French law for retrospective studies of clinical data, participants were individually informed that their data could be used for research (MR-004 CNIL reference methodology). People who objected and those with missing updated contact details were excluded. All the data was retrospectively collected in the electronic health records and de-identified.

2.2 | Study cohort

All consecutive individuals (1) admitted to Grenoble University Hospital (CHU Grenoble Alpes) between January and May 2018, (2) with abdominal CT images that included a complete image of the third lumbar vertebra and sagittal reconstruction, and (3) with their height recorded in the electronic health records, were retrospectively included.

Exclusion criteria were (1) people under 19 years, (2) creatinine level above the upper limit, and (3) images of insufficient quality for manual skeletal muscle segmentation, for example, incomplete slices with muscle area missing, metal artefacts, psoas hematoma, or motion. Data was initially collected for another research project assessing the relationship between low level of serum creatinine and sarcopenia. For this previous project, to limit inclusion of subjects with advanced kidney disease, those with serum creatinine level beyond the upper limit (84 and 104 $\mu\text{mol/L}$ for female and male respectively) were excluded. Similarly, subjects under 19 years were excluded, since the normal range of serum creatinine level is different for these subjects.

Since contrast enhancement does not influence cross-sectional skeletal muscle area, both CT images with and without contrast administration were included.^{17,18} CT images from four different CT machines (Optima CT660 GE Healthcare, Revolution CT GE Healthcare, Revolution HD GE Healthcare, and Siemens Somatom Definition Edge) with different phases of contrast administration, tube current intensities, and tube voltages were included in this study. The convolution kernel was standard ($n = 645$), standard2 ($n = 11$), detail2 ($n = 1$), or bone ($n = 1$) for GE Healthcare and I30f for Siemens ($n = 18$).

2.3 | Data collection

Descriptive data including size, weight, sex, and age were retrospectively collected through a clinical data warehouse called PREDIMED (acronym in French for: Plateforme de Recueil et d'Exploitation des Données bloMEDicales), deployed at the Grenoble Alpes University Hospital.¹⁹ Some missing data was completed manually from electronic health records.

2.4 | Skeletal muscle segmentation

Single CT slices of the third lumbar vertebrae were identified from the Picture Archiving and Communication System (PACS) in Digital Imaging and Communication in Medicine (DICOM) format. We identified the third lumbar vertebrae from the sagittal reconstruction. CT images were collected and de-identified with PREDIMED. They were de-identified by removing all identifying DICOM tags and dose reports. All processing was performed on a single GPU (NVIDIA TITAN RTX) machine with an Intel(R) Xeon(R) W-2135 processor.

2.4.1 | Manual segmentation with sliceOmatic

Skeletal muscle was manually segmented by a trained observer with medical experience (Quentin Boulouard - Q.B.) using SliceOmatic Version 5.0 (Tomovision, Canada) as previously described.⁵ The abdominal muscles (transversus abdominus, external and internal obliques, and rectus abdominus), the paraspinal muscles (erector spinae and quadratus lumborum) and the psoas muscle were segmented using the Hounsfield unit (HU) threshold from -29 to 150 .⁵ To determine the ground truth, each of the 676 segmentations was double-checked by a medical supervising observer (Cécile Bétry - C.B.). This second observer looked at each of the segmented images. She improved the segmentation whenever it was needed. Furthermore, to determine inter-observer reliability, a random sample of 50 CT-scan slices was independently segmented by both observers (Q.B. and C.B.).

2.4.2 | Automated methods: ABACS-SliceOmatic and AutoMATiCA

Skeletal muscle was automatically segmented with (1) ABACS (Voronoi Health Analytics Inc.), a module of the SliceOmatic software solution (Tomovision, version 5.0) and (2) the AutoMATiCA software solution.

ABACS-SliceOmatic is based on a priori information about muscle shape at the third lumbar vertebra level. More precisely, it is based on a finite element method deformable model that incorporates an a priori shaped model within the template-based segmentation framework. It was developed to avoid mislabeling parts of the neighboring organs that have an HU density within the range of muscle tissue HU density.^{10,11}

AutoMATiCA is based on a deep learning algorithm and, more precisely, on convolutional Neural Networks that have proven their effectiveness in the field of medical imaging.²⁰ Paris et al. developed AutoMATiCA as an

open source software solution: it is available at GitLab (https://gitlab.com/Michael_Paris/AutoMATiCA).¹⁴

2.5 | Statistical analysis

Descriptive data is presented as median and interquartile range (IQR). The Dice similarity coefficient (DSC) was used to measure the overlap between segmentation with the manual reference method and segmentation with each of the automated solutions. A value of 1 indicates perfect agreement, whereas 0 indicates no overlap.²¹ Median values and the 5th percentile are presented for the DSC. The median DSC values for each solution were compared with the Mann–Whitney U test. The DSC value for the manual segmentation of 50 slices by both observers was also calculated. Inter-observer agreement was determined with the intra-class correlation coefficients two-way random effect model (ICCs) and 95% confidence intervals. A value < 0.5 indicates poor reliability, a value between 0.5 and 0.75 indicates moderate reliability, a value between 0.75 and 0.9 indicates substantial reliability, and a value > 0.90 indicates excellent reliability.²² To determine the impact of body mass index (BMI), current intensity, voltage and intravenous (IV) contrast administration on the reliability of AutoMATiCA and ABACS-SliceOmatic, mixed-model linear regressions were performed on each variable. The potentially impacting variable was used as a continuous factor or a categorical variable, the two solutions as fixed factors, and the individual as a random factor. The DSC was transformed to allow modeling. The impact of voltage was assessed after classification into 3 classes (100/120/140 mV). Statistical analyses were performed with Stata (version 15). p -values < 0.05 were considered significant.

3 | RESULTS

3.1 | Participant inclusion and characteristics

All participants ($n = 1071$) with abdominal CT images, including sagittal reconstruction, performed between January and May 2018 and with creatinine measurement were screened. Ninety-nine individuals who objected or with missing updated contact details were excluded. After exclusion of those with missing height information ($n = 128$) and poor-quality images ($n = 168$), a total of 676 individuals were included. The study cohort consisted of 365 males (53.8%) and 312 females (46.2%), their characteristics are shown in Table 1. Participants were hospitalized in different units of the hospital, that is, medical units, surgical units, intensive care units, or emergency units.

TABLE 1 Characteristics of the cohort and CT acquisition parameters.

	Median (IQR) or <i>n</i> (%)
Age (years)	61 (44–73)
BMI (kg/m ²)	24.7 (21.6–27.8)
Underweight (%)	50 (7.5%)
Normal (%)	312 (46.6%)
Overweight	204 (30.5%)
Obese	206 (15.4%)
CSMA (cm ²) with SliceOmatic	123 (102–151)
Muscle density (HU)	40.1 (31.9–46.7)
SMI (cm ² /m ²) with SliceOmatic	43.9 (37.8–51.4)
Voltage (kV)	
100	16 (2.4%)
120	657 (97.2%)
140	3 (0.4%)
Current (mA)	155 (150–247)
IV contrast, %	539 (79.7%)

Abbreviations: BMI, body mass index; CSMA, cross sectional muscle area; SMI, skeletal muscle index.

BMI was available for 669 participants.

TABLE 2 Dice similarity coefficient (DSC) values for the paired comparisons of automated and manual segmentation.

	SliceOmatic vs. ABACS-SliceOmatic	SliceOmatic vs. AutoMATiCA
5th percentile	0.836	0.909
Median	0.949	0.969
IQR	(0.924–0.967)	(0.954–0.980)
Number of DSC <0.90	100 (14.8 %)	23 (3.4 %)

3.2 | Inter-observer agreement

Median CSMA was 119 (106–146) and 116 (104–142) cm² for the trained observer (QB) and the supervising observer (CB) respectively. The agreement between both observers was excellent with an ICC of 0.982 (0.967–0.990). Median DSC value was 0.960 (5th percentile: 0.901) for the random sample of 50 CT slices.

3.3 | Comparison of ABACS-SliceOmatic and AutoMATiCA

Median CSMA were 123 (102–149) and 128 (108–158) for ABACS-SliceOmatic and AutoMATiCA respectively. The descriptive values of the DSC for the paired comparisons between manual segmentation and automated segmentation are shown in Table 2 and Figure 1. AutoMATiCA was more reliable than ABACS-SliceOmatic ($p < 0.001$). Nevertheless, in 50 cases (7.4 %), the

DSC value for SliceOmatic versus AutoMATiCA was lower than the value for SliceOmatic versus ABACS-SliceOmatic. Figure 2 reports some erroneous segmentations with ABACS-SliceOmatic, for example, in individuals with sarcopenia (A, C) or obesity (B, C), and in the case of artifacts such as photon starvation (B). The manual segmentation of 50 slices took 29 min for the experimented observer whereas segmentation of the same images took 19 min with ABACS-SliceOmatic and 33 s with AutoMATiCA.

3.4 | Factors affecting the accuracy of the muscle segmentation by ABACS-SliceOmatic and AutoMATiCA

Among the 23 cases with a DSC (SliceOmatic vs AutoMATiCA) value < 0.9, only two had a value (SliceOmatic vs ABACS-SliceOmatic) > 0.9 (Figure S1). Seven of the twenty-three cases (30%) occurred in individuals with a BMI <18.5 kg/m². The results of the mixed-model linear regression analysis for the impact of BMI on the DSC are displayed in Table 3. The BMI and the solution used had a significant effect on the DSC. The interaction in the multivariate analyses was significant, suggesting a different effect of BMI depending on the solution used, with a greater effect of BMI with AutoMATiCA. As shown in Figure S2, the lower the BMI, the lower the DSC. Similarly, the lower the current intensity, the lower the DSC, with no interaction with the solution used. Conversely, voltage and IV contrast did not have a significant effect on the accuracy of muscle segmentation (Table 3).

4 | DISCUSSION

The results of this study showed that AutoMATiCA provided a more reliable automatic segmentation of the muscles at the level the L3 vertebra from CT images than ABACS-SliceOmatic.

4.1 | Reliability of the AutoMATiCA AI-automated software

As far as we know, our study is the first to compare an AI-automated software solution to ABACS-SliceOmatic in an unselected cohort. The DSC values for the inter-observer comparison and between manual and automated segmentation with AutoMATiCA were in the same order of magnitude. Moreover, the DSC value for the external validation (0.969) was also close to the value for the internal validation (0.983) suggesting good generalizability of AutoMATiCA.^{17,23} AutoMATiCA is faster than both manual segmentation and ABACS-SliceOmatic (33 s vs. 29 and 19 min respectively for 50 slices), which is an important feature for clinical

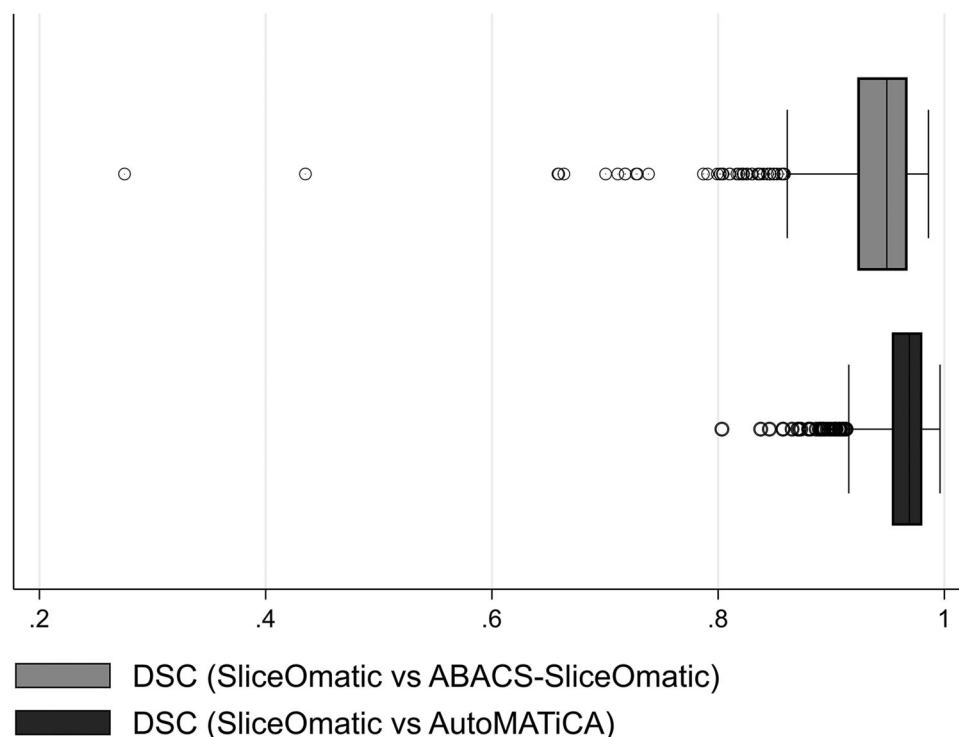


FIGURE 1 Boxplots for dice similarity coefficient (DSC) values as a measure of the overlap between the manual segmentations with SliceOmatic (the ground truth) and the automated segmentations (ABACS-SliceOmatic and AutoMATiCA). A DSC of 1 indicates perfect agreement, whereas 0 indicates no overlap.

practice. These findings support the use of machine learning for automated skeletal muscle segmentation in clinical practice.

4.2 | Limits of ABACS-SliceOmatic in clinical practice

The reliability of ABACS-SliceOmatic has already been assessed.^{15,16} ABACS-SliceOmatic is based on the a priori shape of skeletal muscle.¹⁰ As illustrated in Figure 2, anatomic differences, poor-quality images caused by excessive adipose tissue, severe sarcopenia or artifacts such as photon starvation may lead to erroneous segmentation, with some very low DSC values. Only 3.4% of the DSC were <0.90 for AutoMATiCA (however they remained >0.8): a much higher proportion (14.8%) were <0.90 with ABACS-SliceOmatic, with some are very low DSC values. Thus, segmentation with ABACS-SliceOmatic was erroneous in an unacceptable proportion of cases. Our team previously demonstrated a 13% rate of false negative sarcopenia diagnosis with ABACS-SliceOmatic.¹⁵ In contrast, Cespedes Feliciano et al. concluded that the reliability of automated segmentation with ABACS-SliceOmatic was similar to that of manual segmentation.¹⁶ One explanation for this discrepancy is the difference in quality of the two CT-scan databases. We did not exclude individuals with photon

starvation artefacts relating to the body pressing against the sides of the scanner, whereas Cespedes Feliciano et al. did. Moreover, cases with anatomic abnormalities were estimated at <2% in the study by Cespedes Feliciano et al. We believe that the proportion was much higher in our study. Finally, some of the CT-scans in our study were carried out in emergency situations, which often result in lower image quality.

4.3 | Study limitations

Several potential limitations need to be considered. First, there is no gold standard available in clinical practice for the assessment of muscle mass, even if CT is considered a reference standard.⁸ In this study, we compared automated segmentation to manual segmentation. We demonstrated that the DSC for two observers was high, which suggests that the manual method was of good quality and reliable. Furthermore, we double-checked the segmentation of the 676 CT slices to improve the quality of “the ground truth.”²³ Alternatively, a phantom study could have been performed. Second, we collected data and CT images retrospectively. Nevertheless, we included all consecutive individuals with abdominal CT images and sagittal reconstruction during a 5-month period. Thus, we limited selection bias and ensured that the database was representative of hospitalized

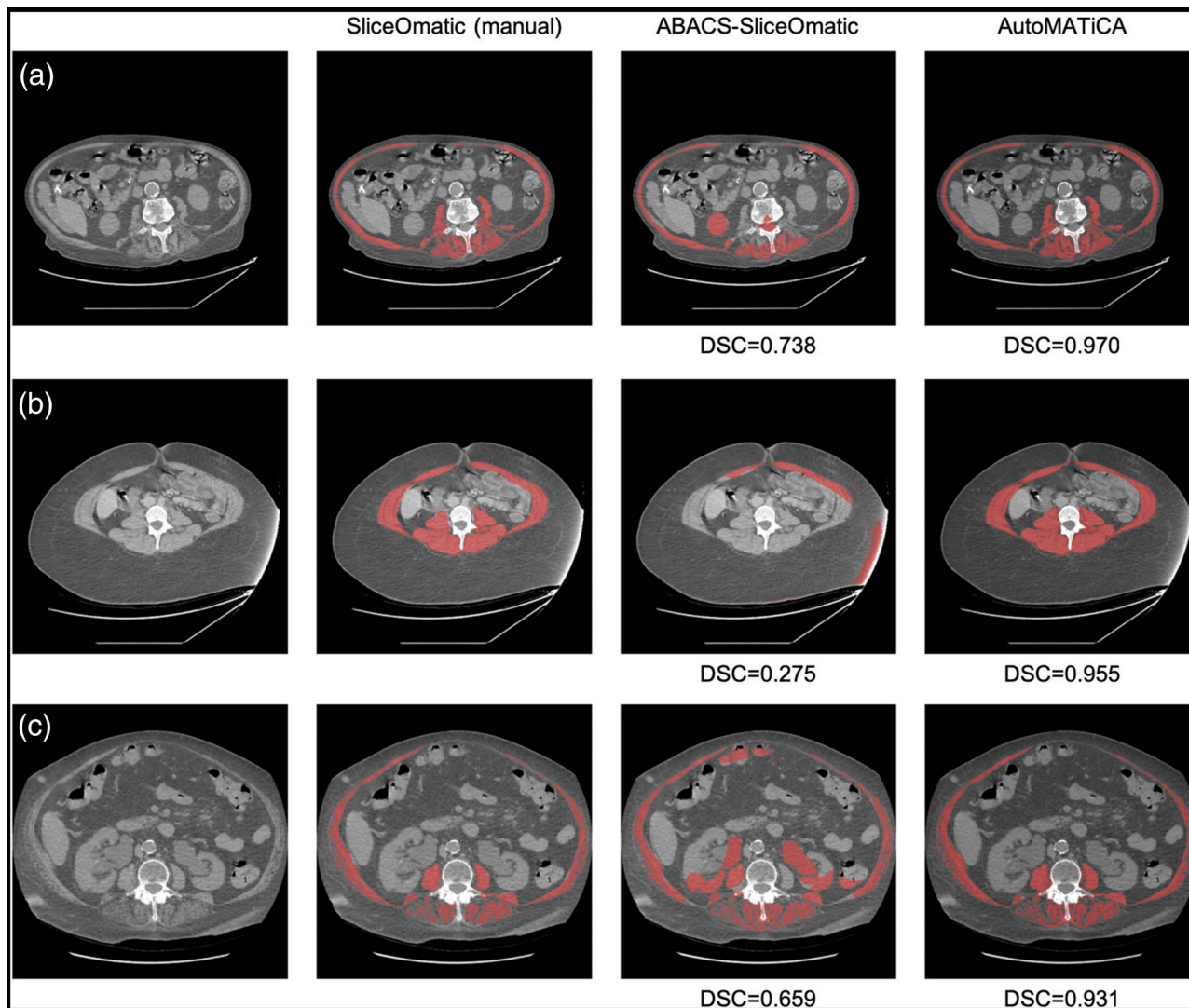


FIGURE 2 Three examples of erroneous automated segmentation with ABACS-SliceOmatic as compared with manual segmentation with SliceOmatic and automated segmentation with AutoMATiCA. The Dice similarity coefficient (DSC) value for the comparison of manual segmentation and each automated segmentation are provided.

individuals. Given the retrospective design, the quality of some CT images was insufficient for manual segmentation, leading to the exclusion of some individuals. In this context, AI-segmentation of a large dataset without human prescreening should be interpreted with caution.

4.4 | Future directions

Future work should focus on two main areas. First, although the reliability of AutoMATiCA was higher than that of ABACS-SliceOmatic, it partially failed in 23 cases, with a DSC value <0.9 , in comparison to the manual method. The DSC value for ABACS-SliceOmatic was also <0.9 in the vast majority of these cases. We demonstrated that both ABACS-SliceOmatic

and AutoMATiCA are less reliable for very low BMI values. We suggest that the number of underweight subjects should be increased in the training cohort for AutoMATiCA to improve the performance of this solution. The proportion of underweight individuals in the study by Paris et al. was 1.3% in the training and validation cohort versus 7.5% in our cohort and 30% in the subgroup of cases with a DSC for AutoMATiCA <0.9 . When using AutoMATiCA in underweight subjects, we suggest that clinicians or researchers should manually check the segmented images. We also demonstrated that reliability is better with higher current intensities, in agreement with a previous study.²⁴ Further work needs to be performed to improve the applicability of deep learning model for varying currents of CT. Second, the automation of segmentation is an important step for sarcopenia diagnosis in clinical practice, but other

TABLE 3 Impact of the BMI, the current intensity, the IV contrast and the voltage on the reliability of AutoMATiCA and ABACS-SliceOmatic using a mixed-model linear regression analysis.

	Univariate analysis		Bivariate with interaction analysis	
	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value
BMI (kg/m ²)	0.04 (0.03–0.05)	<0.001	0.03 (0.02–0.04)	<0.001
Solution				
ABACS-SliceOmatic	Ref		Ref	
AutoMATiCA	0.57 (0.53–0.60)	<0.001	0.56 (0.52–0.60)	<0.001
Interaction for the BMI and the solution			0.02 (0.01–0.03)	<0.001
	Univariate analysis		Bivariate with interaction analysis	
Current (mA)	0.27 (0.18; 0.36)	<0.001	0.24 (0.14; 0.34)	<0.001
Solution				
ABACS-SliceOmatic	Ref		Ref	
AutoMATiCA	0.57 (0.53–0.60)	<0.001	0.53 (0.48; 0.59)	<0.001
Interaction for the current and the solution			0.07 (–0.01; 0.15)	<i>p</i> = 0.08
	Univariate analysis			
IV contrast				
No	Ref			
Yes	0.03 (–0.09; 0.14)	0.67		
	Univariate analysis			
Voltage				
100	–0.14 (–0.44; 0.16)	0.35		
120	Ref			
140	–0.09 (–0.78; 0.60)	0.80		

For the multivariate regression analyses, the BMI was centered at 25 kg/m².

steps remain to be automated. In particular, manual determination of the slice at the L3 level in the CT-scan is time-consuming. Convolutional networks can be used for this step.^{25,26} We are currently validating such algorithms in order to develop a software solution that can fully automate the three following steps (1) identification of the L3 slice in an abdominal CT series, (2) segmentation of skeletal muscle, and (3) skeletal muscle index (SMI) calculation. Such an automatic pipeline has been already developed for muscle segmentation at the T12 level.²⁷

5 | CONCLUSIONS

In conclusion, our work has shown that AutoMATiCA is more reliable than ABACS-SliceOmatic for skeletal muscle segmentation from CT images at the L3 level in a cohort of individuals hospitalized in a French University Hospital. Unlike AutoMATiCA, the performance of ABACS-SliceOmatic is reduced in cases of anatomic differences or poor-quality images. The next step is to develop and validate a neural network to identify the L3 slice in a CT series.

ACKNOWLEDGMENTS

This work was supported by a grant from the Délégation Régionale à la Recherche Clinique du Centre Hospitalier Universitaire Grenoble Alpes (2019). The funding bodies did not have any involvement in the design/conduct of the research, in data analysis/interpretation, or in writing/approval of the manuscript.

CONFLICT OF INTEREST STATEMENT

Katia Charrière, Quentin Boulouard, Svetlana Artemova, Antoine Vilotitch, Jean-Luc Bosson, Alexandre Moreau-Gaudry, Joris Giai, Eric Fontaine, and Cécile Bétry have no relevant conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, CB upon reasonable request.

REFERENCES

1. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing*. 2019;48(1):16–31. doi:10.1093/ageing/afy169

2. Sousa AS, Guerra RS, Fonseca I, Pichel F, Amaral TF. Sarcopenia and length of hospital stay. *Eur J Clin Nutr.* 2016;70(5):595. doi:10.1038/ejcn.2015.207
3. Xu J, Wan CS, Ktoris K, Reijnierse EM, Maier AB. Sarcopenia is associated with mortality in adults: a systematic review and meta-analysis. *Gerontology.* 2022;68(4):1-16. doi:10.1159/000517099
4. Sivaharan A, Boylan L, Witham MD, Nandhra S. Sarcopenia in patients undergoing lower limb bypass surgery is associated with higher mortality and major amputation rates. *Ann Vasc Surg.* 2021;75:227-236. doi:10.1016/j.avsg.2021.02.022
5. Prado CM, Lieffers JR, McCargar LJ, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* 2008;9(7):629-635. doi:10.1016/S1470-2045(08)70153-0
6. Findlay M, White K, Stapleton N, Bauer J. Is sarcopenia a predictor of prognosis for patients undergoing radiotherapy for head and neck cancer? A meta-analysis. *Clin Nutr.* 2021;40(4):1711-1718. doi:10.1016/j.clnu.2020.09.017
7. Cox MC, Booth M, Ghita G, et al. The impact of sarcopenia and acute muscle mass loss on long-term outcomes in critically ill patients with intra-abdominal sepsis. *J Cachexia Sarcopenia Muscle.* 2021;12(5):1203-1213. doi:10.1002/jcsm.12752
8. Albano D, Messina C, Vitale J, Sconfienza LM. Imaging of sarcopenia: old evidence and new insights. *Eur Radiol.* 2020;30(4):2199-2208. doi:10.1007/s00330-019-06573-2
9. Amini B, Boyle SP, Boutin RD, Lenchik L. Approaches to assessment of muscle mass and myosteatosis on computed tomography: a systematic review. *J Gerontol A Biol Sci Med Sci.* 2019;74(10):1671-1678. doi:10.1093/gerona/glz034
10. Popuri K, Cobzas D, Esfandiari N, Baracos V, Jägersand M. Body composition assessment in axial CT images using FEM-based automatic segmentation of skeletal muscle. *IEEE Trans Med Imaging.* 2016;35(2):512-520. doi:10.1109/TMI.2015.2479252
11. Chung H, Cobzas D, Birdsell L, Lieffers J, Baracos V. Automated segmentation of muscle and adipose tissue on CT images for human body composition analysis. *Med Imaging, Visualization, Image-Guided Procedures, and Modeling.* 2009;7261:72610K. doi:10.1117/12.812412
12. Blanc-Durand P, Schiratti JB, Schutte K, et al. Abdominal musculature segmentation and surface prediction from CT using deep learning for sarcopenia assessment. *Diagn Interv Imaging.* 2020;101(12):789-794. doi:10.1016/j.diii.2020.04.011
13. Borrelli P, Kaboteh R, Enqvist O, et al. Artificial intelligence-aided CT segmentation for body composition analysis: a validation study. *Eur Radiol Exp.* 2021;5:11. doi:10.1186/s41747-021-00210-8
14. Paris MT, Tandon P, Heyland DK, et al. Automated body composition analysis of clinically acquired computed tomography scans using neural networks. *Clin Nutr.* 2020;39(10):3049-3055. doi:10.1016/j.clnu.2020.01.008
15. Caudron L, Bussy A, Artemova S, et al. Sarcopenia diagnosis: comparison of automated with manual computed tomography segmentation in clinical routine. *JCSM Rapid Commun.* 2021;4(2):103-110. doi:10.1002/rco2.37
16. Cespedes Feliciano EM, Popuri K, Cobzas D, et al. Evaluation of automated computed tomography segmentation to assess body composition and mortality associations in cancer patients. *J Cachexia Sarcopenia Muscle.* 2020;11(5):1258-1269. doi:10.1002/jcsm.12573. Published online April 20.
17. Paris MT, Furberg HF, Petruzella S, Akin O, Hötter AM, Mourtzakis M. Influence of contrast administration on computed tomography-based analysis of visceral adipose and skeletal muscle tissue in clear cell renal cell carcinoma. *JPEN J Parenter Enteral Nutr.* 2018;42(7). doi:10.1002/jpen.1067. Published online January 19.
18. van Vugt JLA, Coebergh van den Braak RRJ, Schippers HJW, et al. Contrast-enhancement influences skeletal muscle density, but not skeletal muscle mass, measurements on computed tomography. *Clin Nutr.* 2018;37(5):1707-1714. doi:10.1016/j.clnu.2017.07.007. Published online July 14.
19. Artemova S, Madiot PE, Caporossi A, Group P, Mossuz P, Moreau-Gaudry A. PREDIMED: clinical data warehouse of Grenoble Alpes University Hospital. *Stud Health Technol Inform.* 2022;290:1068-1069. doi:10.3233/SHTI190464
20. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Published online May 18, 2015. Accessed July 21, 2021. <https://arxiv.org/abs/1505.04597v1>
21. Dice LR. Measures of the amount of ecologic association between species. *Ecology.* 1945;26(3):297-302. doi:10.2307/1932409
22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012
23. Mahajan V, Venugopal VK, Murugavel M, Mahajan H. The algorithmic audit: working with vendors to validate radiology-AI algorithms-how we do it. *Acad Radiol.* 2020;27(1):132-135. doi:10.1016/j.acra.2019.09.009
24. Kim DW, Ha J, Ko Y, et al. Reliability of skeletal muscle area measurement on CT with different parameters: a phantom study. *Korean J Radiol.* 2021;22(4):624-633. doi:10.3348/kjr.2020.0914
25. Belharbi S, Chatelain C, Héroult R, et al. Spotting L3 slice in CT scans using deep convolutional network and transfer learning. *Comput Biol Med.* 2017;87:95-103. doi:10.1016/j.combiomed.2017.05.018
26. Chen Y, Gao Y, Li K, Zhao L, Zhao J. Vertebrae identification and localization utilizing fully convolutional networks and a hidden markov model. *IEEE Trans Med Imaging.* 2020;39(2):387-399. doi:10.1109/TMI.2019.2927289
27. Lenchik L, Barnard R, Boutin RD, et al. Automated muscle measurement on chest CT predicts all-cause mortality in older adults from the national lung screening trial. *J Gerontol A Biol Sci Med Sci.* 2020;76(2):277-285. doi:10.1093/gerona/glaa141

SUPPORTING INFORMATION

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

How to cite this article: Charrière K, Boulouard Q, Artemova S, et al. A comparative study of two automated solutions for cross-sectional skeletal muscle measurement from abdominal computed tomography images. *Med Phys.* 2023;50:4973–4980. <https://doi.org/10.1002/mp.16261>