Journal of Zhejiang University-SCIENCE B in press
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Research Article

https://doi.org/10.1631/jzus.B2600062

Adverse impacts of body composition abnormalities on outcome after liver transplantation for hepatocellular carcinoma: Development of a multi-task classifier chain surrogate model

Junjie YIN[1,2*], Zhu LI[3*], Jinyan CHEN[1], Qingyang QUE[1], Huigang LI[1], Ruijie ZHAO[1], Jianyong ZHUO[6], Chiyu HE[1], Wei SHEN[1], Peiru ZHANG[5], Yifan GU[1], Chenghao CAO[1], Jiaqi ZHENG[1], Yuhang LI[1], Rongsen WANG[3], Zuyuan LIN[1], Shengjun XU[2], Xuyong WEI[2], Shusen ZHENG[4], Di LU[3]✉, Xiao XU[1,3,4,6]✉

[1] *Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou 310000, China*
[2]*Department of Hepatobiliary and Pancreatic Surgery, Affiliated Hangzhou First People's Hospital, School of Medicine, Westlake University, Hangzhou 310006, China*
[3]*School of Basic Medical Sciences and Forensic Medicine, Hangzhou Medical College, Hangzhou 310000, China*
[4]*NHC Key Laboratory of Combined Multi-organ Transplantation, Hangzhou 310000, China*
[5]*The Fourth School of Clinical Medicine Zhejiang Chinese Medical University, Hangzhou 310003, China*
[6]*Hepatobiliary Center, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210000, China*

Abstract: Background: Sarcopenia and adipose abnormalities are co-occurring risks in liver transplantation for hepatocellular carcinoma (HCC), yet reliance on computed tomography (CT) segmentation limits routine assessment. Thus, we aimed to quantify the cumulative prognostic burden of these correlated phenotypes and develop a multi-task classifier chain surrogate to predict them using standard clinical data. Methods: In a cohort of 409 HCC transplant recipients, we analyzed the impact of four CT-derived phenotypes on recurrence-free survival (RFS) and overall survival (OS). To predict these correlated comorbidities without CT, we designed a Multi-task Classifier Chain framework using 37 routine peri-transplant variables. This architecture was explicitly trained to exploit inter-phenotype dependencies, comparing performance against standard multi-task baselines. Results: All four phenotypes could independently predict poor outcomes. A profound "dose-response" relationship was observed: 5-year OS dropped from 77.3% (≤1 abnormality) to 19.6% (≥3 abnormalities; P<0.0001). In the test set, the Classifier Chain framework significantly outperformed baselines by capturing inter-phenotype dependencies. It achieved high discrimination for sarcopenia (area under the receiver operating characteristic curve (AUC), 0.88), myosteatosis (AUC 0.88), visceral obesity (AUC 0.87), and high subcutaneous adipose tissue density (SATD) (AUC 0.79), yielding the greatest net clinical benefit. Conclusions: Body composition abnormalities exert a severe, additive negative impact on post-transplant outcomes. By modeling inter-phenotype dependencies, our multi-task classifier chain provides an accurate, non-invasive surrogate, enabling robust risk stratification using routine clinical data alone.

Key words: Hepatocellular carcinoma; Liver transplantation; Body composition; Multi-task learning

1 Introduction

Hepatocellular carcinoma (HCC) is one of the leading causes of cancer-related deaths worldwide, with liver transplantation (LT) offering a curative option for selected patients by removing both the tumor and the cirrhotic

✉ Di LU, zjuludi@zju.edu.cn
Xiao XU, zjxu@zju.edu.cn
* The two authors contributed equally to this work
Di LU, https://orcid.org/0000-0002-8724-3739
Xiao XU, https://orcid.org/0000-0002-2761-2811
Junjie YIN, https://orcid.org/0000-0003-2548-4139
Zhu LI, https://orcid.org/0009-0002-9440-4147.

Received Feb. 1, 2026; Revision accepted June 1, 2026;
Crosschecked xxx. xx, 20xx

liver (Reichelt et al., 2022; Eriksen and Møller, 2024; Lin et al., 2025; Sequeira et al., 2025). Despite refinements in selection criteria and perioperative management, post-transplant survival and recurrence remain highly heterogeneous among patients who otherwise meet similar tumor-based criteria, underscoring the importance of host-related factors in prognostic assessment (Rimassa et al, 2021; Reichelt et al, 2022).

Body composition has emerged as a key host determinant in chronic liver disease and transplantation (Reichelt, et al, 2022; Eriksen and Møller, 2024). Third lumbar vertebral level (L3) computed tomography (CT)-derived measures of skeletal muscle mass and radiodensity, as well as visceral and subcutaneous adipose tissue, can capture sarcopenia, myosteatosis and obesity-related phenotypes that are highly prevalent in LT candidates (Singal et al., 2016; Reichelt, et al., 2022; Eriksen and Møller, 2024). Prior studies, including one by our group, have linked sarcopenia and myosteatosis to poorer outcomes after LT for HCC and suggested that adverse body-composition traits may have additive effects on prognosis (Jang et al., 2021; Artru et al., 2022; Liu et al., 2024; Lu et al., 2024). However, most work has evaluated single or composite indices, without systematically dissecting the independent and cumulative prognostic impact of four distinct CT phenotypes—sarcopenia, myosteatosis, visceral obesity, and high subcutaneous adipose tissue density (SATD)—or comparing their contributions to early versus late post-transplant recurrence (Li et al, 2024).

In clinical practice, CT-based body-composition analysis remains constrained by the need for segmentation, dedicated software and technical expertise, whereas peri-transplant laboratory tests, anthropometric indices and demographic variables are routinely available (Eriksen and Møller, 2024; Mishra et al, 2025). A clinically useful CT-free surrogate should therefore not only reproduce individual CT-defined phenotypes but also reflect the biological relatedness among them. Sarcopenia, myosteatosis, visceral obesity, and high SATD are unlikely to behave as independent host traits in patients with chronic liver disease and HCC undergoing LT; rather, they represent interconnected manifestations along the muscle–adipose axis, shaped by overlapping hepatic dysfunction, nutritional depletion, metabolic dysregulation, systemic catabolic stress, and tissue-quality remodeling (Turimov Mustapoevich and Kim, 2023; Li, et al, 2024). This phenotypic relatedness is clinically relevant because a given abnormality may carry information about both its own tissue compartment and the broader host vulnerability state in which other abnormalities co-occur. Conventional single-task models, via estimating each phenotype separately, are therefore structurally limited in their ability to capture shared information and conditional dependencies across correlated body-composition traits. Multi-task learning (MTL) directly addresses this limitation by jointly modeling related phenotypes, allowing shared signals across tasks to inform phenotype-specific predictions (Harutyunyan et al., 2019; Pickhardt et al., 2021; Zhang et al., 2021; Delrieu et al., 2023; Fumagalli et al., 2024). However, such strategies have not been systematically evaluated for deriving routine-data surrogates of multiple CT-defined body-composition abnormalities in HCC patients undergoing LT (Pickhardt, et al, 2021).

In this double-center study of adults undergoing LT for HCC, we therefore aimed to: (i) evaluate the independent and additive prognostic impact of four L3 CT body-composition phenotypes—sarcopenia, myosteatosis, visceral obesity, and high SATD—on survival and early/late recurrence after LT; and (ii) develop and compare multi-task machine-learning models that use routinely available peri-transplant variables to provide surrogates for these CT-defined phenotypes.

2 Methods

2.1 Study design and population

In this double-center retrospective cohort study, consecutive adults (age ≥18 years) who underwent liver transplantation (LT) for hepatocellular carcinoma (HCC) between January 2015 and December 2020 at the First Affiliated Hospital, Zhejiang University School of Medicine, and Shulan (Hangzhou) Hospital (Hangzhou, China) were screened. The inclusion criteria were: (i) pathologically confirmed HCC as the primary indication; (ii) pre-transplant abdominal CT including the L3 level within 3 months; and (iii) complete peri-transplant clinical and follow-up data. Patients with re-transplantation, multi-organ transplantation, non-HCC indications,

missing or poor-quality L3 images, or missing key covariates, or missing follow-up were excluded. A total of 409 recipients met these criteria and were included (Fig. 1).

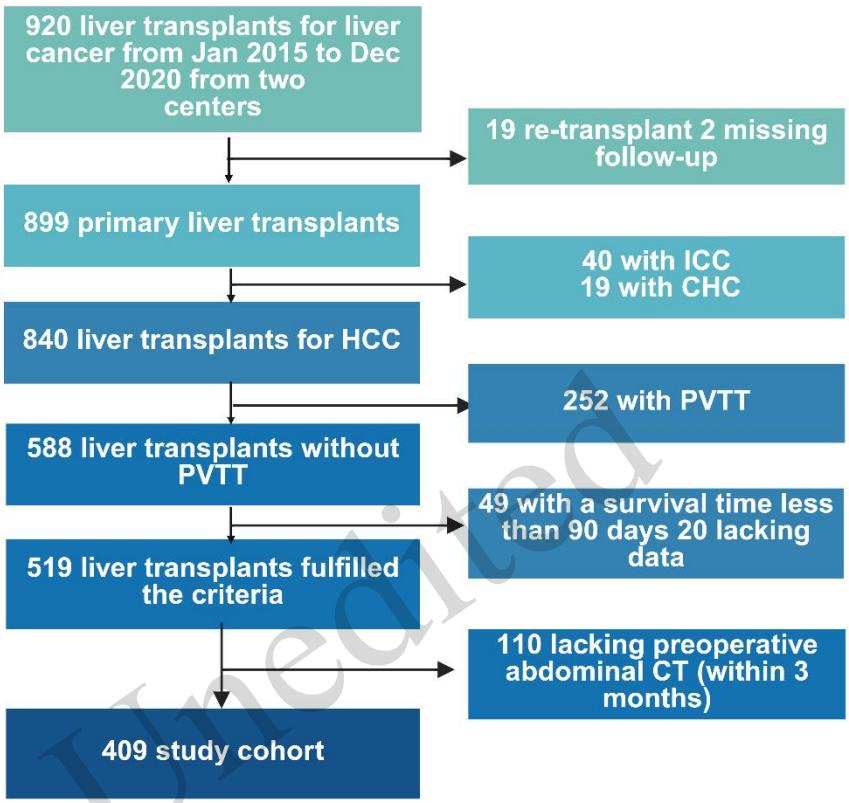

Figure 1. Flowchart of the patient selection procedure.
ICC, intrahepatic cholangiocarcinoma; CHC, combined hepatocellular-cholangiocarcinoma; PVTT, portal vein tumor thrombus; CT, computed tomography.

2.2 Clinical variables and follow-up

Baseline data were retrieved from prospectively maintained transplant databases and included age, sex, body mass index (BMI), aetiology of liver disease, diabetes and other comorbidities, Child–Pugh class, Model for End-stage Liver Disease (MELD) score, history of ascites, variceal bleeding and hepatic encephalopathy, and pre-transplant α-fetoprotein (AFP) (Lencioni and Llovet, 2010). Tumor size and number were obtained from pre-transplant imaging; and differentiation and micro- or macrovascular invasion were determined based on explant pathology. Donor and operative variables included donor age and type, ABO compatibility, cold ischemia time, operative time, and intraoperative blood loss.

Patients were followed up according to institutional protocols with regular clinical visits, laboratory tests and imaging. Recurrence was diagnosed by imaging and/or histology. Overall survival (OS) was defined as the time from LT to death from any cause or last follow-up. Recurrence-free survival (RFS) was defined as the time from LT to first tumor recurrence, death or last follow-up.

2.3 CT acquisition and body composition analysis

Pre-transplant abdominal CT scans were retrieved from radiology archives. Non-enhanced or portal-venous phase images acquired at 120 kV with 5-mm slice thickness on multidetector CT scanners were eligible. For

each patient, a single axial slice at the third lumbar vertebral level was selected using standard anatomical landmarks.

Skeletal muscle, visceral adipose tissue (VAT), and subcutaneous adipose tissue (SAT) were segmented by five trained observers blinded to the outcomes. Skeletal muscle area included psoas, paraspinal and abdominal wall muscles within −29 to +150 Hounsfield units (HU); VAT and SAT were segmented within −150 to −50 HU and −190 to −30 HU, respectively (Okubo et al., 2021). Cross-sectional areas (cm^2) were quantified and normalized to height squared to derive skeletal muscle index (SMI, cm^2/m^2). Mean skeletal muscle radiodensity (SMRA, HU), mean VAT attenuation and mean SAT attenuation were recorded (Liu et al, 2015; Van Der Werf et al, 2018). Extreme outliers were discarded, and the mean of three central measurements per tissue compartment was used for analysis.

2.4 Definition of body-composition abnormalities

Four CT-derived body-composition abnormalities were evaluated: sarcopenia, myosteatosis, visceral obesity, and high subcutaneous adipose tissue density (high SATD) (Cho et al, 2021). Sarcopenia was defined as an L3 skeletal muscle index (SMI) <43.75 cm^2/m^2 in men, based on a previously published Asian cutoff value (Yabusaki et al, 2016). Myosteatosis was defined as reduced skeletal muscle radiodensity (SMRA), reflecting the pathological fatty infiltration of skeletal muscle; an SMRA cutoff of 37.5 HU was used in this study. BMI-stratified myosteatosis thresholds were not applied because our previous study in a similar HCC liver transplantation cohort did not observe a significant association between BMI and myosteatosis (Lu, et al, 2024). Visceral obesity and high SATD were defined using cohort-specific tertile-derived cutoffs based on the distributions of visceral adipose tissue index (VATI) and mean SAT attenuation, respectively, in the present cohort. Each abnormality was coded as present or absent. The cumulative number of abnormalities (0–4) was summed and used to define three burden groups: ≤1, 2 and ≥3 abnormalities.

2.5 Multi-task machine-learning models for CT surrogates

To develop CT-free surrogates for the four abnormalities, we built multi-task models based on 37 routinely available peri-transplant variables (Table 3), including liver and kidney function tests, hematological and metabolic indices, BMI, age, sex, and other standard laboratory measures. Numeric predictors with missing values were imputed using cohort medians. The dataset was randomly split into a training set and a test set in a 70:30 ratio with a fixed seed, stratified by center.

2.6 Four multi-task strategies were compared:

(1) Shared-feature random forest. A global random forest was first fitted across all tasks to rank feature importance. The top 30 shared features were selected, and four task-specific random forests (500 trees each) were trained, one per abnormality.

(2) Classifier chain. A classifier-chain model was constructed using four task-specific random forests in a prespecified fixed task order: sarcopenia, myosteatosis, visceral obesity, and high SATD. This order followed the measurement hierarchy of the CT-derived phenotypes, moving from skeletal muscle quantity to skeletal muscle radiodensity and then from adipose-tissue quantity to adipose-tissue radiodensity. This order was implemented as a modeling sequence and was not intended to represent temporal progression or causal precedence among the four body-composition abnormalities. It was also not selected according to post hoc test-set performance or statistical feature-importance ranking. Because all four phenotypes were derived from the same pre-transplant CT examination, they were interpreted as contemporaneous measurements rather than sequential biological events. The first model used all 37 clinical features. For each subsequent task, the design matrix was augmented with the observed labels of preceding tasks in the training set and their predicted probabilities (“prev_” features) in the test set, thereby allowing the model to encode conditional dependence among co-occurring body-composition abnormalities. Each task-specific random forest included 500 trees.

(3) Multi-task neural network. All predictors were z-score standardized. For each abnormality, a separate feed-forward neural network was trained using the 'nnet' package (20 hidden units, weight decay 0.01, up to 500 iterations) and used to generate predicted probabilities on the test set.
(4) Stacking multi-task learning. For each abnormality, four base learners were trained: random forest (300 trees), extreme gradient boosting (XGBoost) (100 rounds, maximum depth 6, learning rate 0.1), logistic regression (generalized linear model with binomial link), and a small neural network (15 hidden units, decay 0.01, up to 300 iterations). Their predicted probabilities across all tasks formed 16 meta-features (4 models × 4 abnormalities), which were used to train a random forest meta-learner (300 trees) for each abnormality.
On the independent test set, we calculated the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity for each model–task combination, drew receiver operating characteristic (ROC) curves and performed decision-curve analysis to compare the net clinical benefit.

2.7 Statistical analysis

Continuous variables were presented as mean ± standard deviation or median (interquartile range) and compared using the Student's t-test or Mann–Whitney U test, as appropriate. Categorical variables were shown as counts (percentages) and compared using the χ^2 test or Fisher's exact test. Kaplan–Meier curves for OS and RFS were compared with the log-rank test. Cox proportional hazards models were used to identify independent predictors of OS and RFS; variables with P<0.10 in univariable analyses were considered for multivariable models, with checks for collinearity and proportional hazards assumptions. All analyses were performed using R version 4.4.2, and two-sided P values <0.05 were considered statistically significant.

2.8 Ethics considerations

Ethical approval for this double-center retrospective cohort study was granted by the China Liver Transplant Registry (CLTR) Scientific Committee (Approval No. 20220021). The study was conducted in accordance with national regulations, the ethical standards of the responsible committee on human experimentation, and the Declaration of Helsinki and its later amendments. Organ donation and transplantation procedures followed the guidelines set by the China Organ Donation Committee and the Organ Transplant Committee. All donor livers were obtained from donations after citizens' deaths, and no organs were procured from prisoners. Patient information was collected with approval from the ethics committee, and individual informed consent was obtained. Patient confidentiality and anonymity were strictly maintained throughout the study.

3 Results

3.1 Baseline characteristics

A total of 409 adult male patients who underwent LT for HCC between 2015 and 2020 were included. The median age was 54.0 years (interquartile range [IQR] 48.0–60.0), median BMI was 21.5 kg/m^2 (IQR 20.4–23.9), and 86.1% (352/409) had hepatitis B virus (HBV)-related liver disease. The median MELD score was 29.0 (IQR 11.0–39.0) and median AFP was 19.7 ng/mL (IQR 4.4–167.5). Overall, 46.5% of patients were beyond the Milan criteria, 20.8% had more than three tumors, 12.7% had a tumor size >8 cm, and 37.4% had microvascular invasion. During follow-up, 101 patients (24.7%) died and 128 (31.3%) developed recurrence; the median OS and RFS were 1150.0 days (IQR 808.0–1470.0) and 1108.0 days (IQR 716.0–1466.0), respectively. Liver, lung, bone and brain metastases occurred in 16.9%, 19.3%, 8.1%, and 2.0% of patients, respectively (Table 1).
Patients were classified by the cumulative number of abnormal CT-derived body-composition phenotypes (sarcopenia, myosteatosis, visceral obesity, and high SATD) into ≤1 (n=276), 2 (n=101) and ≥3 (n=32) abnormalities. Age increased across these groups (53.0, 56.0 and 57.0 years, P<0.001), while BMI and MELD were similar. The proportions of poor differentiation (23.2%, 30.7%, 43.8%; P=0.026), liver metastasis (12.0%, 21.8%, 43.8%; P<0.001), and lung metastasis (13.4%, 27.7%, 43.8%; P<0.001) rose markedly with

accumulating abnormalities. Deaths (17.4%, 26.7%, 81.2%) and recurrences (23.2%, 41.6%, 68.8%) also increased stepwise (*both P*<0.001), with corresponding declines in median OS (1201.5, 1112.0 and 636.5 days) and RFS (1180.0, 1006.0 and 337.5 days). Overall, sarcopenia, myosteatosis, visceral obesity, and high SATD were present in 24.9%, 30.3%, 34.0%, and 34.0% of patients, respectively, and clustered particularly in those with ≥3 abnormalities (Table 1).

Table 1 Patient characteristics for entire cohort

Characteristic	Overall	≤1 abnormality	2 abnormalities	≥3 abnormalities	*P*
n	409	276	101	32	
Age (years)	54.0 [48.0, 60.0]	53.0 [47.0, 58.0]	56.0 [50.0, 63.0]	57.0 [51.0, 62.2]	<0.001
BMI	21.5 [20.4, 23.9]	21.5 [20.5, 23.8]	21.8 [20.5, 23.9]	21.0 [20.0, 22.6]	0.356
MELD Score	29.0 [11.0, 39.0]	29.0 [11.0, 39.0]	30.0 [13.0, 38.0]	27.5 [13.2, 36.8]	0.995
HBV	352 (86.1)	244 (88.4)	82 (81.2)	26 (81.2)	0.144
AFP (ng/ml)	19.7 [4.4, 167.5]	19.1 [4.2, 153.4]	19.0 [4.3, 131.3]	80.8 [6.3, 822.9]	0.114
Beyond Milan Criteria	190 (46.5)	128 (46.4)	45 (44.6)	17 (53.1)	0.698
Tumor number (>3)	85 (20.8)	53 (19.2)	22 (21.8)	10 (31.2)	0.271
Tumor size (> 8cm)	52 (12.7)	28 (10.1)	17 (16.8)	7 (21.9)	0.061
Microvascular Invasion	153 (37.4)	98 (35.5)	38 (37.6)	17 (53.1)	0.149
Death	101 (24.7)	48 (17.4)	27 (26.7)	26 (81.2)	<0.001
OS (days)	1150.0 [808.0, 1470.0]	1201.5 [919.8, 1511.2]	1112.0 [759.0, 1451.0]	636.5 [369.8, 811.2]	<0.001
Recurrence	128 (31.3)	64 (23.2)	42 (41.6)	22 (68.8)	<0.001
RFS(days)	1108.0 [716.0, 1466.0]	1180.0 [831.2, 1507.0]	1006.0 [497.0, 1451.0]	337.5 [145.2, 720.2]	<0.001
Differentiation (Poor)	109 (26.7)	64 (23.2)	31 (30.7)	14 (43.8)	0.026
Liver metastasis	69 (16.9)	33 (12.0)	22 (21.8)	14 (43.8)	<0.001
Lung metastasis	79 (19.3)	37 (13.4)	28 (27.7)	14 (43.8)	<0.001
Bone metastasis	33 (8.1)	18 (6.5)	11 (10.9)	4 (12.5)	0.244
Brain metastasis	8 (2.0)	4 (1.4)	3 (3.0)	1 (3.1)	0.566
Sarcopenia	102 (24.9)	22 (8.0)	53 (52.5)	27 (84.4)	<0.001
Myosteatosis	124 (30.3)	29 (10.5)	64 (63.4)	31 (96.9)	<0.001
Visceral obesity	139 (34.0)	89 (32.2)	36 (35.6)	14 (43.8)	0.395
High SATD	139 (34.0)	67 (24.3)	46 (45.5)	26 (81.2)	<0.001

Data are presented as median [interquartile range] for continuous variables and n (%) for categorical variables. P values refer to comparisons among the ≤1 abnormality, 2 abnormalities, and ≥3 abnormalities groups; the overall cohort column was not included in the statistical comparison. Continuous variables were compared using the Kruskal–Wallis test, and categorical variables were compared using the χ^2 test or Fisher's exact test, as appropriate. BMI, body mass index; MELD, Model for End-stage Liver Disease; HBV, hepatitis B virus; AFP, alpha-fetoprotein; MVI, microvascular invasion; OS, overall survival; RFS, recurrence-free survival; SATD, subcutaneous adipose tissue density.

3.2 Association of individual body-composition abnormalities with post-transplant outcomes

Kaplan–Meier analysis showed that each of the four CT-derived body-composition abnormalities was associated with worse post-transplant prognosis (Fig. 2). For OS, patients with sarcopenia (*n*=102) had lower 1-, 3- and 5-year survival than those without (90.2%, 66.7% and 63.6% vs. 96.1%, 81.2% and 72.0%; log-rank

P=0.0035; Fig. 2b). Myosteatosis (n=124) showed a similar pattern, with 1-, 3- and 5-year OS of 90.2%, 67.4% and 60.2% vs. 96.5%, 82.0% and 74.2% in patients without myosteatosis (log-rank *P*=0.00047; Fig. 2d). Visceral obesity (n=139) was associated with broadly comparable short-term survival but worse long-term OS: 3- and 5-year OS declined from 81.6% and 73.3% in non-obese patients to 70.2% and 65.1% in those with visceral obesity (log-rank P=0.039; Fig. 2f). High SATD (n=139) identified a subgroup with particularly unfavorable OS, with 1-, 3- and 5-year survival of 89.1%, 70.1% and 59.6% compared with 97.4%, 81.5% and 76.5% in patients with normal SATD (log-rank P=0.0014; Fig. 2h).

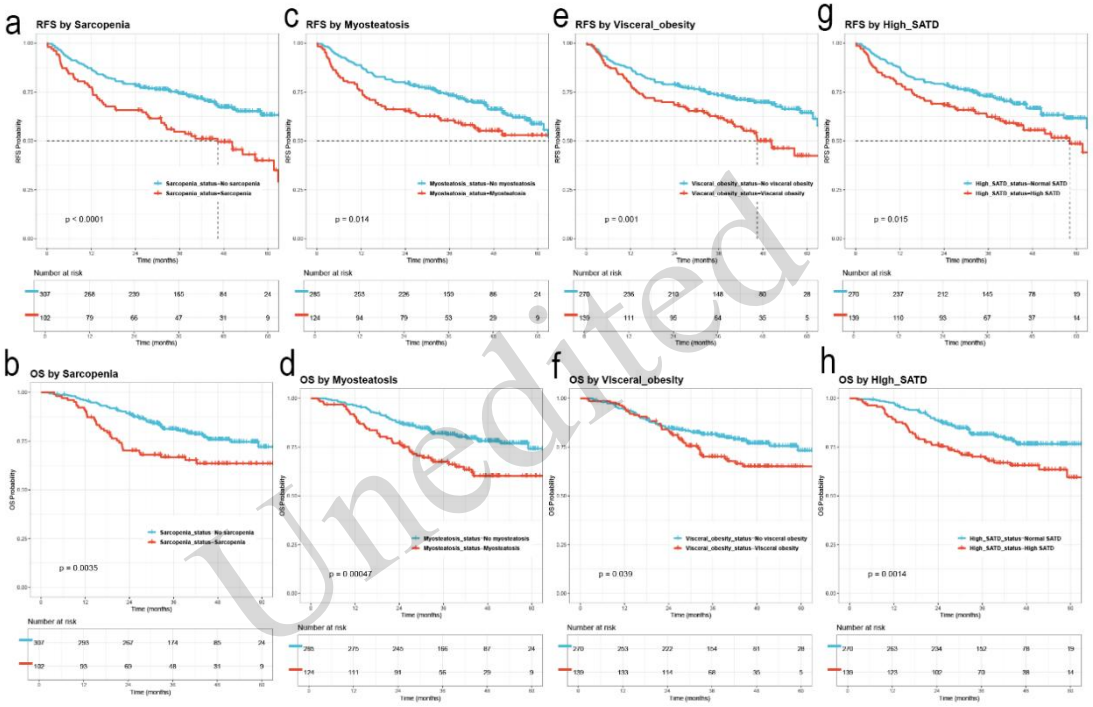

Figure 2. Kaplan-Meier survival analysis of body composition abnormalities in liver transplant recipients. (a) RFS stratified by sarcopenia status. (b) OS stratified by sarcopenia status. (c) RFS stratified by myosteatosis status. (d) OS stratified by myosteatosis status. (e) RFS stratified by visceral obesity status. (f) OS stratified by visceral obesity status. (g) RFS stratified by high SATD status. (h) OS stratified by high SATD status. Survival curves were estimated using the Kaplan–Meier method and compared using the two-sided log-rank test. The survival probability is shown on the y-axis, and follow-up time after liver transplantation is shown in months on the x-axis. The numbers below each panel indicate the number of patients at risk at each time point. Tick marks on the curves indicate censored observations. Dashed horizontal lines indicate a survival probability of 0.50, and dashed vertical lines indicate the corresponding median survival time when estimable. The figure was generated by the authors using R software. RFS, recurrence-free survival; OS, overall survival; HCC, hepatocellular carcinoma; LT, liver transplantation; SATD, subcutaneous adipose tissue density.

The same gradient was observed for RFS. Sarcopenic patients had 1-, 3- and 5-year RFS of 77.5%, 54.8% and 40.0%, compared with 87.0%, 74.3% and 63.3% in those without sarcopenia (log-rank P<0.0001; Fig. 2a). In patients with myosteatosis, 1-, 3- and 5-year RFS were 75.8%, 60.5% and 53.0%, versus 88.4%, 73.2% and 58.8% in those without myosteatosis (log-rank P=0.014; Fig. 2c). Visceral obesity was associated with clearly inferior RFS, with 1-, 3- and 5-year RFS of 79.1%, 61.7% and 42.4% versus 87.4%, 73.3% and 64.5% in patients without (log-rank P=0.001; Fig. 2e). High SATD also conferred worse RFS, with 1-, 3- and 5-year RFS of 79.1%, 62.2% and 48.6% compared with 87.4%, 73.0% and 61.6% in patients with normal SATD (log-rank P=0.015; Fig. 2g).

In Cox regression for RFS (Table 2), classical tumor-related factors—including being beyond the Milan criteria (hazard ratio (HR) 1.955, 95% confidence interval (CI) 1.277–2.992; P=0.002), pre-transplant AFP>400 ng/mL (HR 1.853, 95% CI 1.25–2.747; P=0.002) and microvascular invasion (HR 1.639, 95% CI 1.144–2.348; P=0.007)—remained independently associated with earlier recurrence. Importantly, all four body-composition abnormalities were also independent predictors of shorter RFS after adjustment: sarcopenia (HR 1.692, 95% CI 1.142–2.507; P=0.009), myosteatosis (HR 1.584, 95% CI 1.073–2.337; P=0.021), visceral obesity (HR 2.819, 95% CI 1.886–4.213; P<0.001), and high SATD (HR 2.314, 95% CI 1.572–3.406; P<0.001).

Table 2 Cox regression analysis of the variables of RFS in 409 HCC patients after liver transplantation

Characteristic	Univariable predictors of RFS		Multivariable predictors of RFS	
Variable	HR (95% CI)	P	HR (95% CI)	P
Age (years)	1.003 (0.985-1.022)	0.734	1.01 (0.99-1.03)	0.322
BMI	0.913 (0.86-0.971)	0.004	0.956 (0.897-1.02)	0.175
MELD Score	1.019 (1.006-1.033)	0.005	1.006 (0.991-1.021)	0.429
HBV	1.086 (0.671-1.758)	0.737	1.102 (0.66-1.838)	0.711
Beyond Milan Criteria	2.704 (1.926-3.795)	<0.001	1.955 (1.277-2.992)	**0.002**
Pre-transplant AFP > 400 (ng/ml)	2.141 (1.485-3.088)	<0.001	1.853 (1.25-2.747)	**0.002**
Tumor number (>3)	2.458 (1.743-3.466)	<0.001	1.196 (0.791-1.809)	0.396
Tumor size (> 8cm)	2.951 (2.005-4.345)	<0.001	1.132 (0.717-1.788)	0.594
Microvascular Invasion	2.218 (1.605-3.065)	<0.001	1.639 (1.144-2.348)	**0.007**
Differentiation (Poor)	1.517 (1.076-2.14)	0.017	1.177 (0.822-1.684)	0.373
Sarcopenia	1.969 (1.411-2.749)	<0.001	1.692 (1.142-2.507)	**0.009**
Myosteatosis	1.519 (1.085-2.126)	0.015	1.584 (1.073-2.337)	**0.021**
Visceral obesity	1.717 (1.24-2.378)	0.001	2.819 (1.886-4.213)	**<0.001**
High SATD	1.498 (1.08-2.078)	0.015	2.314 (1.572-3.406)	**<0.001**

Data are presented as hazard ratios (HRs) with 95% confidence intervals (CIs). HRs for continuous variables are expressed per one-unit increase. For categorical variables, the reference groups were absence of the corresponding condition, within Milan criteria, pre-transplant AFP ≤400 ng/mL, tumor number ≤3, tumor size ≤8 cm, and non-poor differentiation, as appropriate. P values were calculated using Cox proportional hazards regression. AFP, alpha-fetoprotein; BMI, body mass index; CI, confidence interval; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HR, hazard ratio; MELD, Model for End-stage Liver Disease; RFS, recurrence-free survival; SATD, subcutaneous adipose tissue density.

3.3 Cumulative impact of body-composition abnormalities on post-transplant outcomes

We next examined the cumulative impact of body-composition abnormalities by stratifying patients according to the number of concurrent abnormalities (≤1, 2, ≥3; Fig. 3a-3b). For OS, 1-, 3- and 5-year survival in patients with ≤1 abnormality were 97.1%, 84.7% and 77.3%, compared with 94.0%, 76.9% and 68.1% in those with 2 abnormalities and 75.0%, 19.6% and <20% in those with ≥3 abnormalities (pairwise log-rank $P=1.0\times10^{-4}$ for 2 vs ≤1; $P<2.0\times10^{-16}$ for ≥3 vs ≤1; $P=5.4\times10^{-7}$ for ≥3 vs 2; Fig. 3a).

A similar stepwise gradient was observed for RFS. Patients with ≤1 abnormality had 1-, 3- and 5-year RFS of 90.2%, 78.1% and 68.2%, which declined to 81.2%, 61.5% and 41.6% in those with 2 abnormalities and to 46.9%, 17.9% and <20% in those with ≥3 abnormalities (pairwise log-rank P=0.024 for 2 vs ≤1; $P<2.0\times10^{-16}$ for ≥3 vs ≤1; $P=2.8\times10^{-10}$ for ≥3 vs 2; Fig. 3b).

To account for imbalances in tumor-related factors, we performed propensity score matching using age, AFP, tumor size >8 cm, poor differentiation, and microvascular invasion. After matching, the baseline characteristics between groups were well balanced (Table S1), and the stepwise deterioration in both OS and RFS across increasing numbers of body-composition abnormalities persisted (Fig. 3c-3d), supporting a robust additive effect of body-composition abnormality burden on post-transplant prognosis.

In addition, to test whether the four body-composition abnormalities exerted independent effects rather than

acting as surrogates for one another, we fitted Cox models that included only the four abnormalities mutually adjusted (Fig. 3e-3f). In these body composition–only models, all four abnormalities remained significantly associated with both RFS (myosteatosis: P=0.016; visceral obesity, high SATD and sarcopenia: all P<0.001; Fig. 3e) and OS (myosteatosis: P=0.016; visceral obesity, high SATD and sarcopenia: all P<0.001; Fig. 3f), with hazard ratios consistently above 1.0 on the log scale. These analyses support that sarcopenia, myosteatosis, visceral obesity, and high SATD each contribute independently to post-transplant risk.

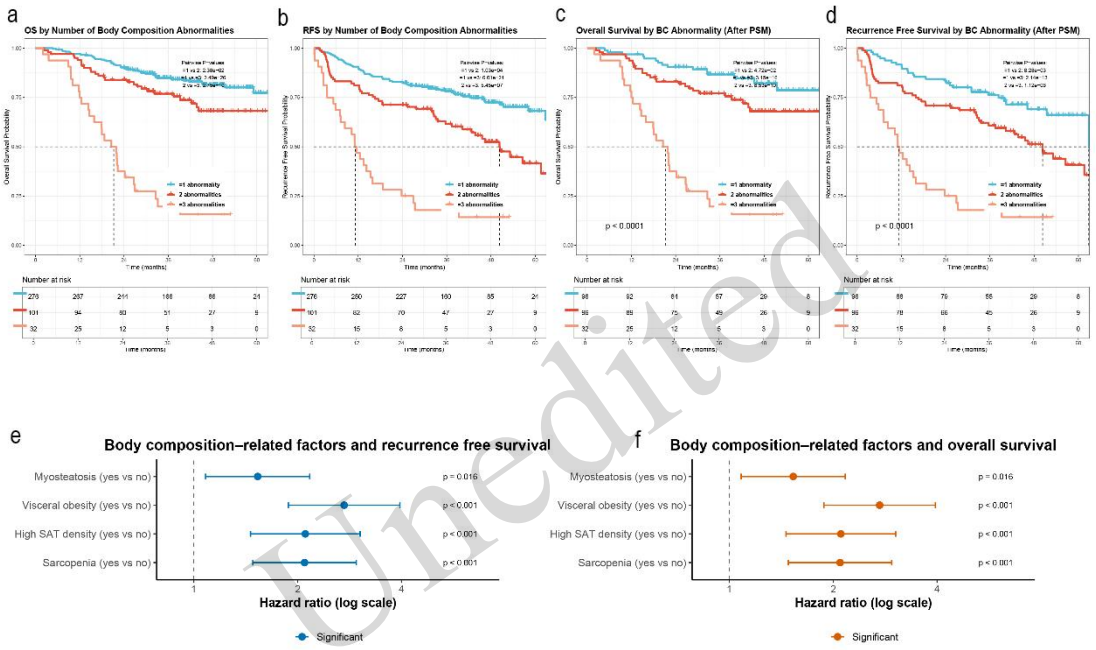

Figure 3. Cumulative impact of body composition abnormalities on post-transplant survival and recurrence free survival. (a) OS stratified by the number of concurrent body composition abnormalities (≤1, 2, and ≥3). (b) RFS stratified by the number of concurrent body composition abnormalities (≤1, 2, and ≥3). (c) OS stratified by cumulative abnormality burden after propensity score matching. (d) RFS stratified by cumulative abnormality burden after propensity score matching. (e) Forest plot of multivariable Cox regression analysis for RFS. (f) Forest plot of multivariable Cox regression analysis for OS. Survival curves were estimated using the Kaplan–Meier method and compared using the log-rank test. Numbers at risk are shown below each curve; tick marks indicate censored observations. Dots and horizontal lines indicate hazard ratios and 95% confidence intervals, respectively. OS, overall survival; RFS, recurrence-free survival; PSM, propensity score matching; HR, hazard ratio; CI, confidence interval.

3.4 Training and performance of multi-task surrogate models

We next compared four multi-task strategies for predicting the four CT-defined body-composition abnormalities—sarcopenia, myosteatosis, visceral obesity, and high SATD—using 37 routinely available peri-transplant variables (Table 3). The evaluated strategies included a shared-feature random forest, classifier chain, multi-task neural network, and stacking multi-task learning; the detailed model specifications are described in the Methods section. For the classifier-chain model, the four tasks were implemented in the following prespecified fixed order: sarcopenia, myosteatosis, visceral obesity, and high SATD. This order was chosen to mirror the measurement hierarchy of the CT-derived phenotypes, namely skeletal muscle quantity, skeletal muscle radiodensity, adipose-tissue quantity, and adipose-tissue radiodensity. This sequence was not

intended to imply temporal or causal progression, because all four abnormalities were assessed from the same pre-transplant CT examination. It was not determined by statistical feature importance or by post hoc test-set performance; rather, it was used as a consistent modeling sequence that allowed preceding task predictions to be incorporated into subsequent prediction tasks, thereby capturing conditional dependence among co-occurring body-composition abnormalities.

Table 3: Input Features for Multi-Task Learning Model

Variable	Overall	≤1 abnormality	2 abnormalities	≥3 abnormalities	p
n	409	276	101	32	
AGE (median [IQR])	54.00 [48.00, 60.00]	53.00 [47.00, 58.00]	56.00 [50.00, 63.00]	57.00 [51.00, 62.25]	<0.001
BMI (median [IQR])	21.51 [20.44, 23.88]	21.47 [20.52, 23.84]	21.77 [20.52, 23.94]	21.03 [19.99, 22.58]	0.356
AFP (median [IQR])	19.70 [4.40, 167.50]	19.10 [4.20, 153.40]	19.00 [4.30, 131.30]	80.85 [6.32, 822.85]	0.114
ALB (median [IQR])	37.10 [33.60, 40.80]	37.40 [33.60, 40.80]	37.00 [34.00, 41.30]	35.95 [32.48, 38.05]	0.37
AG (median [IQR])	1.30 [1.08, 1.60]	1.32 [1.10, 1.60]	1.27 [1.03, 1.60]	1.20 [1.00, 1.60]	0.305
ALT (median [IQR])	32.00 [22.00, 50.00]	32.00 [22.00, 48.00]	34.00 [23.00, 60.00]	29.00 [22.00, 41.00]	0.427
AST (median [IQR])	39.00 [28.00, 59.00]	39.00 [28.00, 56.00]	43.00 [27.00, 87.00]	39.00 [33.25, 56.00]	0.181
INR (median [IQR])	1.23 [1.09, 1.47]	1.23 [1.10, 1.47]	1.21 [1.07, 1.42]	1.27 [1.15, 1.45]	0.266
Mono (median [IQR])	0.08 [0.06, 0.11]	0.09 [0.07, 0.11]	0.08 [0.06, 0.11]	0.08 [0.07, 0.13]	0.187
Neu (median [IQR])	0.68 [0.59, 0.76]	0.67 [0.58, 0.75]	0.68 [0.60, 0.78]	0.71 [0.60, 0.80]	0.219
PLT (median [IQR])	71.00 [43.00, 115.00]	69.50 [41.00, 112.25]	77.00 [54.00, 131.00]	83.50 [54.00, 106.25]	0.082
WBC (median [IQR])	4.00 [2.70, 5.80]	3.90 [2.70, 5.50]	4.10 [2.70, 6.30]	3.85 [2.58, 6.15]	0.849
Lym (median [IQR])	21.60 [14.40, 28.50]	22.35 [14.67, 29.50]	20.80 [14.00, 27.20]	18.70 [11.45, 24.95]	0.154
Eos (median [IQR])	2.30 [1.10, 3.70]	2.30 [1.20, 3.73]	2.20 [1.10, 3.60]	1.55 [0.85, 3.75]	0.325
Bas (median [IQR])	0.40 [0.20, 0.60]	0.40 [0.30, 0.70]	0.40 [0.20, 0.60]	0.40 [0.20, 0.80]	0.353
Hb (median [IQR])	117.00 [98.00, 137.00]	122.50 [100.00, 139.00]	110.00 [89.00, 131.00]	111.00 [88.25, 119.00]	0.001
RBC (median [IQR])	3.73 [3.09, 4.40]	3.79 [3.24, 4.47]	3.55 [3.00, 4.24]	3.26 [2.88, 3.88]	0.004
MCV (median [IQR])	93.50 [89.40, 98.80]	93.35 [89.30, 98.40]	94.50 [90.50, 99.70]	93.85 [89.93, 97.80]	0.5
MCH (median [IQR])	31.80 [29.80, 33.50]	31.90 [29.50, 33.60]	31.70 [30.30, 33.30]	31.35 [29.27, 33.95]	0.838
MCHC (median [IQR])	336.00 [328.00, 346.00]	336.50 [329.00, 346.00]	335.00 [328.00, 343.00]	336.50 [318.00, 348.25]	0.409
RDW (median [IQR])	14.40 [13.20, 15.90]	14.30 [13.17, 15.62]	14.40 [13.20, 16.10]	15.80 [13.67, 17.72]	0.026
PCT (median [IQR])	0.08 [0.06, 0.13]	0.08 [0.05, 0.13]	0.09 [0.06, 0.14]	0.09 [0.08, 0.13]	0.331
GGT (median [IQR])	63.00 [34.00, 120.00]	62.00 [34.00, 112.25]	58.00 [33.00, 135.00]	89.00 [50.75, 134.00]	0.069
ALP (median [IQR])	108.00 [79.00, 146.00]	107.50 [79.75, 142.25]	108.00 [78.00, 155.00]	123.00 [82.50, 165.75]	0.443
T-bil (median [IQR])	23.70 [13.20, 47.00]	22.25 [13.00, 43.12]	23.00 [14.00, 58.00]	29.30 [19.32, 61.75]	0.101

D-bil (median [IQR])	10.70 [5.30, 23.10]	10.00 [5.00, 20.33]	11.00 [5.00, 32.10]	18.10 [7.60, 38.67]	0.025
ChE (median [IQR])	3523.00 [2391.00, 5358.00]	3772.00 [2541.75, 5568.25]	3279.00 [1953.00, 5317.00]	2730.50 [1753.75, 3895.50]	0.007
TBA (median [IQR])	32.00 [11.00, 78.30]	33.05 [9.97, 78.95]	30.70 [13.00, 64.70]	38.95 [13.58, 96.03]	0.861
GPDA (median [IQR])	84.00 [65.00, 106.00]	88.00 [66.75, 107.25]	75.00 [61.00, 103.00]	73.50 [58.50, 99.00]	0.106
AFU (median [IQR])	29.60 [23.00, 38.00]	30.00 [23.00, 38.00]	29.60 [21.90, 37.00]	28.95 [23.88, 35.02]	0.822
ADA (median [IQR])	20.00 [15.10, 26.00]	20.00 [15.00, 26.02]	20.00 [15.00, 25.00]	19.85 [15.65, 23.75]	0.845
TG (median [IQR])	0.89 [0.65, 1.27]	0.92 [0.67, 1.28]	0.82 [0.63, 1.18]	0.94 [0.64, 1.34]	0.311
TC (median [IQR])	3.03 [2.23, 3.96]	3.08 [2.32, 4.00]	3.11 [2.14, 4.00]	2.70 [2.18, 3.40]	0.212
HDL (median [IQR])	0.84 [0.56, 1.15]	0.90 [0.61, 1.21]	0.77 [0.44, 1.02]	0.58 [0.36, 0.80]	<0.001
LDL (median [IQR])	1.56 [1.09, 2.20]	1.58 [1.09, 2.20]	1.58 [1.04, 2.26]	1.41 [1.25, 1.80]	0.785
VLDL (median [IQR])	0.49 [0.30, 0.84]	0.46 [0.31, 0.82]	0.56 [0.28, 0.90]	0.54 [0.27, 0.76]	0.855
fastingblood (median [IQR])	5.69 [4.88, 7.20]	5.73 [4.89, 7.38]	5.83 [4.93, 6.87]	5.23 [4.66, 5.97]	0.126

Data are presented as median [interquartile range] for continuous variables. P values refer to comparisons among the ≤1 abnormality, 2 abnormalities, and ≥3 abnormalities groups; the overall cohort column was not included in the statistical comparison. Continuous variables were compared using the Kruskal–Wallis test. AFP, alpha-fetoprotein; ALB, albumin; AG, albumin-to-globulin ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase; INR, international normalized ratio; Mono, monocyte fraction; Neu, neutrophil fraction; Lym, lymphocyte percentage; Eos, eosinophil percentage; Bas, basophil percentage; WBC, white blood cell count; RBC, red blood cell count; Hb, hemoglobin; PLT, platelet count; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red cell distribution width; PCT, plateletcrit; GGT, gamma-glutamyl transferase; ALP, alkaline phosphatase; T-bil, total bilirubin; D-bil, direct bilirubin; ChE, cholinesterase; TBA, total bile acid; GPDA, glycylproline dipeptidyl aminopeptidase; AFU, alpha-L-fucosidase; ADA, adenosine deaminase; TG, triglyceride; TC, total cholesterol; HDL, high-density lipoprotein cholesterol; LDL, low-density lipoprotein cholesterol; VLDL, very-low-density lipoprotein cholesterol; BMI, body mass index; IQR, interquartile range.

On the held-out test set, all four strategies achieved clinically useful discrimination, whereas the classifier chain clearly outperformed the others across all four tasks (Fig. 4a). For high SATD, myosteatosis, sarcopenia and visceral obesity, the AUCs of the classifier chain were 0.789, 0.878, 0.882, and 0.865, respectively, compared with 0.703, 0.812, 0.811, and 0.759 for the shared-feature model, 0.638, 0.742, 0.786, and 0.661 for the multi-task neural network, and 0.678, 0.780, 0.780, and 0.711 for the stacking MTL model. The mean AUC across the four tasks reached 0.853 for the classifier chain, versus 0.771 for shared features, 0.749 for stacking MTL, and 0.707 for the multi-task neural network. The ROC curves and decision-curve analyses (Fig. 4b-4c) showed that classifier-chain models generally dominated the other approaches across a wide range of thresholds and yielded the greatest net clinical benefit at threshold probabilities of approximately 10–30%. The predicted risk distributions from the classifier chain demonstrated clear separation between patients with and without each abnormality (Fig. 4d). Moreover, in the joint plot of AUC versus mean net benefit (Fig. 4e), the classifier-chain models clustered in the upper-right region, confirming them as the preferred multi-task surrogate framework for approximating CT-derived body-composition abnormalities from routine peri-transplant data.

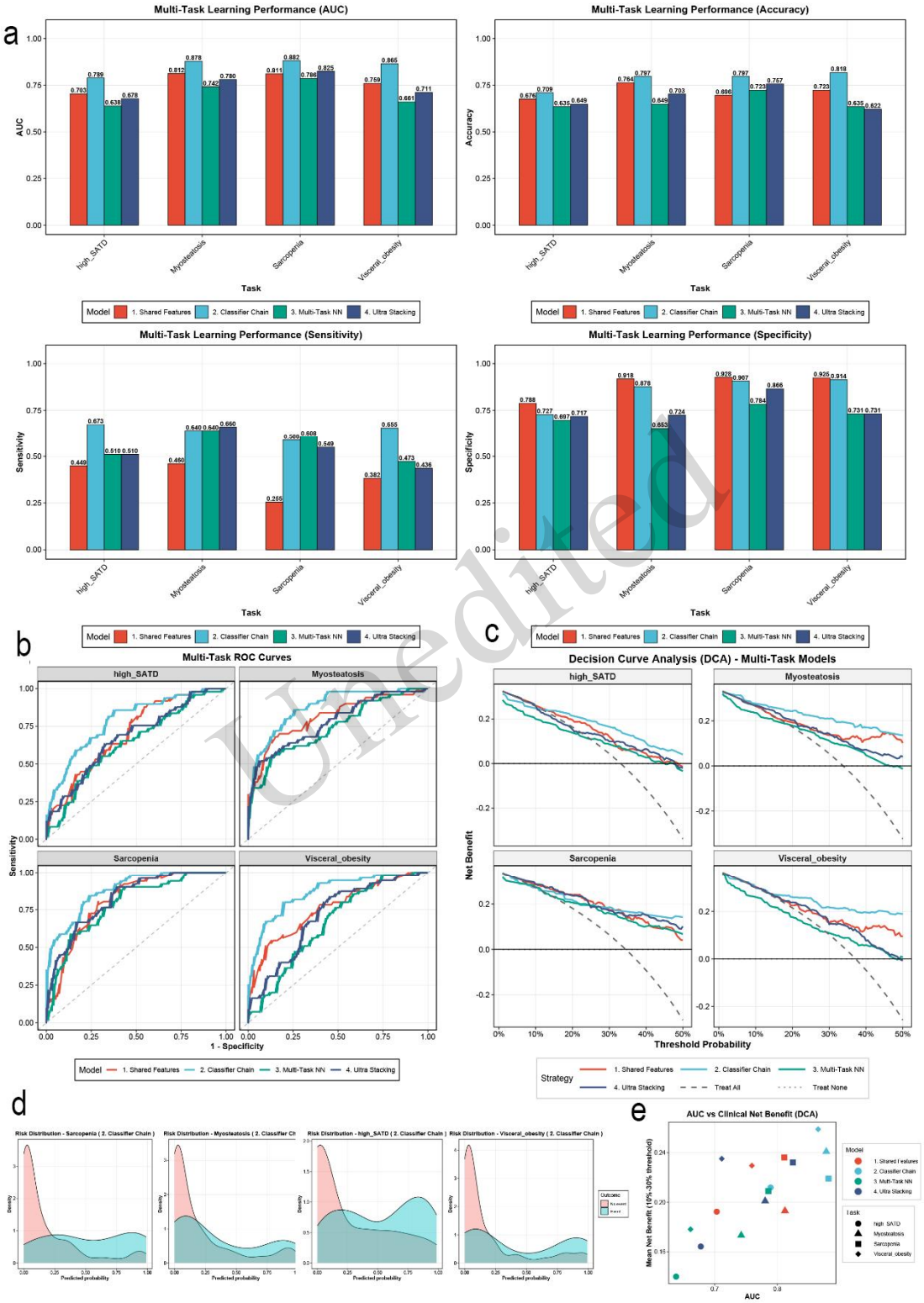

Figure 4. Performance evaluation and clinical utility of multi-task learning models for body composition abnormality prediction. (a) Test-set AUC, accuracy, sensitivity, and specificity of four multi-task models across the four body-composition abnormalities. (b) Test-set ROC curves of the four multi-task models for each body-composition abnormality. (c) Clinical decision curves of the four multi-task models for the four body-composition abnormalities in the test set. (d) Test-set predicted risk distributions of the classifier chain model for each body-composition abnormality, stratified by presence versus absence of the

abnormality. (e) AUC versus mean clinical net benefit (10%–30% threshold) for all model–task combinations. Colors denote models; shapes denote abnormalities. AUC, area under the receiver operating characteristic curve; ROC, receiver operating characteristic; NN, neural network; SATD, subcutaneous adipose tissue density.

3.5 Feature–task dependencies and correlations between body-composition abnormalities

To better understand how the classifier-chain model linked clinical features to each body-composition abnormality, we examined XGBoost-based feature importance and chain dependencies (Fig. 5). At the single-task level (Fig. 5a), sarcopenia prediction was mainly driven by markers of hepatobiliary injury and nutritional status, with GGT, HDL cholesterol, bile acid–related markers (e.g. TBA, GPDA), and red cell indices among the top contributors. Myosteatosis prediction was dominated by age and BMI, followed by ALP, AFP and the propagated sarcopenia probability (prev_Sarcopenia), indicating that muscle radiodensity shared predictive information with muscle quantity rather than being modeled as an isolated phenotype. Visceral obesity relied primarily on cholinesterase, bile acids and triglycerides (ChE, TBA, TG), together with albumin–globulin balance, consistent with a metabolic–hepatic signature. High SATD prediction was driven by coagulation and metabolic markers (INR, TG, creatinine, PCT) as well as the propagated visceral-obesity probability (prev_Visceral_obesity), suggesting a shared predictive structure between adipose-tissue burden and adipose-tissue radiodensity.

The cross-task heatmap (Fig. 5b) highlighted a set of shared predictors—such as ChE, INR, TG, TBA, HDL, BMI and age—that contributed meaningfully to several abnormalities, alongside more task-specific drivers (e.g., GGT for sarcopenia, ChE for visceral obesity, or INR/TG for high SATD). Chain-dependency analysis (Fig. 5c) showed that predictions from earlier tasks added non-negligible information to later tasks, with the strongest links observed from visceral obesity to high SATD and from sarcopenia to visceral obesity. The propagated “prev_” variables were therefore interpreted as conditional prediction features rather than markers of temporal disease progression. Because all four abnormalities were defined from the same pre-transplant CT examination, these variables were used to capture cross-phenotype dependence among co-occurring body-composition abnormalities, not to imply biological chronology.

Comorbidity-pattern analysis in the validation set (Fig. 5d) further illustrated how these abnormalities co-occurred in real patients: isolated visceral obesity and high SATD were most frequent, but combinations such as sarcopenia + high SATD, sarcopenia + myosteatosis, and myosteatosis + visceral obesity were also common, forming a spectrum from single to multiple concurrent abnormalities. Finally, the bump chart of feature-rank trajectories (Fig. 5e) distinguished shared from task-volatile predictors across the four classifier-chain tasks. Relatively stable trajectories indicated predictors shared across multiple body-composition abnormalities, whereas larger upward or downward rank shifts indicated task-specific predictive signals that may help distinguish muscle-centered from adipose-centered phenotypes. Together, these analyses indicate that the classifier chain captures both shared metabolic–hepatic signatures and distinct feature patterns for each body-composition abnormality, while it leverages correlations between abnormalities to improve prediction.

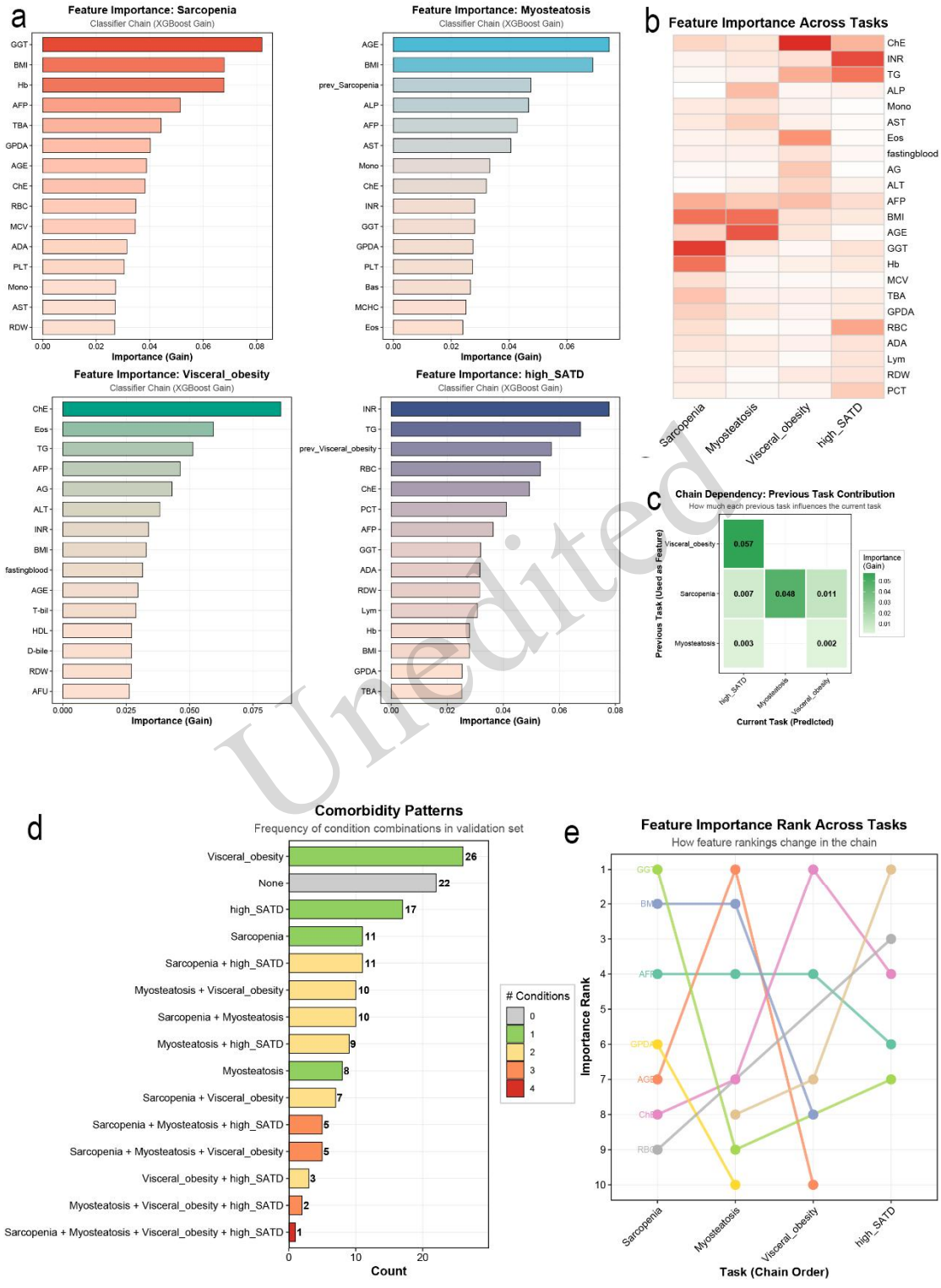

Figure 5. Feature importance and chain dependency analysis of the Classifier Chain model.
(a) Horizontal bar plots displaying the top 15 predictive features ranked by XGBoost gain for each body composition abnormality. Features prefixed with "prev_" denote chain-propagated predictions from preceding tasks. (b) Heatmap illustrating feature importance (gain) across all four prediction tasks. Rows represent features; columns represent tasks. Hierarchical clustering was applied to rows. (c) Heatmap showing the contribution of preceding task predictions to subsequent tasks in the Classifier Chain

architecture. Cell values represent XGBoost gain for chain-propagated features (prev_*). (d) Bar plot displaying the frequency of comorbidity patterns (co-occurrence of multiple abnormalities) in the test set, colored by the number of concurrent conditions. (e) Bump chart showing feature-importance rank trajectories across the four sequential classifier-chain tasks. Features with larger upward or downward rank shifts were interpreted as task-volatile predictors, whereas relatively stable trajectories indicated predictors shared across multiple body-composition abnormalities. ChE, cholinesterase; INR, international normalized ratio; TG, triglyceride; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; AFP, alpha-fetoprotein; BMI, body mass index; GGT, gamma-glutamyl transferase; Hb, hemoglobin; MCV, mean corpuscular volume; TBA, total bile acid; GPDA, glycylproline dipeptidyl aminopeptidase; RBC, red blood cell; RDW, red cell distribution width; PCT, plateletcrit; XGBoost, extreme gradient boosting.

4 Discussion

In this two-center HCC–LT cohort, four CT-defined body-composition abnormalities—sarcopenia, myosteatosis, visceral obesity and high SATD—were each associated with earlier recurrence and shorter overall survival. The risk gradients for both RFS and OS increased stepwise with the number of abnormalities, and the curves for patients with ≥3 abnormalities separated very early and remained widely divergent. When the four abnormalities were mutually adjusted, all remained significant, indicating that they capture partly distinct aspects of host vulnerability rather than a single frailty construct. These results support viewing body composition as a multidimensional prognostic domain that complements conventional tumor-based criteria (Chen et al., 2024; Eriksen and Møller, 2024; Lin, et al., 2025).

These findings are broadly consistent with previous HCC-LT studies showing that adverse body-composition phenotypes are associated with inferior post-transplant outcomes. Kim et al. reported that sarcopenia was associated with post-transplant tumor recurrence after living-donor LT in male patients with HCC beyond the Milan criteria (Kim et al, 2018). Grąt et al. further evaluated abdominal adipose and skeletal muscle compartments in HCC transplant recipients and identified excessive subcutaneous adipose tissue as a recurrence-related host factor (Grąt et al., 2019). More recently, Lu et al. showed that pre-transplant myosteatosis aggravated the adverse prognostic impact of muscle loss in male HCC-LT recipients, while Li et al. reported that CT-derived skeletal muscle parameters were associated with post-LT survival in HCC patients (Li, et al, 2024; Lu, et al, 2024). Together, these studies support the clinical relevance of host body composition beyond conventional tumor-related factors. The present study extends these earlier findings by evaluating sarcopenia, myosteatosis, visceral obesity, and high SATD in parallel, thereby capturing both muscle- and adipose-related dimensions of host vulnerability.

The definitions of these four CT-derived phenotypes should be interpreted according to both the clinical context and the availability of established standards. For sarcopenia, several CT-based cutoffs have been proposed in patients with chronic liver disease, liver transplantation candidates and cancer populations (Wang and Yang, 2025). Because the present cohort consisted predominantly of Asian male HCC liver transplantation recipients, we used a previously published Asian male SMI cutoff to define sarcopenia. For myosteatosis, previously reported muscle attenuation thresholds are often BMI-stratified and were mostly derived from general oncology cohorts (Martin et al, 2013). In this study, we used an SMRA cutoff of 37.5 HU to reflect the prognostic relevance of muscle radiodensity in this transplant cohort, and BMI-stratified thresholds were not applied because our previous study in a similar HCC liver transplantation cohort did not observe a significant association between BMI and myosteatosis (Lu, et al, 2024). For visceral obesity and high SATD, no universally accepted cutoff has been established in HCC liver transplantation recipients (Eriksen and Møller, 2024). Moreover, previous studies have set heterogeneous adipose tissue parameters, including VAT area, VATI, SATI, and adipose tissue attenuation (Van Dijk et al, 2017). We therefore defined visceral obesity and high SATD using cohort-specific tertile-derived cutoffs based on the corresponding CT-derived adipose parameters in the present cohort. These tertile-derived definitions were intended for within-cohort prognostic stratification and should not be regarded as universal diagnostic thresholds. Therefore, external validation is needed before these definitions can be applied to other populations.

The dose–response relationship observed in our cohort further refines this concept. Previous studies have generally treated body composition as an individual exposure, a selected compartment or a binary composite phenotype, whereas our abnormality-burden analysis captured the progressive accumulation of muscle and adipose abnormalities within the same host. The marked decline in 5-year OS from 77.3% in patients with ≤1 abnormality to 19.6% in those with ≥3 abnormalities, together with the parallel deterioration in RFS, suggests that these phenotypes are not interchangeable markers of a single frailty state but they rather appear to provide partially non-redundant information on host vulnerability. This interpretation is also supported by the mutually adjusted body-composition models, in which all four abnormalities remained associated with post-transplant outcomes. Therefore, the accumulation and co-occurrence of body-composition abnormalities may represent a correlated host-risk structure, rather than a simple summation of isolated CT findings.

Our multi-task learning framework adds a pragmatic layer on top of the CT findings. Using 37 routinely available peri-transplant features, the classifier-chain model achieved good discrimination for all four abnormalities and consistently outperformed shared-feature random forests, a multi-task neural network and a stacking ensemble. Feature-importance and chain-dependency analyses showed that the model leveraged both shared metabolic–hepatic markers (age, BMI, bile-acid and lipid indices, ChE, INR) and information propagated from preceding tasks, mirroring the clinical co-occurrence of these abnormalities. This suggests that the CT-derived body-composition risk can be approximated, at least in part, in settings where full image-based analysis is not feasible, and that different abnormalities may be driven by distinct but overlapping biological pathways.

Predictive modeling studies for sarcopenia or related body-composition impairment have begun to emerge, including clinical screening approaches and automated imaging-based methods for sarcopenia or body-composition assessment (Liu, et al, 2024). However, most of these approaches have addressed different populations, imaging workflows or prediction targets. Some models are designed to automate CT segmentation or quantify body-composition compartments directly from imaging, whereas others focus on sarcopenia screening in non-transplant settings. In contrast, our modeling objective was to approximate multiple baseline CT-defined body-composition phenotypes from routinely available peri-transplant variables in HCC-LT recipients. The classifier-chain framework was therefore designed not as an automated CT-segmentation tool or a direct survival-prediction model but as a routine-data surrogate for correlated CT-defined phenotypes. By incorporating preceding task predictions into subsequent tasks, the model explicitly considered inter-phenotype dependence among co-occurring muscle and adipose abnormalities.

The steep dose-response gradient observed across abnormality-burden groups also has potential implications for pre-transplant risk enrichment. Patients with multiple CT-defined abnormalities represented a small but clinically fragile subgroup, characterized by markedly reduced long-term OS and RFS. In settings where formal CT segmentation is unavailable or impractical, a routine-data surrogate may help identify candidates who are likely to harbor multi-phenotype body-composition impairment. Rather than functioning as a standalone treatment decision tool, such a model could help prioritize high-burden candidates for earlier multidisciplinary assessment, including nutritional evaluation, individualized protein–energy optimization, supervised exercise-based conditioning when clinically feasible, metabolic risk management, and closer peri-transplant surveillance. This interpretation remains hypothesis-generating. The present retrospective study does not establish that prehabilitation reverses CT-defined abnormalities or improves post-transplant oncological outcomes. Moreover, optimization strategies should not delay oncologically appropriate transplantation. Prospective interventional studies are required to determine whether targeted prehabilitation in patients with high body-composition burden can improve functional reserve and translate into better post-transplant outcomes (Duarte-Rojo et al, 2021).

Several limitations of this study should be acknowledged. First, it was retrospective and derived from two HBV-predominant centers, which limits immediate extrapolation to non-viral HCC cohorts. This issue is

particularly relevant for metabolic dysfunction-associated steatotic liver disease (MASLD)-related HCC, in which visceral adiposity, impaired muscle quality and cardiometabolic dysfunction are central host features. Although the muscle–adipose framework evaluated here is biologically relevant to MASLD, the performance and calibration of the routine-data surrogate in MASLD-enriched cohorts require dedicated external evaluation. Second, differences in BMI, diabetes prevalence, lipid and glucose profiles, inflammatory status, sex composition, and competing cardiometabolic risks may shift feature weights, calibration, and optimal cut-offs. External validation and, if necessary, recalibration in MASLD-enriched and other non-HBV transplant cohorts are therefore required before broader application (Czarnecka et al, 2023). Third, body composition was assessed at a single pre-transplant L3 slice (Brown et al, 2023); longitudinal changes and more granular radiomics or volumetric metrics were not captured (Brown, et al., 2023; Huang et al., 2025; Nowak et al., 2025). Fourth, the tertile-derived definitions of visceral obesity and high SATD were cohort-specific and were used for prognostic stratification rather than universal diagnosis; therefore, these adipose-related thresholds may not directly generalise to other populations (Shi et al., 2023; Müller et al., 2024; Li et al., 2025). In addition, the surrogate models were internally validated only and were trained to predict CT-defined abnormalities rather than time-to-event outcomes.

Finally, the classifier-chain model was also implemented using a fixed task order, and its performance may be influenced by task sequencing. This order should therefore be regarded as an implementation-specific modeling choice rather than evidence of a temporal biological pathway. Future studies should evaluate alternative chain orders or ensemble classifier-chain strategies to further assess model robustness.

Despite these constraints, the work collectively supports a simple message: the accumulation of muscle and fat abnormalities before LT matters, and it can be quantified and approximated using information that clinicians already collect. Future studies should test these models prospectively, examine whether modifying body composition or its surrogate scores alters post-transplant risk, and explore integration with dynamic tumor and immunological markers (Ge et al, 2025). Such efforts may help move HCC–LT decision-making from "tumor-only" criteria towards truly integrated host–tumor profiling and more personalized long-term care (Avramidou et al, 2025).

Data availability statement

The datasets and analysis code used in this study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by the Key Research & Development Plan of Zhejiang Province (No. 2024C03051), the Major Research Plan of the National Natural Science Foundation of China (No. 92159202), the National Key Research and Development Program of China (No. 2021YFA1100500), the Construction Fund of Key Medical Disciplines of Hangzhou (No. C020200093), and the Medical and Health Science and Technology Development Plan of Zhejiang Province (No. 2025KY1086). The authors thank all members of the Key Laboratory of Integrated Oncology and Intelligent Medicine of Zhejiang Province for their valuable advice and discussion during the preparation of this manuscript.

Author contributions

Junjie YIN and Zhu LI performed the data analysis, software implementation, model visualization, and manuscript drafting. Jinyan CHEN, Qingyang QUE, Huigang LI, Ruijie ZHAO, Jianyong ZHUO, Jiaqi ZHENG, and Zuyuan LIN contributed to clinical data collection and data verification. Zhu LI, Jinyan CHEN,

Ruijie ZHAO, Zuyuan LIN, Chiyu HE, Yifan GU, Chenghao CAO, and Yuhang LI performed image assessment and visualization. Wei SHEN, Peiru ZHANG, Rongsen WANG, Shengjun XU, Xuyong WEI, and Shusen ZHENG contributed to interpretation of results and critical revision of the manuscript. Di LU and Xiao XU conceived and supervised the study, contributed to the study design, interpretation of results, critical revision of the manuscript, and funding acquisition. All authors have read and approved the final manuscript, 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

Junjie YIN, Zhu LI, Jinyan CHEN, Qingyang QUE, Huigang LI, Ruijie ZHAO, Jianyong ZHUO, Chiyu HE, Wei SHEN, Peiru ZHANG, Yifan GU, Chenghao CAO, Jiaqi ZHENG, Yuhang LI, Rongsen WANG, Zuyuan LIN, Shengjun XU, Xuyong WEI, Shusen ZHENG, Di LU, and Xiao XU declare that they have no conflicts of interest.

Ethical approval for this retrospective cohort study was granted by the CLTR Scientific Committee (Approval No. 20220021). All procedures followed were in accordance with national regulations, the ethical standards of the responsible committee on human experimentation, and the Declaration of Helsinki and its later amendments. Organ donation and transplantation procedures followed the guidelines set by the China Organ Donation Committee and the Organ Transplant Committee. All donor livers were obtained from donations after citizens' deaths, and no organs were procured from prisoners. Patient information was collected with approval from the ethics committee, and individual informed consent was obtained. Patient confidentiality and anonymity were strictly maintained throughout the study.

Consent for publication was not applicable because no identifiable patient information is included in this article.

Declaration on the use of generative AI tools

No generative AI tools were used in the preparation of this manuscript.

References

Artru F, Le Goffic C, Pageaux GP, et al., 2022. Sarcopenia should be evaluated in patients with acute-on-chronic liver failure and candidates for liver transplantation. *J Hepatol*, 76(4):983-985. https://doi.org/10.1016/j.jhep.2021.09.004

Avramidou E, Todorov D, Katsanos G, et al., 2025. Ai innovations in liver transplantation: From big data to better outcomes. *Livers*, 5(1):14.

Brown S, Richardson B, Bouquet E, et al., 2023. Cirrhosis-related sarcopenia may not resolve after liver transplantation. *JHEP Rep*, 5(11):100881. https://doi.org/10.1016/j.jhepr.2023.100881

Chen T, Li S, Deng D, et al., 2024. Key role of interferon regulatory factor 1 (irf-1) in regulating liver disease: Progress and outlook. *J Zhejiang Univ Sci B*, 25(6):451-470. https://doi.org/10.1631/jzus.B2300159

Cho YS, Lee HY, Jeong JY, et al., 2021. Computed tomography-determined body composition abnormalities usefully predict long-term mortality in patients with liver cirrhosis. *J Comput Assist Tomogr*, 45(5):684-690. https://doi.org/10.1097/rct.0000000000001207

Czarnecka K, Czarnecka P, Tronina O, et al., 2023. Metabolic profile of liver transplant recipients and determinants of their body fat distribution. *Obesities*, 3(3):225-242.

Delrieu L, Blanc D, Bouhamama A, et al., 2023. Automatic deep learning method for third lumbar selection and body composition evaluation on ct scans of cancer patients. *Front Nucl Med*, 3:1292676. https://doi.org/10.3389/fnume.2023.1292676

Duarte-Rojo A, Bloomer PM, Rogers RJ, et al., 2021. Introducing el-fit (exercise and liver fitness): A smartphone app to prehabilitate and monitor liver transplant candidates. *Liver Transpl*, 27(4):502-512. https://doi.org/10.1002/lt.25950

Eriksen CS, Møller S, 2024. Quantitative assessment of body composition in cirrhosis. *Diagnostics (Basel)*, 14(19):2191. https://doi.org/10.3390/diagnostics14192191

Fumagalli IA, Le ST, Peng PD, et al., 2024. Automated ct analysis of body composition as a frailty biomarker in abdominal surgery. *JAMA Surgery*, 159(7):766-774. https://doi.org/10.1001/jamasurg.2024.0628

Ge T, Chen Y, Pang L, et al., 2025. Roles of panoptosis and related genes in acute liver failure: Neoteric insight from

bioinformatics analysis and animal experiment verification. *J Zhejiang Univ Sci B*, 26(4):353-370. https://doi.org/10.1631/jzus.B2300678

Grąt K, Pacho R, Grąt M, et al., 2019. Impact of body composition on the risk of hepatocellular carcinoma recurrence after liver transplantation. *Journal of Clinical Medicine*, 8(10):1672.

Harutyunyan H, Khachatrian H, Kale DC, et al., 2019. Multitask learning and benchmarking with clinical time series data. *Sci Data*, 6(1):96. https://doi.org/10.1038/s41597-019-0103-9

Huang Y, Cun H, Mou Z, et al., 2025. Multiparameter body composition analysis on chest ct predicts clinical outcomes in resectable non-small cell lung cancer. *Insights into Imaging*, 16(1):32. https://doi.org/10.1186/s13244-025-01910-0

Jang HY, Choi GH, Hwang SH, et al., 2021. Sarcopenia and visceral adiposity predict poor overall survival in hepatocellular carcinoma patients after curative hepatic resection. *Translational Cancer Research*, 10(2):854-866.

Kim YR, Park S, Han S, et al., 2018. Sarcopenia as a predictor of post-transplant tumor recurrence after living donor liver transplantation for hepatocellular carcinoma beyond the milan criteria. *Scientific Reports*, 8(1):7157. https://doi.org/10.1038/s41598-018-25628-w

Lencioni R, Llovet JM, 2010. Modified recist (mrecist) assessment for hepatocellular carcinoma. *Semin Liver Dis*, 30(1):52-60. https://doi.org/10.1055/s-0030-1247132

Li W, Zhu H, Dong HZ, et al., 2025. Impact of body composition parameters, age, and tumor staging on gastric cancer prognosis. *Eur J Cancer Prev*, 34(3):267-275. https://doi.org/10.1097/cej.0000000000000917

Li Z, Zhao Y, Xie Y, et al., 2024. Impact of ct-relevant skeletal muscle parameters on post-liver transplantation survival in patients with hepatocellular carcinoma. *Hepatol Int*, 18(5):1516-1527. https://doi.org/10.1007/s12072-024-10708-z

Lin ZY, Lu D, Wu SD, et al., 2025. Body composition predicts prognosis of patients with hepatocellular carcinoma after liver transplantation: A multicenter study. *Hepatobiliary Pancreat Dis Int*, https://doi.org/10.1016/j.hbpd.2025.11.002

Liu B, Teng F, Fu H, et al., 2015. Excessive intraoperative blood loss independently predicts recurrence of hepatocellular carcinoma after liver transplantation. *BMC Gastroenterology*, 15(1):138. https://doi.org/10.1186/s12876-015-0364-5

Liu Z, Wu Y, Khan AA, et al., 2024. Deep learning-based radiomics allows for a more accurate assessment of sarcopenia as a prognostic factor in hepatocellular carcinoma. *J Zhejiang Univ Sci B*, 25(1):83-90. https://doi.org/10.1631/jzus.B2300363

Lu D, Hu Z, Chen H, et al., 2024. Myosteatosis and muscle loss impact liver transplant outcomes in male patients with hepatocellular carcinoma. *J Cachexia Sarcopenia Muscle*, 15(5):2071-2083. https://doi.org/10.1002/jcsm.13554

Martin L, Birdsell L, Macdonald N, et al., 2013. Cancer cachexia in the age of obesity: Skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol*, 31(12):1539-1547. https://doi.org/10.1200/jco.2012.45.2722

Mishra A, Thapa R, Bigam K, et al., 2025. Body composition assessment provides prognostic information in patients with cancer affected by chronic graft vs. Host disease. *J Cachexia Sarcopenia Muscle*, 16(2):e13759. https://doi.org/10.1002/jcsm.13759

Müller L, Mähringer-Kunz A, Auer TA, et al., 2024. Ai-derived body composition parameters as prognostic factors in patients with hcc undergoing tace in a multicenter study. *JHEP Reports*, 6(8):101125. https://doi.org/10.1016/j.jhepr.2024.101125

Nowak M, Núñez L, Hill CE, et al., 2025. Automated volumetric mri quantification of body fat and skeletal muscle in the uk biobank: Technical validation against expert manual segmentations and comparison with single-slice techniques. *Abdominal Radiology*, https://doi.org/10.1007/s00261-025-05170-w

Okubo S, Shindoh J, Kobayashi Y, et al., 2021. Adipose tissue distribution predicts prognosis of cirrhotic patients undergoing hepatectomy for hepatocellular carcinoma. *Ann Surg Oncol*, 28(11):6738-6746. https://doi.org/10.1245/s10434-021-09658-9

Pickhardt PJ, Summers RM, Garrett JW, 2021. Automated ct-based body composition analysis: A golden opportunity. *Korean J Radiol*, 22(12):1934-1937. https://doi.org/10.3348/kjr.2021.0775

Reichelt S, Pratschke J, Engelmann C, et al., 2022. Body composition and the skeletal muscle compartment in liver transplantation: Turning challenges into opportunities. *American Journal of Transplantation*, 22(8):1943-1957. https://doi.org/10.1111/ajt.17089

Rimassa L, Personeni N, Czauderna C, et al., 2021. Systemic treatment of hcc in special populations. *J Hepatol*, 74(4):931-943. https://doi.org/10.1016/j.jhep.2020.11.026

Sequeira LM, Ozturk NB, Sierra L, et al., 2025. Hepatocellular carcinoma and the role of liver transplantation: An update and review. *J Clin Transl Hepatol*, 13(4):327-338. https://doi.org/10.14218/jcth.2024.00432

Shi S, Zhao Y-X, Fan J-L, et al., 2023. Development and external validation of a nomogram including body composition parameters for predicting early recurrence of hepatocellular carcinoma after hepatectomy. *Academic Radiology*, 30(12):2940-2953. https://doi.org/10.1016/j.acra.2023.05.022

Singal AG, Zhang P, Waljee AK, et al., 2016. Body composition features predict overall survival in patients with hepatocellular carcinoma. *Clin Transl Gastroenterol*, 7(5):e172. https://doi.org/10.1038/ctg.2016.31

Turimov Mustapoevich D, Kim W, 2023. Machine learning applications in sarcopenia detection and management: A comprehensive survey. *Healthcare*, 11(18):2483.

Van Der Werf A, Langius JaE, De Van Der Schueren MaE, et al., 2018. Percentiles for skeletal muscle index, area and radiation attenuation based on computed tomography imaging in a healthy caucasian population. *Eur J Clin Nutr*, 72(2):288-296. https://doi.org/10.1038/s41430-017-0034-5

Van Dijk DP, Bakens MJ, Coolsen MM, et al., 2017. Low skeletal muscle radiation attenuation and visceral adiposity are associated with overall survival and surgical site infections in patients with pancreatic cancer. *J Cachexia Sarcopenia Muscle*, 8(2):317-326. https://doi.org/10.1002/jcsm.12155

Wang Z, Yang B, 2025. Pro-aging metabolic reprogramming: A unified theory of aging. *Engineering*, 44:37-43. https://doi.org/ 10.1016/j.eng.2024.09.010

Yabusaki N, Fujii T, Yamada S, et al., 2016. Adverse impact of low skeletal muscle index on the prognosis of hepatocellular carcinoma after hepatic resection. *International Journal of Surgery*, 30:136-142. https://doi.org/ 10.1016/j.ijsu.2016.04.049

Zhang X, Xing Y, Sun K, et al., 2021. Omiembed: A unified multi-task deep learning framework for multi-omics data. *Cancers (Basel)*, 13(12) https://doi.org/10.3390/cancers13123047

Supplementary information

Table S1 Patient characteristics for the entire cohort after propensity score matching.