Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) 2024 25(1):83-90
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Correspondence

https://doi.org/10.1631/jzus.B2300363

Check for updates

Deep learning-based radiomics allows for a more accurate assessment of sarcopenia as a prognostic factor in hepatocellular carcinoma

Zhikun LIU[1,5*], Yichao WU[2,5*], Abid Ali KHAN[2,5], Lun LU[3], Jianguo WANG[1,5], Jun CHEN[1,5], Ningyang JIA[3✉], Shusen ZHENG[4,5,6✉], Xiao XU[1,2,5✉]

[1]*Key Laboratory of Integrated Oncology and Intelligent Medicine of Zhejiang Province, Department of Hepatobiliary and Pancreatic Surgery, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China*
[2]*Zhejiang University School of Medicine, Hangzhou 310058, China*
[3]*Department of Radiology, Easter Hepatobiliary Surgery Hospital, Naval Medical University, Shanghai 200438, China*
[4]*Department of Hepatobiliary and Pancreatic Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*
[5]*NHC Key Laboratory of Combined Multi-organ Transplantation, Hangzhou 310003, China*
[6]*Department of Hepatobiliary and Pancreatic Surgery, Shulan (Hangzhou) Hospital, Hangzhou 310022, China*

Hepatocellular carcinoma (HCC) is one of the most common malignancies and is a major cause of cancer-related mortalities worldwide (Forner et al., 2018; He et al., 2023). Sarcopenia is a syndrome characterized by an accelerated loss of skeletal muscle (SM) mass that may be age-related or the result of malnutrition in cancer patients (Cruz-Jentoft and Sayer, 2019). Preoperative sarcopenia in HCC patients treated with hepatectomy or liver transplantation is an independent risk factor for poor survival (Voron et al., 2015; van Vugt et al., 2016). Previous studies have used various criteria to define sarcopenia, including muscle area and density. However, the lack of standardized diagnostic methods for sarcopenia limits their clinical use. In 2018, the European Working Group on Sarcopenia in Older People (EWGSOP) renewed a consensus on the definition of sarcopenia: low muscle strength, loss of muscle quantity, and poor physical performance (Cruz-Jentoft et al., 2019). Radiological imaging-based measurement of muscle quantity or mass is most commonly used to evaluate the degree of sarcopenia. The gold standard is to measure the SM and/or psoas muscle (PM) area using abdominal computed tomography (CT) at the third lumbar vertebra (L3), as it is linearly correlated to whole-body SM mass (van Vugt et al., 2016). According to a "North American Expert Opinion Statement on Sarcopenia," SM index (SMI) is the preferred measure of sarcopenia (Carey et al., 2019). The variability between morphometric muscle indexes revealed that they have different clinical relevance and are generally not applicable to broader populations (Esser et al., 2019).

Radiomics refers to the large-scale, algorithm-based, quantitative analysis of imaging features, which can reveal disease features and underlying pathophysiological features. Deep learning is a data-driven analytical method that allows for the mining of images with hidden clinical value, and has shown promising results in diagnosing HCC and determining its pathology, prognosis, and response to therapy (Jin et al., 2021). Coupled with radiomics, deep learning can automatically learn features from imaging labels and be used to plan individualized treatment of cancer patients (Bi et al., 2019). In this study, we assessed muscle mass and determined its prognosis significance in HCC patients using a deep learning-based radiomic survival model implemented by tensorflow (TFDeepSurv).

✉ Xiao XU, zjxu@zju.edu.cn
Shusen ZHENG, zyzss@zju.edu.cn
Ningyang JIA, ningyangjia@163.com
* The two authors contributed equally to this work
Xiao XU, https://orcid.org/0000-0002-2761-2811
Shusen ZHENG, https://orcid.org/0000-0003-1459-8261
Ningyang JIA, https://orcid.org/0000-0001-9587-3637

Received June 9, 2023; Revision accepted Aug. 25, 2023;
Crosschecked Dec. 21, 2023

© Zhejiang University Press 2024

This study included HCC patients who had undergone hepatectomy (n=492, discovery cohort, divided into the training and internal test sets at a 7:3 ratio) and liver transplantation (n=173, external LT test set) at the Department of Hepatobiliary and Pancreatic Surgery of the First Affiliated Hospital, Zhejiang University School of Medicine (Hangzhou, China), and 161 patients who had undergone hepatectomy at Easter Hepatobiliary Surgery Hospital, Naval Medical University (Shanghai, China) (external test set). Patients were eligible for inclusion if they had primary HCC confirmed by pathology and had undergone abdominal CT within the month before surgery. Patients with unsatisfactory image quality or missing clinical data were excluded. The clinical characteristics of the subjects in the training and validation sets are summarized in Table 1. The discovery and the external test cohorts were similar in terms of baseline clinical characteristics. In the external LT test cohort, the majority of recipients were female and the tumors were mostly multiple. Radiomic features from CT scans at L3 were used to develop the TFDeepSurv model. A flowchart of the research process is shown in Fig. 1. The theories of AutoEncoder and the TFDeepSurv network are shown in the supplementary section materials and methods.

Plain abdominal CT images of the patients were obtained from the hospital's picture archiving and communication systems in digital imaging and communications in medicine (DICOM) format. The details of the CT images can be found in the supplementary

Table 1 Subject characteristics of discovery and test sets

Characteristics	Discovery cohort (n=492)	External test (n=161)	External LT test (n=173)
Age (years)*	55.9±11.1	55.0±11.2	52.4±9.3
BMI (kg/m^2)*	22.9±2.9	24.0±3.4	23.3±3.3
Gender			
Male	418 (85.0%)	139 (86.3%)	23 (13.3%)
Female	74 (15.0%)	22 (13.7%)	150 (86.7%)
Tumor size			
Total diameter≥5 cm	128 (41.3%)[a]	97 (60.2%)	56 (32.4%)[b]
Tumor number			
Single	415 (84.3%)[c]	150 (93.2%)	77 (44.5%)[d]
Vascular invasion			
Yes	70 (14.2%)[e]	NA[f]	46 (26.6%)

* Data are shown as mean±standard deviation. [a] Miss data (n=35); [b] Miss data (n=1); [c] Miss data (n=35); [d] Miss data (n=1); [e] Miss data (n=24); [f] Miss data (n=161). BMI, body mass index; NA, not available.

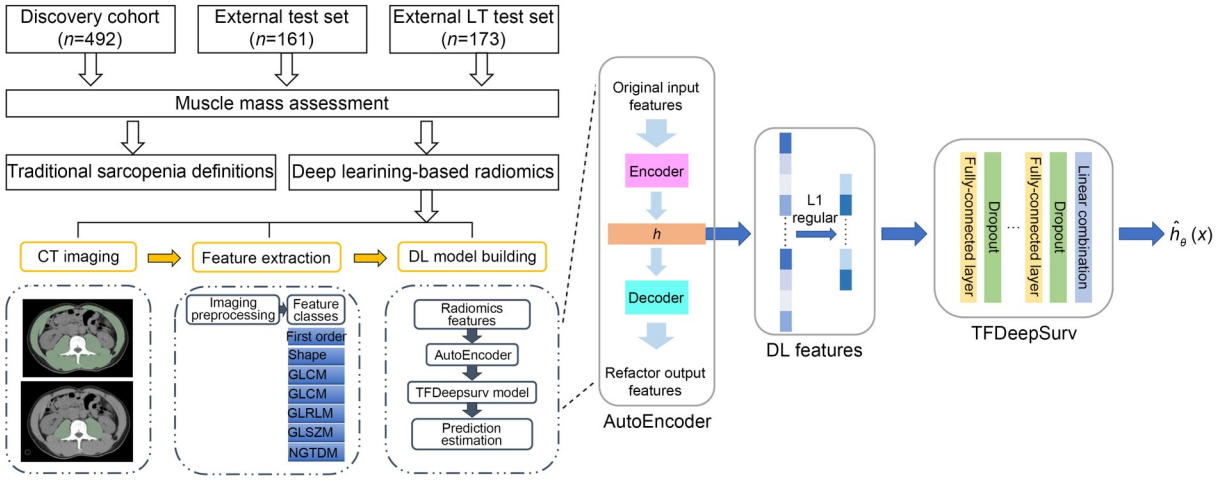

Fig. 1 Overview of the research process. The left panel shows the overall workflow of the proposed study. The right panel shows the architecture of the deep neural network prediction model for the prognosis of hepatocellular carcinoma (HCC) based on AutoEncoder and TFDeepSurv. CT: computed tomography; DL: deep learning; TFDeepSurv: deep learning-based radiomic survival model implemented by tensorflow.

section materials and methods. After the optimal window width and position were set, region of interest (ROI) segmentation was performed with a three-dimensional (3D) slicer. ROIs in SM and PM were manually delineated at the L3 level of the CT in each patient by an experienced clinician, and the delineated images were reviewed by a senior clinician. The patients' clinical data were obtained from the hospital's medical records, such as survival time and recurrence-free time.

According to Carey et al. (2017), sarcopenia is indicated by an SMI of ≤50 cm^2/m^2 in males and ≤39 cm^2/m^2 in females, while according to Golse et al. (2017), the PM area (PMA) is <15.61 cm^2 in males and <14.64 cm^2 in females. According to Hamaguchi et al. (2016), the PM index (PMI) is <6.36 cm^2/m^2 in men and <3.92 cm^2/m^2 in women, and according to Yoo et al. (2017), a PM density (PD) of <38.5 Hounsfield unit (HU) indicates sarcopenia. Fujiwara et al. (2015) used mean muscle attenuation (MA) as an indicator of intermuscular fat (IMF) deposition and found that a positive MA result is ≤44.4 HU in men and ≤39.3 HU in women. The prevalence of preoperative sarcopenia in patients with HCC varies according to different definitions of sarcopenia. In our cohort, based on definitions by Golse et al. (2017), Hamaguchi et al. (2016), Yoo et al. (2017), Carey et al. (2017), and Fujiwara et al. (2015), sarcopenia was present in 216 (43.9%), 263 (53.5%), 20 (4.1%), 290 (58.9%), and 352 (71.5%) patients, respectively. Its prevalence among 173 LT cases ranged from 2.9% to 78.6% (Table 2).

In the discovery cohort, patients with sarcopenia at the time of operation had significantly lower overall survival (OS) according to three muscle indices (PMI, SMI, and PMA), while based on other muscle indices (IMF and PD), sarcopenia was not associated with OS (Fig. 2). However, sarcopenia was not associated with recurrence-free survival (RFS) according to any muscle indices (Fig. S1). Additionally, only the patients with sarcopenia based on PMI had a trend of lower OS than those without sarcopenia in the external test cohort (Fig. S2), and the patients with sarcopenia based on PD had a significantly lower OS than the patients without sarcopenia in the LT cohort (Fig. S3). To summarize, sarcopenia was more associated with OS. Finally, the TFDeepSurv model was constructed to predict OS.

Radiomics feature extraction was carried out using PyRadiomics, a tool that extracts standardized radiomic features from imaging data. Firstly, we extracted 1343 radiomics features from muscle images, including three standardized feature classes: first-order statistics, shape descriptors, and texture features (including a grey level co-occurrence matrix, a grey level run length matrix (GLCM), a grey level size zone matrix (GLSZM), a grey level dependence matrix (GLDM), and a neighboring gray tone difference matrix (NGTDM)). Features not only contain the SM area but

Table 2 Prevalence of sarcopenia in different definitions and its association with outcomes

Definition of sarcopenia and muscle mass assessment	Discovery cohort (*n*=492)			External test set (*n*=161)		External LT test set (*n*=173)		
	n (%)	RFS (HR, 95% CI)	OS (HR, 95% CI)	*n* (%)	OS (HR, 95% CI)	*n* (%)	RFS (HR, 95% CI)	OS (HR, 95% CI)
Golse et al. (2017)	216	1.04	1.58	54	1.58	71	1.24	1.59
M: PMA<1561 mm^2	(43.9%)	(0.80–1.35)	(1.07–2.35)	(33.5%)	(0.81–3.06)	(41.0%)	(0.64–2.38)	(0.76–3.31)
F: PMA<1464 mm^2		*P*=0.76	*P*=0.02		*P*=0.18		*P*=0.52	*P*=0.22
Hamaguchi et al. (2016)	263	1.14	1.57	81	1.27	100	1.37	1.47
M: PMI<6.36 cm^2/m^2	(53.5%)	(0.88–1.48)	(1.05–2.35)	(50.3%)	(0.66–2.45)	(57.8%)	(0.70–2.65)	(0.70–3.13)
F: PMI<3.92 cm^2/m^2		*P*=0.31	*P*=0.03		*P*=0.48		*P*=0.36	*P*=0.31
Yoo et al. (2017)	20	1.27	2.11	0	NA	5	5.98	5.47
PD<38.5 HU	(4.1%)	(0.68–2.40)	(0.92–4.85)			(2.9%)	(1.33–26.90)	(1.25–23.92)
		P=0.45	*P*=0.08				*P*=0.02	*P*=0.02
Carey et al. (2017)	290	1.10	1.52	97	1.92	87	1.08	1.23
M: SMI≤50 cm^2/m^2	(58.9%)	(0.84–1.43)	(1.00–2.31)	(60.2%)	(0.92–3.98)	(50.3%)	(0.55–2.11)	(0.57–2.66)
F: SMI≤39 cm^2/m^2		*P*=0.49	*P*=0.05		*P*=0.08		*P*=0.83	*P*=0.60
Fujiwara et al. (2015)	352	1.31	1.52	8	0.52	136	1.34	2.44
M: MA≤44.4 HU	(71.5%)	(0.97–1.76)	(0.96–2.40)	(5.0%)	(0.07–3.82)	(78.6%)	(0.61–2.95)	(0.84–7.05)
F: MA≤39.3 HU		*P*=0.07	*P*=0.08		*P*=0.52		*P*=0.47	*P*=0.10

LT, liver transplantation; RFS, recurrence-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval; M: male; F: female; PMA, psoas muscle area; PMI, psoas muscle index; PD, psoas muscle density; SMI, skeletal muscle index; MA, mean muscle attenuation; HU, Hounsfield unit; NA, not available.

also its density, shape, and textural features. In total, 1343 radiomic features were extracted from SMs containing psoas at L3 level CT images, which comprised various muscle mass information. To assess the performance of numerous radiomics features, a 3D-vector was obtained after using principal component analysis (PCA) to reduce dimensionality. As expected, radiomics features also have good power to discriminate between various sarcopenia definitions, except for the definition by Yoo et al. (2017) (Fig. 3), indicating that radiomics features of muscle are associated with sarcopenia.

Radiomic features have a lot of redundancy and many features are linearly related to each other. Therefore, we used AutoEncoder to reduce the dimensionality

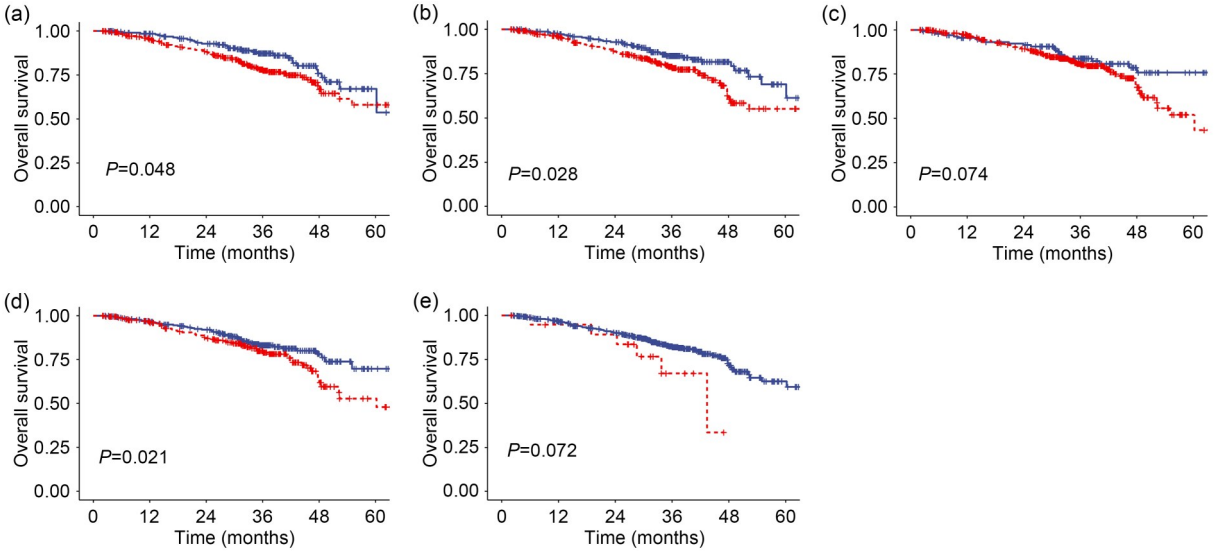

Fig. 2 Role of traditional sarcopenia definitions in the prognosis of overall survival in patients with hepatocellular carcinoma (HCC) who have undergone liver resection. (a) Skeletal muscle index (SMI); (b) Psoas muscle index (PMI); (c) Intermuscular fat (IMF); (d) Psoas muscle area (PMA); (e) Psoas muscle density (PD). Red, sarcopenia; blue, non-sarcopenia.

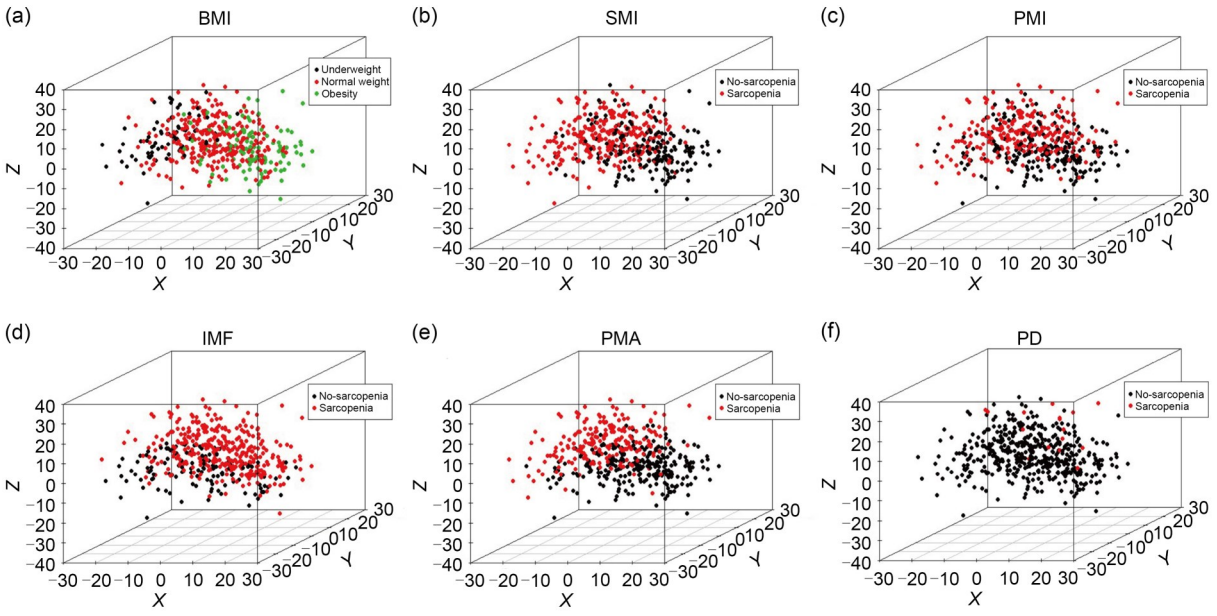

Fig. 3 Ability of radiomics features to discriminate between different sarcopenia definitions. (a) Body mass index (BMI); (b) Skeletal muscle index (SMI); (c) Psoas muscle index (PMI); (d) Intermuscular fat (IMF); (e) Psoas muscle area (PMA); (f) Psoas muscle density (PD).

of original radiomics features to obtain deep learning features (relatively small and not related to each other), and TFDeepSurv to construct a prognosis prediction model. TFDeepSurv was constructed as follows: to speed up the analysis and fully use its features, 1343 radiomics features were compressed into 100 variables using AutoEncoder, and the TFDeepSurv survival network was trained based on those variables. Figs. 4a and 4b show the training process. The medium-based cut-off of the TFDeepSurv stratified the HCC patients in the training set into two risk groups (Fig. 4c), and its prognostic stratification power was confirmed in the internal (Fig. 4d), the external (Fig. 4e), and the LT test sets (Fig. 4f) via Kaplan-Meier survival analysis.

The time-dependent area under the receiver operating characteristic curve (AUC) and the concordance index (C-index) values are depicted in Fig. 5. Compared to the other sarcopenia definitions, TFDeepSurv maintained higher AUC values in the internal test set (Fig. 5a), the external test set (Fig. 5c), and the external LT test set (Fig. 5e). As anticipated, TFDeepSurv exhibited a significant predictive ability in the internal test set (C-index, 0.730; 95% confidence interval (CI), 0.707–0.753; Fig. 5b), the external test set (C-index, 0.667; 95% CI, 0.651–0.682; Fig. 5d), and the external LT test set (C-index, 0.653; 95% CI, 0.633–0.677; Fig. 5f). This is the first attempt to decode SM mass with radiomics in tumor patients, independent of any definition of sarcopenia. The principal finding was that, despite substantial heterogeneity among patients from multiple centers, TFDeepSurv still exhibited a superior performance to the traditional definitions of sarcopenia in the test sets.

SM mass is an important predictor of HCC outcomes, but many questions, such as "what is the best modality for assessing muscle?" "what are the ideal timing and frequency of muscle mass assessment?" and "how to incorporate the method of assessment into clinical decision-making?" remain unanswered. Sarcopenia, which is a metric of SM depletion, is used to

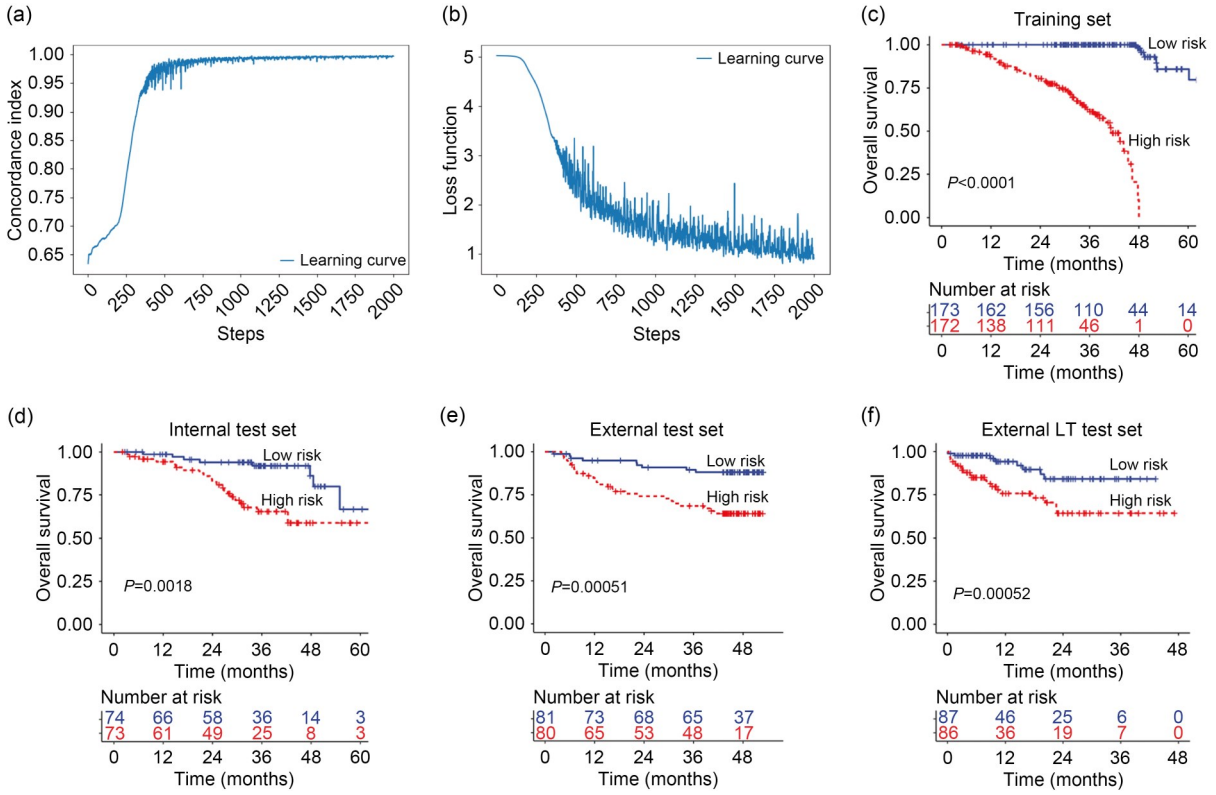

Fig. 4 Development of the TFDeepSurv and Kaplan-Meier survival curves of training and test sets. (a, b) The trend of the concordance index (a) and loss function (b) of the deep neural network training process. (c–f) The TFDeepSurv divided patients with hepatocellular carcinoma (HCC) into high- and low-risk groups with a significantly different overall survival (OS) in the training group (c), the internal test set (d), the external test set (e), and the external liver transplantation (LT) test set (f). TFDeepSurv: deep learning-based radiomic survival model implemented by tensorflow.

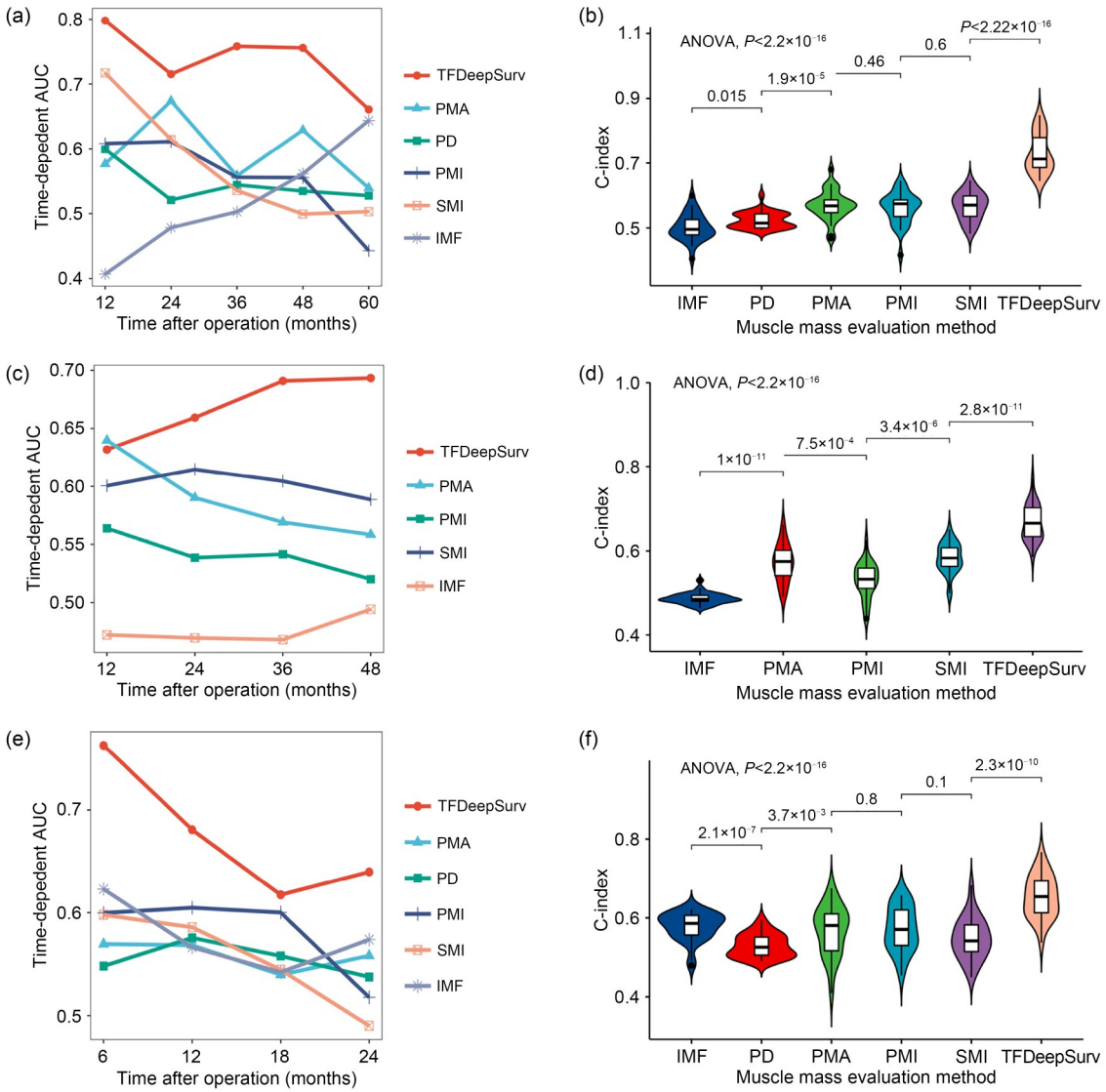

Fig. 5 Comparison of the predictive performance of TFDeepSurv vs. various sarcopenia definitions. (a, b) Time-dependent area under the receiver operating characteristic curve (AUC) (a) and C-index (b) in the internal test set. (c, d) Time-dependent AUC (c) and C-index (d) in the external test set. (e, f) Time-dependent AUC (e) and C-index (f) in the external liver transplantation (LT) test set. TFDeepSurv: deep learning-based radiomic survival model implemented by tensorflow; PMA: psoas muscle area; PD: psoas muscle density; PMI: psoas muscle index; SMI: skeletal muscle index; IMF: intermuscular fat; ANOVA: analysis of variance.

assess muscle mass in clinical practice. Despite the existence of multiple definitions of sarcopenia, there is no consensus on the threshold value associated with poor survival that accurately stratifies patients awaiting surgery. However, based on the traditional definitions, the prevalence of sarcopenia varied from 2.9% to 78.6% in our cohort. Until now, radiomics has rarely been used to assess sarcopenia. Kim (2021) successfully identified sarcopenia based on radiomic features and machine learning methods, which proved that radiomic features could decode the phenomenon in the muscle. However, this was a pilot study of only a limited number of patients and lacked sufficient validation. Chen et al. (2022) proposed a new CT radiomics-based method for diagnosing sarcopenia, which could effectively improve the predictive accuracy of prognosis compared to the traditional sarcopenia definitions. However, the above research defined sarcopenia using different criteria. There are no established evaluative indicators or cut-off values for each traditional definition of sarcopenia, which may limit their application to different populations. Hence, there is a desperate need for

an objective muscle mass evaluation system assisted by deep learning and radiomics.

This study is limited by its retrospective nature. The use of varying CT imaging equipment in diverse patient populations may display different muscle composition phenotypes, potentially leading to biases during feature extraction. Thus, to validate the robustness and generalizability of our findings, larger prospective studies are warranted.

In conclusion, we developed a TFDeepSurv system for accessing SMs and predicting the survival of patients with HCC who have undergone liver resection and transplantation. The prognostic stratification power of TFDeepSurv was superior to other traditional definitions of sarcopenia, so it can serve as a visual prognostic tool that can assess muscles and help clinicians to identify patients with a high mortality risk and plan their treatment.

Data availability statement

The dataset used or analyzed during the current study is available from the corresponding author on reasonable request.

Materials and methods

Detailed methods are provided in the electronic supplementary materials of this paper.

Acknowledgments

This work was supported by the Key Program of Provincial Natural Foundation of Zhejiang Province (No. LZ22H180003), the National Natural Science Foundation of China (Nos. 92159202 and 81802889), and the Key Research & Development Program of Zhejiang Province (No. 2022C03108). We are grateful to Wiley Editing Services for assisting in language editing assistance (Order No. TNIPJ_1).

Author contributions

Xiao XU, Shusen ZHENG, Ningyang JIA, and Zhikun LIU performed the conceptualization; Yichao WU, Lun LU, Ningyang JIA, Abid Ali KHAN, Jianguo WANG, and Jun CHEN contributed to the data curation; Xiao XU, Zhikun LIU, and Jun CHEN performed the funding acquisition; Zhikun LIU and Yichao WU performed the methodology; Zhikun LIU, Yichao WU, and Lun LU performed the investigation; Xiao XU and Zhikun LIU performed the project administration; Yichao WU, Zhikun LIU, and Abid Ali KHAN wrote the manuscript; Xiao XU, Shusen ZHENG, and Ningyang JIA contributed to writing – review & editing. All authors have read and approved the final manuscript, and therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Xiao XU is an editorial board member for *Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)* and was not involved in the editorial review or the decision to publish this article.

Zhikun LIU, Yichao WU, Abid Ali KHAN, Lun LU, Jianguo WANG, Jun CHEN, Ningyang JIA, Shusen ZHENG, and Xiao XU declared that they have no conflict of interest.

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (the Ethics Committee of Affiliated Hangzhou First People's Hospital) and with the Helsinki Declaration of 1975, as revised in 2013. The individual consent for this retrospective analysis was waived.

References

Bi WL, Hosny A, Schabath MB, et al., 2019. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin*, 69(2):127-157. https://doi.org/10.3322/caac.21552

Carey EJ, Lai JC, Wang CW, et al., 2017. A multicenter study to define sarcopenia in patients with end-stage liver disease. *Liver Transpl*, 23(5):625-633. https://doi.org/10.1002/lt.24750

Carey EJ, Lai JC, Sonnenday C, et al., 2019. A North American expert opinion statement on sarcopenia in liver transplantation. *Hepatology*, 70(5):1816-1829. https://doi.org/10.1002/hep.30828

Chen XD, Chen WJ, Huang ZX, et al., 2022. Establish a new diagnosis of sarcopenia based on extracted radiomic features to predict prognosis of patients with gastric cancer. *Front Nutr*, 9:850929. https://doi.org/10.3389/fnut.2022.850929

Cruz-Jentoft AJ, Sayer AA, 2019. Sarcopenia. *Lancet*, 393(10191):2636-2646. https://doi.org/10.1016/s0140-6736(19)31138-9

Cruz-Jentoft AJ, Bahat G, Bauer J, et al., 2019. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing*, 48(4):601. https://doi.org/10.1093/ageing/afz046

Esser H, Resch T, Pamminger M, et al., 2019. Preoperative assessment of muscle mass using computerized tomography scans to predict outcomes following orthotopic liver transplantation. *Transplantation*, 103(12):2506-2514. https://doi.org/10.1097/tp.0000000000002759

Forner A, Reig M, Bruix J, 2018. Hepatocellular carcinoma. *Lancet*, 391(10127):1301-1314. https://doi.org/10.1016/s0140-6736(18)30010-2

Fujiwara N, Nakagawa H, Kudo Y, et al., 2015. Sarcopenia, intramuscular fat deposition, and visceral adiposity independently predict the outcomes of hepatocellular carcinoma. *J Hepatol*, 63(1):131-140. https://doi.org/10.1016/j.jhep.2015.02.031

Golse N, Bucur PO, Ciacio O, et al., 2017. A new definition of sarcopenia in patients with cirrhosis undergoing liver transplantation. *Liver Transpl*, 23(2):143-154. https://doi.org/10.1002/lt.24671

Hamaguchi Y, Kaido T, Okumura S, et al., 2016. Proposal for new diagnostic criteria for low skeletal muscle mass based

on computed tomography imaging in Asian adults. *Nutrition*, 32(11-12):1200-1205. https://doi.org/10.1016/j.nut.2016.04.003

He ZQ, She XM, Liu ZY, et al., 2023. Advances in postoperative prognostic models for hepatocellular carcinoma. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 24(3):191-206. https://doi.org/10.1631/jzus.B2200067

Jin JY, Yao Z, Zhang T, et al., 2021. Deep learning radiomics model accurately predicts hepatocellular carcinoma occurrence in chronic hepatitis b patients: a five-year follow-up. *Am J Cancer Res*, 11(2):576-589.

Kim YJ, 2021. Machine learning models for sarcopenia identification based on radiomic features of muscles in computed tomography. *Int J Environ Res Public Health*, 18(16):8710. https://doi.org/10.3390/ijerph18168710

van Vugt JLA, Levolger S, de Bruin RWF, et al., 2016. Systematic review and meta-analysis of the impact of computed tomography-assessed skeletal muscle mass on outcome in patients awaiting or undergoing liver transplantation. *Am J Transpl*, 16(8):2277-2292. https://doi.org/10.1111/ajt.13732

Voron T, Tselikas L, Pietrasz D, et al., 2015. Sarcopenia impacts on short- and long-term results of hepatectomy for hepatocellular carcinoma. *Ann Surg*, 261(6):1173-1183. https://doi.org/10.1097/sla.0000000000000743

Yoo T, Lo WD, Evans DC, 2017. Computed tomography measured psoas density predicts outcomes in trauma. *Surgery*, 162(2):377-384. https://doi.org/10.1016/j.surg.2017.03.014

Supplementary information

Materials and methods; Figs. S1–S3