Explainable deep learning model WAL-net for individualized assessment of potentially reversible malnutrition in patients with cancer: a multicenter cohort study

Liangyu Yin[1*]; Ning Tong[2]; Na Li[2]; Jie Liu[2]; Wei Li[3]; Jiuwei Cui[3]; Zengqing Guo[4]; Qinghua Yao[5]; Fuxiang Zhou[6]; Ming Liu[7]; Zhikang Chen[8]; Huiqing Yu[9]; Tao Li[10]; Zengning Li[11]; Pingping Jia[12]; Chunhua Song[13]; Hongxia Xu[2*], and Hanping Shi[12, 14, 15*]

[1]Department of Nephrology, Chongqing Key Laboratory of Prevention and Treatment of Kidney Disease, Chongqing Clinical Research Center of Kidney and Urology Diseases, Xinqiao Hospital, Army Medical University (Third Military Medical University), Chongqing 400037, China

[2]Department of Clinical Nutrition, Daping Hospital, Army Medical University (Third Military Medical University), Chongqing 400042, China

[3]Cancer Center of the First Hospital of Jilin University, Changchun, Jilin 130021, China

[4]Department of Medical Oncology, Fujian Cancer Hospital, Fujian Medical University Cancer Hospital, Fuzhou, Fujian 350014, China

[5]Department of Integrated Chinese and Western Medicine, Cancer Hospital of the University of Chinese Academy of Science (Zhejiang Cancer Hospital), Hangzhou, Zhejiang 310022, China

[6]Department of Oncology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei 430071, China

[7]Department of Colorectal Surgery, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang 150001, China

This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI

10.1017/S000711452510384X

The British Journal of Nutrition is published by Cambridge University Press on behalf of The Nutrition Society

[8]Department of Colorectal and Anal Surgery, Xiangya Hospital of Central South University, Changsha, Hunan 410008, China

[9]Department of Palliative Care and Department of Geriatric Oncology, Chongqing University Cancer Hospital, Chongqing 400030, China

[10]Department of Radiation Oncology, Sichuan Cancer Hospital and Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan 610041, China

[11]Department of Clinical Nutrition, The First Hospital of Hebei Medical University, Shijiazhuang, Hebei 050031, China

[12]Department of Gastrointestinal Surgery and Department of Clinical Nutrition, Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China

[13]Department of Epidemiology, College of Public Health, Zhengzhou University, Zhengzhou, Henan 450001, China

[14]Key Laboratory of Cancer FSMP for State Market Regulation, Beijing 100038, China

[15]Lead Contact and Principal Investigator of the Investigation on Nutrition Status and Its Clinical Outcome of Common Cancers (INSCOC) Project, China

***Correspondence authors: Liangyu Yin**, MD, PhD, Department of Nephrology, Xinqiao Hospital, Army Medical University (Third Military Medical University), Xinqiao Main Street No.83, Shapingba District, Chongqing 400037, China. **Email:** liangyuyin1988@tmmu.edu.cn, liangyuyin1988@qq.com or **Hongxia Xu**, MD, PhD, Department of Clinical Nutrition, Daping Hospital, Army Medical University (Third Military Medical University), Chongqing 400042, China. Tel: +86 23 6872 9641, **Email:** hx_xu2015@163.com or **Hanping Shi**, MD, PhD, Department of Gastrointestinal Surgery and Department of Clinical Nutrition, Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China; Key Laboratory of Cancer FSMP for State Market Regulation Beijing 100038, China. **Email:** shihp@ccmu.edu.cn

Short title: Reversible malnutrition in cancer

ABSTRACT

Persistent malnutrition is associated with poor clinical outcomes in cancer. However, assessing its reversibility can be challenging. The present study aimed to utilize machine learning (ML) to predict reversible malnutrition (RM) in patients with cancer. A multicenter cohort study including hospitalized oncology patients. Malnutrition was diagnosed using an international consensus. RM was defined as a positive diagnosis of malnutrition upon patient admission which turned negative one month later. Time-series data on body weight and skeletal muscle were modeled using a long short-term memory (LSTM) architecture to predict RM. The model was named as WAL-net, and its performance, explainability, clinical relevance and generalizability were evaluated. We investigated 4254 patients with cancer-associated malnutrition (discovery set=2977, test set=1277). There were 2783 men and 1471 women (median age=61 years). RM was identified in 754 (17.7%) patients. RM/non-RM groups showed distinct patterns of weight and muscle dynamics, and RM was negatively correlated with the progressive stages of cancer cachexia ($r=-0.340$, $P<0.001$). WAL-net was the state-of-the-art model among all ML algorithms evaluated, demonstrating favorable performance to predict RM in the test set (AUC=0.924, 95%CI=0.904-0.944) and an external validation set (n=798, AUC=0.909, 95%CI=0.876-0.943). Model-predicted RM using baseline information was associated with lower future risks of underweight, sarcopenia, performance status decline and progression of malnutrition (all $P<0.05$). This study presents an explainable deep learning model, the WAL-net, for early identification of RM in patients with cancer. These findings might help the management of cancer-associated malnutrition to optimize patient outcomes in multidisciplinary cancer care.

Keywords: Malnutrition; GLIM; Cancer; Machine learning; Recurrent neural network

Introduction

Malnutrition is a major global public health problem affecting more than one billion of the world's population [1, 2]. It is also a highly prevalent disorder in oncology practice [3-5] with a prevalence of 21% to 72% [6-8]. Malnutrition can impede the efficacy and safety of anticancer therapies [3], increase healthcare resource utilization [4] and lead to multiple adverse outcomes [9, 10]. In fact, an estimated 10%-20% of cancer deaths are solely ascribable to malnutrition [11]. However, due to a lack of proper diagnostic techniques and sufficient attention, cancer-associated malnutrition remains largely underreported [12], misclassified [13] or left untreated [14]. Therefore, active diagnosis, surveillance and intervention of malnutrition are imperative in all cancer patients to minimize or reverse its negative impact on patient outcomes [3, 4, 11, 15].

The practical criteria used to diagnose malnutrition vary across different institutions [6, 12-14], making it difficult to implement a universally standardized framework for patient management. To address this challenge, an international consensus-based conceptual framework was proposed in 2019 [16], the Global Leadership Initiative on Malnutrition (GLIM), for diagnosing malnutrition in adults in clinical settings. Briefly, for patients who are screened positive for nutritional risk, at least one phenotypic criterion and one etiological criterion are required to diagnose malnutrition [16]. This new framework effectively incorporates the evolving understanding and current evidence on malnutrition, making it a promising tool with global acceptance potential [5, 10, 17-21].

Since its release, the GLIM framework has proven to be valuable in diagnosing malnutrition and predicting clinical outcomes across a wide range of diseases [9, 10, 22]. In the context of cancer, evidence from our research [20, 21, 23, 24] and other institutions [10, 17-19, 25, 26] consistently demonstrates that GLIM-defined malnutrition is associated with worse clinical outcomes, including postoperative complications [10, 17, 18, 25, 26], length of hospital stay [23, 26], in-hospital mortality [26], thirty-day mortality [19], disease-free survival [18, 25] and/or overall survival [19-21, 23, 24]. Moreover, the GLIM framework has been found to outperform the International Classification of Diseases 10th Revision criteria (ICD-10) in identifying more malnourished patients and establishing a stronger correlation with surgical risk [17]. These findings underscore the superiority and clinical relevance of the GLIM framework in patients with cancer [24]. However, it is important to note that existing studies have focused on the presence of malnutrition at a single time point, mainly at baseline [9, 10, 17, 18, 20, 22, 24, 26, 27]. To our knowledge, no study has investigated the transition of GLIM-defined malnutrition within a clinically operational timeframe in cancer. Considering the variability in responses to

malnutrition and its treatment among individuals [3, 11], an approach that can identify patients who may benefit from multidisciplinary treatment and/or are likely to recover from malnutrition would be valuable for making individualized management decisions.

Weight loss, body mass index (BMI) and muscle mass are key elements used to identify, phenotype and grade malnutrition [16]. Additionally, these factors are useful in predicting future risks of weight loss [28] and muscle loss [29]. However, no studies have yet used weight and muscle information to predict the future outcomes of malnutrition. Based on this understanding, we hypothesize that body weight and muscle dynamic data routinely accessible upon patient admission can independently predict the fate of malnutrition. In this study, we proposed a deep learning model for early identification of reversible malnutrition (RM) in patients with cancer. The primary objective of this study was to enhance the decision-making of cancer-associated malnutrition to optimize patient outcomes.

Materials and Methods

Study design and population

This was a retrospective cohort study conducted at multiple centers in China with prospectively collected data. Patients were derived from the Investigation on Nutrition Status and its Clinical Outcome of Common Cancers (INSCOC) which was registered online (http://www.chictr.org.cn/showproj.aspx?proj=31813, identifier: ChiCTR1800020329). The INSCOC is a clinical research project initiated by the Chinese Society for Nutritional Oncology since 2013. The aim of INSCOC is to determine the prevalence of malnutrition in inpatients with cancer in China and to examine its relationship with the clinical outcomes. The project protocol has been published previously [30] and the inclusion and exclusion criteria for participants are detailed in **Table S1**. In accordance with these criteria, we initially included 30766 patients aged over 18 years who were firstly diagnosed with cancer and/or were hospitalized for anti-cancer treatment from December 2013 to May 2021. Further exclusions were made for patients who lacked required body weight dynamic data (n=16054), had outlier values (n=71), had unclear pathological results (n=567) and did not meet the baseline malnutrition criteria according to the GLIM framework (n=9832) [16]. This left 4254 patients with diagnosed malnutrition for formal analysis. A flowchart illustrating the patient inclusion process is presented in **Figure S1**. A cohort including 798 malnourished oncology patients (the Yunnan cohort), which was independent of the study population, was used for further model validation. The study was approved by the Ethics Committee of Army Medical Center of PLA (approval number: 2018-22) and written consent was provided by all patients. All data were de-identified before analysis and the principles of the Declaration of Helsinki were followed.

The Strengthening and Reporting of Observational Studies in Epidemiology Statement was followed.

Data acquisition and handling

Data were collected through in-person interview and physical examination by project-trained researchers within the first 48 hours after patient admission: patient age, sex, smoking status (active tobacco smoker in the past one year), alcohol drinking (once a week or more frequent alcohol consumption in the past one year, regardless of amount and type), tea drinking (once a week or more frequent tea consumption in the past one year, regardless of amount and type), residency (urban vs. rural), occupation (blue vs white collar), body height, body weight (six months before, one month before, at the time of, and one month after baseline, these time points were abbreviated as –6, –1, baseline and +1), BMI, weight loss, mid-arm circumference (MAC), triceps skinfold thickness (TSF), handgrip strength (HGS), mid-arm muscle circumference (MAMC), calf circumference (CC), appendicular skeletal muscle mass index (ASMI), food intake, gastrointestinal (GI) symptoms, the Eastern Cooperative Oncology Group (ECOG) physical performance score , the Nutritional Risk Screening 2002 (NRS2002) score [31], the Patient-Generated Subjective Global Assessment (PG-SGA) score [32] and the quality of life (QoL) score (using the global QoL score of the European Organization for Research and Treatment of Cancer QLQ-C30 scale).

The detailed approaches used, including the formulas, procedures and devices used to obtain the anthropometric indices (height, weight, BMI, weight loss, MAC, TSF, HGS, MAMC, CC and ASMI) are shown in **Table S2**. The BMI (kg/m^2) was as also categorized as underweight (<18.5), normal (18.5 to <24), overweight (24 to <28) or obese (≥28) according to the Chinese recommendations [33]. The GI symptoms of patients were assessed in accordance with the PG-SGA [32]. Cancer cachexia was retrospectively diagnosed and staged using the Fearon's framework [34]. The clinical characteristics recorded during hospitalization, including the comorbidities, cancer site, clinical tumor stage, anticancer therapies received, nutritional support received, laboratory indices, ECOG score one month after admission was retrospectively retrieved from electronic medical records. Body weight one month after admission was either measured (if hospital stay ≥30 days), or patient-reported via a follow-up after discharge (if hospital stay <30 days).

Definitions of malnutrition and reversible malnutrition

Malnutrition was retrospectively diagnosed based on the GLIM criteria [16]. Briefly, for patients who were at nutritional risk (NRS2002 ≥ 3), at least one phenotypic criterion and one etiologic criterion needed to be met to diagnose malnutrition. The phenotypic criteria were: involuntary

weight loss 5–10% within the past 6 months, or 10–20% beyond 6 months; BMI<18.5 kg/m^2 if <70 years, <20 kg/m^2 if ≥70 years [4]; or reduced muscle mass (ASMI, male <7.0 kg/m^2 or female <5.4 kg/m^2, based on the Asian Working Group for Sarcopenia 2019 Consensus [35]). For the etiologic criteria, the entire study population was considered positive for the disease burden-related etiologic criterion since all patients were pathologically diagnosed with cancer and/or hospitalized for cancer treatment [4]. The diagnosis of malnutrition was independently performed at baseline and one month after baseline. RM was the primary outcome of the study, defined as a diagnosis of malnutrition at baseline (upon patient admission) turning negative one month later (all three phenotypic criteria being negative).

Machine learning models building

The study population (n=4254) was shuffled and randomly split into a discovery set (n=2977, 70%) for model training and a holdout test set (n=1277, 30%) to evaluate the model performance. The Yunnan cohort (n=798) was used as an external validation set (**Figure S2**). Based on the study hypothesis, feature selection was performed by manually selecting variables related to body weight and muscle dynamics that can be obtained upon patient admission. This approach was found to be feasible in our previous research [36, 37]. Six variables, including BMI and ASMI at -6, -1, and baseline, were defined as input variables, and RM (yes vs. no) was defined as the binary outcome variable. The input data were standardized using a Z-score approach before modeling (Equation 1, where x represents raw score, μ represents mean, σ represents standard deviation).

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

Since the input variables are time-series data, we used a long short-term memory (LSTM) recurrent neural network architecture for modeling to capture the sequential information. The model was named as WAL-net (representing weight, appendicular skeletal muscle and LSTM-based deep neural network). The WAL-net was trained using a 32-sample mini-batch technique with a binary cross-entropy loss function (Equation 2, where N represents the total number of mini-batch samples, i represents an index that iterates over each sample from 1 to N, y_i represents the true label of the i-th sample, p represents probability) and an Adam optimizer at a learning rate of 0.001.

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log\left(p(y_i)\right) + (1 - y_i) \cdot log(1 - p(y_i)) \quad (2)$$

A detailed architecture of the WAL-net is shown in **Figure S3**. Accuracy, area under the curve (AUC), recall, precision, F1 score, Kappa and Matthews correlation coefficient (MCC) were

used as model comparison metrics, with higher values indicating better prediction performance [38]. We also utilized sensitivity, specificity, positive predictive value, and negative predictive value to evaluate and demonstrate the performance of the model, as these metrics are more commonly used in clinical scenarios [4, 5, 39]. To verify the superiority of the WAL-net, another 19 conventional machine learning (ML) algorithms were independently developed for the same binary classification task. A ten-fold cross-validation with ten iterations technique (**Figure S4**) was used to aggregate model performance metrics for comparison and to select model with optimal hyper-parameters. The optimal model among the 19 ML algorithms was then compared with the WAL-net in the holdout test set. The WAL-net was also assessed in different subgroups of the test set, including age, sex, cancer site, clinical stage, curative surgery, curative chemotherapy, nutritional support, C-reactive protein, ECOG score and PG-SGA score to evaluate potential effect modifications. The explainability of the model was evaluated using a shapley additive explanations (SHAP) method at both the group and individual levels. A decision curve analysis (DCA) was used to evaluate the model's clinical usefulness. Finally, the WAL-net was deployed as a prototypic web-based application for online prediction. The project code and files have been stored online in our GitHub repository (https://github.com/kevinlyy/rmalnutrition) for public access.

Statistical analysis

Continuous data were expressed as the medians (interquartile range) and compared using a Wilcoxon's rank-sum test. Continuous data with multiple groups was compared using a pairwise Wilcoxon's rank-sum test and *P* values of multiple comparisons were adjusted using a Bonferroni's method. Categorical data were expressed as numbers (percentage) and were compared using a Chi-squared test. Cutoff values of the input variables to predict RM were calculated by maximizing the Youden's index (sensitivity + specificity − 1) in the discovery set. The generated cutoffs were then assessed in the holdout test set. Model-predicted probability greater than 0.5 was defined as the cutoff point for belonging to the RM group. All reported *P* values were two-sided and considered significant at $P < 0.05$. All analyses were performed using R (version 4.3.1, Foundation for Statistical Computing, Vienna, Austria). Deep learning and other ML algorithms were conducted in Python (version 3.9.11, The Python Software Foundation, USA).

Results

Population overview

The baseline characteristics of the study population are shown in the overall column of **Table 1**. There were 4254 patients with a median age of 61 years, including 2783 males and 1471 females. The tumors were most frequently located in the lung (n=984, 23.1%), colorectum (n=870, 20.5%), stomach (n=749, 17.6%), esophagus (n=416, 9.8%) and nasopharynx (n=267, 6.3%). The number of cases for all cancer sites is shown in **Figure 1A**. The predominant clinical tumor stages were II (32.4%) and III (31.2%).

Reversible malnutrition

RM was found in 754 (17.7%) patients, while the remaining patients exhibited at least one positive phenotypic criterion one month following baseline. The prevalence of RM was further analyzed based on the specific cancer sites as shown in **Figure 1B**. The RM was most frequently observed in prostate cancer (34.8%), brain cancer (25.0%) and multiple myeloma (25.0%), while being least prevalent in cancers located in the biliary tract (5.3%), ovary (5.4%) and esophagus (8.4%).

Dynamics of body weight and skeletal muscle

The BMI and ASMI dynamics over the four time points (−6, −1, baseline, and +1) were assessed (**Figure 1C-1F** and **Table 1**). Across the four time points, there was a significant downward trend in BMI within the study population (**Figure 1C**). In stratified analysis, RM and non-RM groups showed significant difference in the patterns of BMI changes (**Figure 1D**). Briefly, the RM group has higher BMI than the non-RM group at the −6, −1 and +1 time points, but lower BMI at baseline. In-group multiple comparison showed that the BMI change was significant in all time intervals except for the −6 vs −1 (P=0.060), −6 vs +1 (P=0.740) and −1 vs +1 (P=0.430) intervals for the RM group. In contrast, the BMI change was significant in all time intervals except for the −1 to 0 interval for the non-RM group. Regarding ASMI, there was a significant downward trend within the study population across the whole pre-admission interval (−6 to 0) but not the 0 to +1 interval (P=0.930, **Figure 1E**). In stratified analysis (**Figure 1F**), the RM group has higher ASMI than the non-RM group at all four time points. In-group multiple comparison showed that the ASMI was significantly increased in the 0 to +1 interval (P=0.049) for the RM group. In contrast, the ASMI change was only insignificant in the 0 to +1 interval (P=0.220) for the non-RM group. Heatmaps for the distribution of BMI and ASMI at different time points stratified by cancer site in the overall, RM and non-RM groups are shown in **Figure S5**. The results were similar to the findings observed in the overall population, indicating distinct patterns of BMI and ASMI dynamics between the RM and

non-RM groups.

Reversible malnutrition and patient characteristics

The baseline characteristics of the study population, as stratified by the RM status are shown in **Table 1**. RM was significantly associated with higher values/rates of male sex, smoking, tea drinking, diabetes, curative chemotherapy, total protein, creatinine, albumin, urea nitrogen, prealbumin, glucose, triglycerides, alkaline phosphatase, hemoglobin, lymphocytes, red blood cells, body height, MAC, TSF, HGS, MAMC, CC, normal physical performance and global quality of life score. In contrast, RM was associated with lower values/rates of age, blue collar occupation, curative radiotherapy, nutritional support, platelets, C-reactive protein, neutrophil to lymphocyte ratio, anorexia, nausea, vomiting, constipation, dry mouth, taste changes, smell changes, dysphagia, early satiety and the PG-SGA score. In addition, the cancer sites, nutritional support type, food intake and the ECOG score category were also different between the RM and non-RM groups (all $P<0.05$).

Reversible malnutrition and stages of cancer cachexia

The association between RM and stages of cancer cachexia was analyzed in the overall population and in subgroups of cancer site (**Figure 2** and **Table S3**). Compared to the non-RM group, the RM group was associated higher non-cachexia (41.5% vs. 9.8%) and pre-cachexia rates (4.6% vs. 2.4%) rates, but with lower cachexia (cachexia = 53.8% vs. 87.1%) and refractory cachexia (0% vs. 0.7%, $P < 0.001$) rates. Notably, all patients with refractory cachexia were classified into the non-RM group. Moreover, a spearman's rank correlation analysis revealed a negative relationship between the progressive stages of cancer cachexia and RM ($r = -0.340$, $P < 0.001$). Subgroup analysis in gastrointestinal, respiratory, hematologic and others cancers revealed similar results (all $P<0.001$).

Model training, comparison and assessment

A graphical workflow of the WAL-net is shown in **Figure 3A**. The WAL-net starts with a LSTM layer to extract latent information from the sequential data of BMI and ASMI at three time points. After data processing with two fully-connected layers separated by a rectified linear unit (ReLU) layer, the model outputs a probability vector of the two outcome labels, which is subsequently normalized using the softmax function.

Distributions of the study variables in the discovery and test sets are shown in **Table 1**. We first trained the 19 ML models by setting the six variables of BMI and ASMI as the input variables and the RM (yes vs. no) as the outcome variable in the discovery set. The ten-fold cross-validated results of the 19 ML models are shown in **Figure 3B**. The multi-layer perceptron (MLP) model showed the highest performance, with an accuracy = 0.899, AUC =

0.839, recall = 0.536, precision = 0.831, F1 score = 0.647, Kappa = 0.591 and MCC = 0.613. Therefore, the MLP model was selected among the 19 conventional ML models for future use, with its hyper-parameters shown in **Table S4**.

Subsequently, the WAL-net was trained in the discovery set and was compared to the MLP model in the holdout test set (**Figure 3C-D** and **Table 2**). The WAL-net showed higher performance (AUC=0.924, 95%CI=0.904 to 0.944) than the MLP model (AUC=0.899, 95%CI=0.874 to 0.924). A Delong's test indicated that the AUC difference was statistically significant (P = 0.005). For other metrics, the WAL-net showed higher accuracy (0.924 vs. 0.908), Kappa index (0.728 vs. 0.660), sensitivity (0.878 vs. 0.850), specificity (0.932 vs. 0.917), positive predictive value (0.690 vs. 0.615) and negative predictive value (0.978 vs. 0.975) than the MLP model. Thus the WAL-net was finally selected for future analysis as the state-of-the-art (SOTA) model in our specific task. The precision-recall (PR) curve, Kolmogorov Smirnov (KS) statistic plot, cumulative gains curve, lift curve, calibration curve and confusion matrix and of the WAL-net in the test data are shown in **Figure 4A-F**. The area and micro-average area under the PR curve were 0.805 (class 1, the RM group) and 0.958, respectively (**Figure 4A**). The KS statistic plot, cumulative gains curve, lift curve, calibration curve showed that the WAL-net has strong discrimination power and can effectively separate the positive and negative classes (**Figure 4B-E**). A Hosmer and Lemeshow goodness-of-fit test also supported the good consistency between model prediction and actual observation (P = 0.212, **Figure 4E**). The confusion matrix showed that there were 1015 (class 0, non-RM) and 165 (class 1, RM) patients who were correctly classified in the test data (**Figure 4F**).

Model performance in subgroups

The subgroup performance of the WAL-net was assessed in the test data (**Table 2**). Overall, the WAL-net sustained its good performance in different subgroups investigated. Specifically, by setting the AUC as the reference metric of model performance, the WAL-net showed relatively higher performance (defined as AUC > 0.950) in the subgroup of patients who only received enteral nutritional support (AUC=0.980, 95%CI=0.954 to 1.000) and patients with a PG-SGA score of 2 to 3 (AUC=0.957, 95%CI=0.921 to 0.993).

Model explainability

We evaluated the predictors of the WAL-net that contributed to the prediction on a global level. The SHAP summary plot showed that ASMI at baseline was the leading contributor to the high likelihood of RM, with a mean SHAP value of 1.590. In addition, the mean SHAP values for BMI, ASMI−1, ASMI−6, BMI−1 and BMI−6 were 0.806, 0.731, 0.417, 0.145 and 0.143, respectively (**Figure 5A**). We also examined the individual-level risk predictions and their

sources of risk specified by the SHAP values. For the patient with the highest predicted SHAP value (e.g. 1), ASMI (0.69), BMI (0.11), BMI−1 (0.05), ASMI−1 (0.02), BMI−6 (0.02) and ASMI−6 (0.01) were the sources of risk that led to the high SHAP value (**Figure 5B**). On the other hand, for the patient with the lowest SHAP value (e.g. 0), BMI−6 (0.10), ASMI (−0.08), BMI−1 (−0.04), ASMI−6 (−0.04), ASMI−1 (−0.03) and BMI (−0.01) were the sources of risk that led to the low SHAP value (**Figure 5C**).

Clinical relevance

To examine the clinical relevance of the WAL-net, we analyzed the associations of model-predicted RM with the clinical outcomes one month after admission (**Table 3**). Model-predicted RM was associated with higher values/rates of follow up BMI and ASMI. Patients predicted as RM by the WAL-net at baseline were less likely to be classified in the underweight group and low muscle mass group one month after admission. In addition, patients with predicted RM were more likely to be classified in the normal physical performance one month after admission. Notably, the number of positive phenotypic criteria of GLIM were less likely to increase one month after admission in the RM group (all P<0.05). Additionally, DCA analysis in the test set showed that if the threshold probability of a patient was >0.01, using the WAL-net to predict the probability of RM adds more benefits than either the treat-all-patients scheme or the treat-none scheme (**Figure S6**).

Cutoff value for each input variable

To facilitate clinical use, the cutoffs (kg/m^2) for the six input variables to predict RM were independently calculated in the discovery set (sex-specific cutoffs were calculated for ASMI) and then assessed in the test set (**Table S5**). For the three time points (−6, −1 and baseline), the BMI cutoffs were <21.2, <20.6 and <18.5, respectively. Likewise, the ASMI cutoffs were men <7.3 or women <5.3, men <7.0 or women <5.3, and men <7.1 or women <5.3 for the three time points, respectively. After dichotomized using these cutoffs, the baseline BMI showed the highest prognostic value to predict RM in the test set (AUC=0.771, 95%CI=0.739 to 0.802).

Independent model validation

The WAL-net was further assessed for its classification power in the Yunnan cohort. Baseline characteristic of this cohort is shown in **Table S6**. RM was found in 106 (13.3%) patients. The model's good performance sustained, with an AUC (95%CI)=0.909 (0.876, 0.943), accuracy (95%CI)=0.940 (0.921, 0.955), Kappa=0.686, sensitivity=0.953, specificity=0.939, PPV=0.575, and NPV=0.996 (**Table S7**).

Model deployment

Prototypic applications were developed to utilize the WAL-net. These applications were designed to receive user input BMI and ASMI parameters and provide instant RM predictions. These applications deliver both the predicted class and the corresponding probability or confidence as outputs. The graphical user interface of the applications are presented in **Figure S7**. The serialized model objects that support future reuse and the application-related files have been stored online in our GitHub repository for public access (https://github.com/kevinlyy/rmalnutrition).

Discussion

In this study, we focused on addressing a significant real-world challenge regarding the reversibility of malnutrition in patients with cancer. To our knowledge, this is the first large-scale study that predicts RM using an explainable, LSTM-based deep learning model with high performance (AUC = 0.924). We have also confirmed an important hypothesis: the dynamic information of body weight and predicted appendicular skeletal muscle, which is routinely available upon patient admission, can independently predict the future trajectory of cancer-associated malnutrition. As our model does not require consideration of the treatment the patient will subsequently receive, this constitutes its ability for early identification. The findings underscore the importance of body weight and muscle mass surveillance within clinically operational timeframes for oncology patients. The model we developed, along with other findings presented in the study, might assist clinicians with decision-making to help guide more individualized management strategies of malnutrition in cancer care.

Compared to previous studies that only reported the baseline prevalence of GLIM-defined malnutrition in patients with cancer [4, 5], the present study provides more clinically relevant information for decision-making by reporting the cancer-specific prevalence of RM (**Figure 1B)**. Clinicians in oncology practice may typically prioritize cancer sites with higher incidence rates of malnutrition, such as pancreatic cancer [40] and gastrointestinal cancers [41]. However, our observations revealed a distinct pattern in the reversibility of malnutrition across different cancer sites compared to their prevalence. For instance, prostate and brain cancer are generally associated with a low incidence of malnutrition [4], but they exhibited higher rates of RM (**Figure 1B**). These findings underscore the importance of early identification and targeted management strategies, particularly for patients with RM in these specific cancer sites. Implementing such strategies increases the likelihood of achieving reversal and restoring weight/muscle mass [15]. In contrast, malnutrition in some cancers, such as those located in the

biliary tract, ovary and esophagus, is more difficult to ameliorate (RM% < 10%). This suggests that malnutrition occurring in these cancer sites is more refractory. These results (non-RM% of GI cancers=85.1%) partially align with previous studies reporting the statistics of refractory cachexia in patients with cancer [42, 43], which showed that GI cancers were the most refractory. A recent systematic review found that in patients with incurable solid cancer, nutritional intervention, either alone or as part of a multimodal approach, has improved quality of life, body weight, and nutritional intake [44]. Another review also indicated that multimodal interventions, including nutritional support, are effective in preventing unintentional weight loss in cancer patients [45]. These lines of evidence, along with our research findings, collectively suggest that for patients classified as non-RM, additional efforts and specialized interventions may be required to address the unique challenges associated with these cancers and potentially improve outcomes [43]. Similarly, this decision strategy may also be applicable to the RM/non-RM status in the overall population. In this study, although the RM group received less nutritional support than the non-RM group, it was associated with improved nutrition-related outcomes. This implies that the RM/non-RM status predicted by the model might provide additional information to distinguish whether a malnourished patient is likely or unlikely to benefit from nutritional or multidisciplinary intervention. However, epidemiological statistics on cancer-specific prevalence of RM remain scare, especially those derived from Asian populations. As we only include malnourished patients, future studies should aggregate data on both cancer-specific malnutrition prevalence and RM prevalence within the same ethnic groups. Additionally, elucidating the underlying mechanisms behind the differing prevalence of RM in various cancer sites is crucial.

A potential drawback of the GLIM framework is that it does not provide overall nutritional/multidisciplinary treatment recommendations with respect to different malnutrition phenotypes and severities [16]. To inform treatment strategies for patients with GLIM-defined malnutrition, one possible solution is to refer to the cancer cachexia guideline [34], as malnutrition with inflammation and cachexia have been described as interchangeable concepts in some literature [1, 46]. Fearon's framework for diagnosing cancer cachexia defines progressive stages of cachexia (from pre-cachexia to cachexia to refractory cachexia) [34]. Early monitoring and preventive intervention are recommended for patients with pre-cachexia, as once cachexia is well-established, treatment becomes more difficult [47]. For patients at the cachexia stage, multimodal management according to phenotype is needed to prioritize those reversible contributory factors. [34]. For refractory cachexia, medical interventions may be futile or inappropriately invasive [3, 48]. However, it is important to note that although the diagnostic

parameters used in Fearon's framework [34] are the same as the phenotype criteria in the GLIM [16], the cutoff values for specific variables such as weight loss and muscle loss differ. Additionally, other frameworks may employ different criteria for diagnosing cachexia [49, 50]. Therefore, treatment strategies applicable to the stages of cancer cachexia may not necessarily fully applicable to the different phenotypes and grades defined by GLIM. In contrast, our model was directly developed based on GLIM, making it more suitable for decision-making in the context of malnutrition. As the global recognition of GLIM continues to increase [27], our model is promising to serve as a useful tool to evaluate the burdens and costs of anti-malnutrition therapies against the expected benefits for patients.

Another noteworthy point to consider is that our findings reflect the real-world challenges that exist in current treatment strategies for cancer-associated malnutrition in China [30]. Even among patients diagnosed with malnutrition, the coverage rate of nutritional support remains low (36.9%), and only a small proportion of malnourished patients can be reversed (17.7%). We observed in our own practice that clinicians may instinctively prioritize patients with more severe malnutrition-related phenotypes, while these patients may face greater challenges in deriving benefits from nutritional support. Indeed, patients in the non-RM group exhibited a higher prevalence of various gastrointestinal symptoms, reduced food intake, nutritional risk, low BMI, impaired physical status, advanced cachexia stage, etc. However, despite a higher rate of nutritional support during hospitalization in the non-RM group, there was no significant improvement in clinical benefits compared to the RM group. Because the severity of the primary disease, as indicated by the clinical tumor stage, are not significantly different between the non-RM and RM groups. It might be true that a diagnosis of malnutrition alone is insufficient to guide nutritional intervention strategies. A refined patient classification is necessary to provide guidance for more individualized treatment approaches. For example, creating a staging system that reflects the pathophysiological progression of malnutrition and treatment response, rather than solely relying on the severity of phenotypic parameters. Another possible reason for the low rate of RM is: the nutritional interventions in the present cohort were not strictly individualized and were not combined with other multimodal approaches such as physical exercise program, pharmacological interventions and psychological support [30]. Therefore, further studies in patients with individualized nutrition support data [51], especially multimodal anti-malnutrition data [3], are imperative to provide greater insights regarding the clinical usefulness of the model.

The study has several potential limitations that must be noted. First, malnutrition and RM were both diagnosed using the GLIM framework that includes the weight loss as a major

criterion. Because weight loss was calculated based on patient-reported historic weight data, the impact of recall bias on the classification of malnutrition and RM cannot be eliminated [39]. Future studies with a prospective design and measured body weight values are needed to replicate our findings. Second, the ASMI was derived from an equation validated for use in Asians rather than measured using technologies such as dual energy x-ray imaging or bioimpedance analysis. However, the equation showed good consistency with dual energy x-ray imaging [52], and body weight and height are simple to obtain in almost any institutions, which should increase the usability of the study parameters in different scenarios. Nevertheless, future studies using more advanced technologies to measure appendicular muscle mass are needed. Third, the model utilizes weight and skeletal muscle dynamic information from three time points for predicting RM. Future studies need to investigate whether incorporating additional time points would result in further improvements in prediction performance. However, the weight loss information over past one and/or six months has been included in the most popular nutritional assessment tools which are easily accessible in oncology practice [16, 31, 32], further increasing the generalizability of our model. Fourth, the present study only observed RM one month after baseline, future studies employ other time points are needed. Future studies with a larger sample size for a wider spectrum of diseases are needed to address the above issues.

Conclusion

In conclusion, we have provided the first report on the overall and cancer-specific prevalence of RM in a nationwide, multicenter cancer cohort. We have confirmed the hypothesis that the dynamic information of body weight and skeletal muscle, which is routinely available upon patient admission, can independently predict the future trajectory of malnutrition with high accuracy. We developed and validated an explainable deep learning model for the early identification of RM. This model utilizes six easily accessible variables related to body weight and appendicular skeletal muscle dynamics, which showed good performance to predict RM. These findings might assist clinicians or nutritionists in decision-making to help guide management strategies of cancer-associated malnutrition and optimize the allocation of healthcare resources in cancer care.

Data availability

Data described in the manuscript will not be made available because of multiple confidentiality agreements signed with participating hospitals.

Acknowledgements

Conflicts of interest: The authors declare that they have no conflicts of interest.

Funding: This work was supported by the National Natural Science Foundation of China (82304131), the Natural Science Foundation of Chongqing, China (CSTB2024NSCQ-MSX1233), the Young Doctoral Talent Incubation Program of the Xinqiao Hospital, Army Medical University (2024YQB033) and the National Key Research and Development Program of China (2017YFC1309200).

Author contributions: LY: Conceptualization, Methodology, Software, Data curation, Writing—Original draft preparation, Funding acquisition, and Revising; NL and JL: Data curation and Project administration; WL, JC, ZG, QY, FZ, ML, ZC, HY, TL, ZL, PJ and CS: Methodology, Data curation, and Project administration; HX: Project administration, Data curation, Methodology, and Reviewing; HS: Conceptualization, Supervision, Project administration, Methodology, Software and Funding acquisition, and Reviewing.

The authors would like to thank the INSCOC project members for their substantial work on data collection and patient follow-up.

Abbreviations: GLIM, the Global Leadership Initiative on Malnutrition; ICD-10, the International Classification of Diseases 10th Revision criteria; BMI, body mass index; RM, reversible malnutrition; INSCOC, the Investigation on Nutrition Status and its Clinical Outcome of Common Cancers; MAC, mid-arm circumference; TSF, triceps skinfold thickness; HGS, handgrip strength; MAMC, mid-arm muscle circumference; CC, calf circumference; ASMI, appendicular skeletal muscle mass index; GI, gastrointestinal; ECOG, the Eastern Cooperative Oncology Group; NRS2002, the Nutritional Risk Screening 2002; PG-SGA, the Patient-Generated Subjective Global Assessment; QoL, quality of life; LSTM, long short-term memory; WAL, weight, appendicular skeletal muscle and long short-term memory; AUC, area under the curve; MCC, Matthews correlation coefficient; ML, machine learning; SHAP, shapley additive explanations; DCA, decision curve analysis; ReLU, rectified linear unit; MLP, multi-layer perceptron; SOTA, state-of-the-art; PR, precision-recall; KS, Kolmogorov Smirnov

Registration information: The data described in the manuscript were derived from the INSCOC project, which was registered online at https://www.chictr.org.cn (URL: http://www.chictr.org.cn/showproj.aspx?proj=31813; ID: ChiCTR1800020329)

References

1. Cederholm T, Barazzoni R, Austin P, et al. (2017) ESPEN guidelines on definitions and terminology of clinical nutrition. *Clinical nutrition* **36**:49-64.
2. Wang Z, Barazzoni R, Shi H. (2023) Medical nutrition is one basic human right. *Precision Nutrition* **2**:e00026.
3. Arends J, Bachmann P, Baracos V, et al. (2017) ESPEN guidelines on nutrition in cancer patients. *Clinical nutrition* **36**:11-48.
4. Yin L, Song C, Cui J, et al. (2021) A fusion decision system to identify and grade malnutrition in cancer patients: Machine learning reveals feasible workflow from representative real-world data. *Clinical nutrition* **40**:4958-70.
5. Yin L, Lin X, Liu J, et al. (2021) Classification Tree-Based Machine Learning to Visualize and Validate a Decision Tool for Identifying Malnutrition in Cancer Patients. *JPEN Journal of parenteral and enteral nutrition* **45**:1736-48.
6. Gioulbasanis I, Baracos VE, Giannousi Z, et al. (2011) Baseline nutritional evaluation in metastatic lung cancer patients: Mini Nutritional Assessment versus weight loss history. *Annals of oncology : official journal of the European Society for Medical Oncology* **22**:835-41.
7. Almasaudi AS, McSorley ST, Dolan RD, et al. (2019) The relation between Malnutrition Universal Screening Tool (MUST), computed tomography-derived body composition, systemic inflammation, and clinical outcomes in patients undergoing surgery for colorectal cancer. *The American journal of clinical nutrition* **110**:1327-34.
8. Hu C, Barazzoni R, Shi H. (2023) Nutritional care is the first-line therapy for many conditions. *Precision Nutrition* **2**:e00059.
9. Rosato E, Gigante A, Colalillo A, et al. (2023) GLIM-diagnosed malnutrition predicts mortality and risk of hospitalization in systemic sclerosis: A retrospective study. *Eur J Intern Med*.
10. Brown D, Loeliger J, Stewart J, et al. (2023) Relationship between global leadership initiative on malnutrition (GLIM) defined malnutrition and survival, length of stay and post-operative complications in people with cancer: A systematic review. *Clinical nutrition* **42**:255-68.

11. Arends J, Baracos V, Bertz H, et al. (2017) ESPEN expert group recommendations for action against cancer-related malnutrition. *Clinical nutrition* **36**:1187-96.

12. Gyan E, Raynard B, Durand JP, et al. (2018) Malnutrition in Patients With Cancer: Comparison of Perceptions by Patients, Relatives, and Physicians-Results of the NutriCancer2012 Study. *JPEN Journal of parenteral and enteral nutrition* **42**:255-60.

13. Attar A, Malka D, Sabate JM, et al. (2012) Malnutrition is high and underestimated during chemotherapy in gastrointestinal cancer: an AGEO prospective cross-sectional multicenter study. *Nutrition and cancer* **64**:535-42.

14. Hebuterne X, Lemarie E, Michallet M, et al. (2014) Prevalence of malnutrition and current use of nutrition support in patients with cancer. *JPEN Journal of parenteral and enteral nutrition* **38**:196-204.

15. Kipouros M, Vamvakari K, Kalafati IP, et al. (2023) The Level of Adherence to the ESPEN Guidelines for Energy and Protein Intake Prospectively Influences Weight Loss and Nutritional Status in Patients with Cancer. *Nutrients* **15**.

16. Cederholm T, Jensen GL, Correia M, et al. (2019) GLIM criteria for the diagnosis of malnutrition - A consensus report from the global clinical nutrition community. *Clinical nutrition* **38**:1-9.

17. Murnane LC, Forsyth AK, Koukounaras J, et al. (2023) Malnutrition defined by GLIM criteria identifies a higher incidence of malnutrition and is associated with pulmonary complications after oesophagogastric cancer surgery, compared to ICD-10-defined malnutrition. *Journal of surgical oncology* **128**:769-80.

18. Matsui R, Rifu K, Watanabe J, et al. (2023) Impact of malnutrition as defined by the GLIM criteria on treatment outcomes in patients with cancer: A systematic review and meta-analysis. *Clinical nutrition* **42**:615-24.

19. Kiss N, Steer B, de van der Schueren M, et al. (2023) Machine learning models to predict outcomes at 30-days using Global Leadership Initiative on Malnutrition combinations with and without muscle mass in people with cancer. *Journal of cachexia, sarcopenia and muscle* **14**:1815-23.

20. Yin L, Chong F, Huo Z, et al. (2022) GLIM-defined malnutrition and overall survival in cancer patients: A meta-analysis. *JPEN Journal of parenteral and enteral nutrition*.

21. Yin L, Lin X, Li N, et al. (2021) Evaluation of the Global Leadership Initiative on Malnutrition Criteria Using Different Muscle Mass Indices for Diagnosing Malnutrition and Predicting Survival in Lung Cancer Patients. *JPEN Journal of parenteral and enteral nutrition* **45**:607-17.

22. Milanez DSJ, Razzera EL, Lima J, et al. (2023) Feasibility and criterion validity of the GLIM criteria in the critically ill: A prospective cohort study. *JPEN Journal of parenteral and enteral nutrition* **47**:754-65.

23. Yin L, Fan Y, Lin X, et al. (2021) Fat mass assessment using the triceps skinfold thickness enhances the prognostic value of the Global Leadership Initiative on Malnutrition criteria in patients with lung cancer. *The British journal of nutrition*:1-11.

24. Yin L, Cheng N, Chen P, et al. (2021) Association of Malnutrition, as Defined by the PG-SGA, ESPEN 2015, and GLIM Criteria, With Complications in Esophageal Cancer Patients After Esophagectomy. *Frontiers in Nutrition* **8**.

25. Peng D, Zong K, Yang H, et al. (2022) Malnutrition diagnosed by the Global Leadership Initiative on Malnutrition criteria predicting survival and clinical outcomes of patients with cancer: A systematic review and meta-analysis. *Front Nutr* **9**:1053165.

26. Liu C, Lu Z, Li Z, et al. (2021) Influence of Malnutrition According to the GLIM Criteria on the Clinical Outcomes of Hospitalized Patients With Cancer. *Front Nutr* **8**:774636.

27. Barazzoni R, Jensen GL, Correia M, et al. (2022) Guidance for assessment of the muscle mass phenotypic criterion for the Global Leadership Initiative on Malnutrition (GLIM) diagnosis of malnutrition. *Clinical nutrition* **41**:1425-33.

28. O'Donoghue N, Shrotriya S, Aktas A, et al. (2019) Clinical significance of weight changes at diagnosis in solid tumours. *Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer* **27**:2725-33.

29. Hsu WH, Ko AT, Weng CS, et al. (2023) Explainable machine learning model for predicting skeletal muscle loss during surgery and adjuvant chemotherapy in ovarian cancer. *Journal of cachexia, sarcopenia and muscle* **14**:2044-53.

30. Xu H, Song C, Yin L, et al. (2022) Extension protocol for the Investigation on Nutrition Status and Clinical Outcome of Patients with Common Cancers in China (INSCOC) study: 2021 update. *Precision Nutrition* **1**:e00014.

31. Kondrup J, Allison SP, Elia M, et al. (2003) ESPEN guidelines for nutrition screening 2002. *Clinical nutrition* **22**:415-21.

32. Ottery FD. (1994) Rethinking nutritional support of the cancer patient: the new field of nutritional oncology. *Semin Oncol* **21**:770-8.

33. Chen C, Lu FC, Department of Disease Control Ministry of Health PRC. (2004) The guidelines for prevention and control of overweight and obesity in Chinese adults. *Biomedical and environmental sciences : BES* **17 Suppl**:1-36.

34. Fearon K, Strasser F, Anker SD, et al. (2011) Definition and classification of cancer cachexia: an international consensus. *The Lancet Oncology* **12**:489-95.

35. Chen LK, Woo J, Assantachai P, et al. (2020) Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment. *Journal of the American Medical Directors Association* **21**:300-7 e2.

36. Yin L, Li N, Lin X, et al. (2025) Early identification of potentially reversible cancer cachexia using explainable machine learning driven by body weight dynamics: a multicenter cohort study. *The American journal of clinical nutrition* **121**:535-47.

37. Yin L, Zhang L, Li L, et al. (2024) Exploring the optimal indicator of short-term peridiagnosis weight dynamics to predict cancer survival: A multicentre cohort study. *Journal of cachexia, sarcopenia and muscle* **15**:1177-86.

38. Yin L, Liu J, Liu M, et al. (2023) Ensemble learning system to identify nutritional risk and malnutrition in cancer patients without weight loss information. *Science China Life sciences* **66**:1200-3.

39. Yin L, Cui J, Lin X, et al. (2022) Identifying cancer cachexia in patients without weight loss information: machine learning approaches to address a real-world challenge. *The American journal of clinical nutrition* **116**:1229-39.

40. Bouloubasi Z, Karayiannis D, Pafili Z, et al. (2023) Re-assessing the role of peri-operative nutritional therapy in patients with pancreatic cancer undergoing surgery: a narrative review. *Nutr Res Rev*:1-10.

41. Yaceczko S, Baltz J. (2023) Evaluation of nutrition components within prehabilitation programs in gastrointestinal cancers: Is prehab worth the hype? *Nutrition in clinical practice : official publication of the American Society for Parenteral and Enteral Nutrition*.

42. Zhou T, Yang K, Thapa S, et al. (2017) Differences in Symptom Burden Among Cancer Patients With Different Stages of Cachexia. *J Pain Symptom Manage* **53**:919-26.

43. Nakajima N. (2021) Differential Diagnosis of Cachexia and Refractory Cachexia and the Impact of Appropriate Nutritional Intervention for Cachexia on Survival in Terminal Cancer Patients. *Nutrients* **13**.

44. McLuskie A, Bowers M, Bayly J, et al. (2025) Nutritional interventions in randomised clinical trials for people with incurable solid cancer: A systematic review. *Clinical nutrition* **44**:201-19.

45. Baguley BJ, Edbrooke L, Denehy L, et al. (2024) A rapid review of nutrition and exercise approaches to managing unintentional weight loss, muscle loss, and malnutrition in cancer. *The oncologist*.

46. Muscaritoli M, Imbimbo G, Jager-Wittenaar H, et al. (2023) Disease-related malnutrition with inflammation and cachexia. *Clinical nutrition* **42**:1475-9.

47. Penet MF, Bhujwalla ZM. (2015) Cancer cachexia, recent advances, and future directions. *Cancer journal* **21**:117-22.

48. Roeland EJ, Bohlke K, Baracos VE, et al. (2020) Management of Cancer Cachexia: ASCO Guideline. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **38**:2438-53.

49. Evans WJ, Morley JE, Argiles J, et al. (2008) Cachexia: a new definition. *Clinical nutrition* **27**:793-9.

50. Argiles JM, Lopez-Soriano FJ, Toledo M, et al. (2011) The cachexia score (CASCO): a new tool for staging cachectic cancer patients. *Journal of cachexia, sarcopenia and muscle* **2**:87-93.

51. Schuetz P, Fehr R, Baechli V, et al. (2019) Individualised nutritional support in medical inpatients at nutritional risk: a randomised clinical trial. *Lancet* **393**:2312-21.

52. Wen X, Wang M, Jiang CM, et al. (2011) Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. *Asia Pacific journal of clinical nutrition* **20**:551-6.

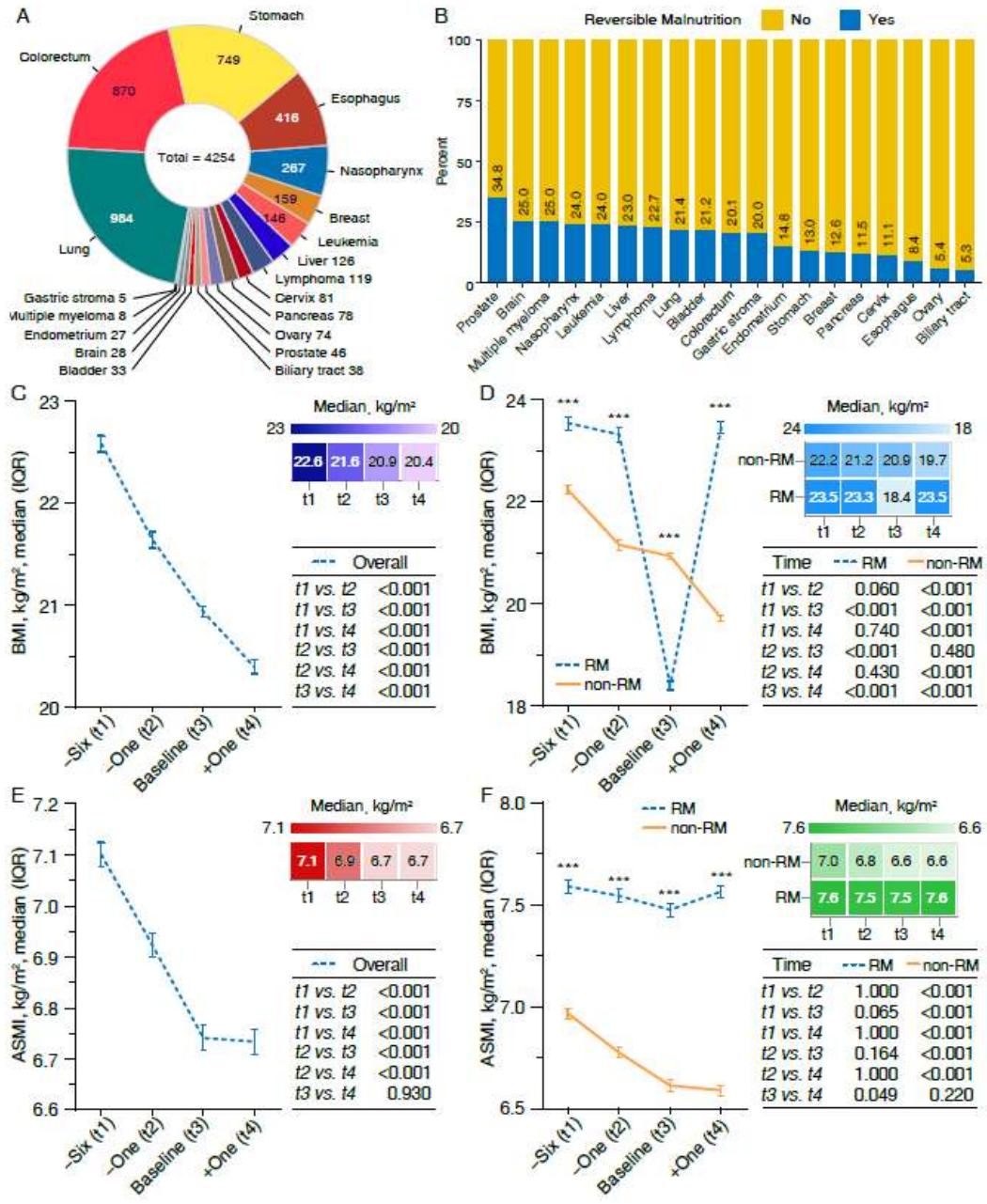

Figure 1. Prevalence of reversible malnutrition and analysis on the dynamics of body weight and muscle. BMI, body mass index; IQR, interquartile range; RM, reversible malnutrition; −Six (t1), six months before baseline; −One (t2), one month before baseline; baseline (t3), upon patient admission; +One (t4), one month after baseline; ASMI, appendicular skeletal muscle mass index; ***, $P < 0.001$. (A) Number of cases for each cancer site. (B) Cancer-specific prevalence of reversible malnutrition. (C) Body mass index dynamics across the four time points. (D) Body mass index dynamics across the four time points stratified by the RM status. (E) Appendicular skeletal muscle index dynamics across the four time points. (F) Appendicular skeletal muscle index dynamics across the four time points stratified by the RM status.

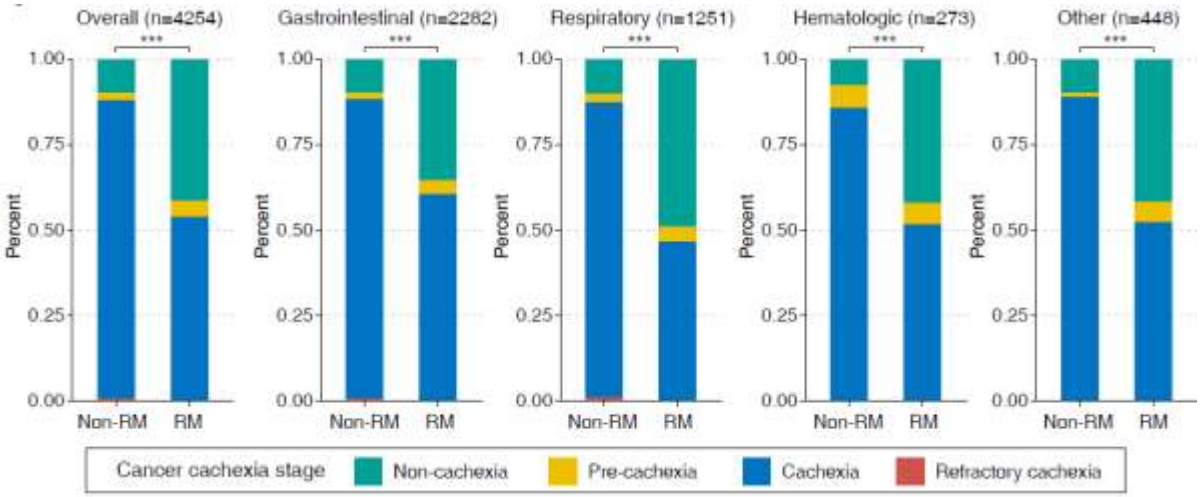

Figure 2. Association between reversible malnutrition and stages of cancer cachexia in the overall population and in subgroups of cancer sites. RM, reversible cachexia; ***, $P < 0.001$.

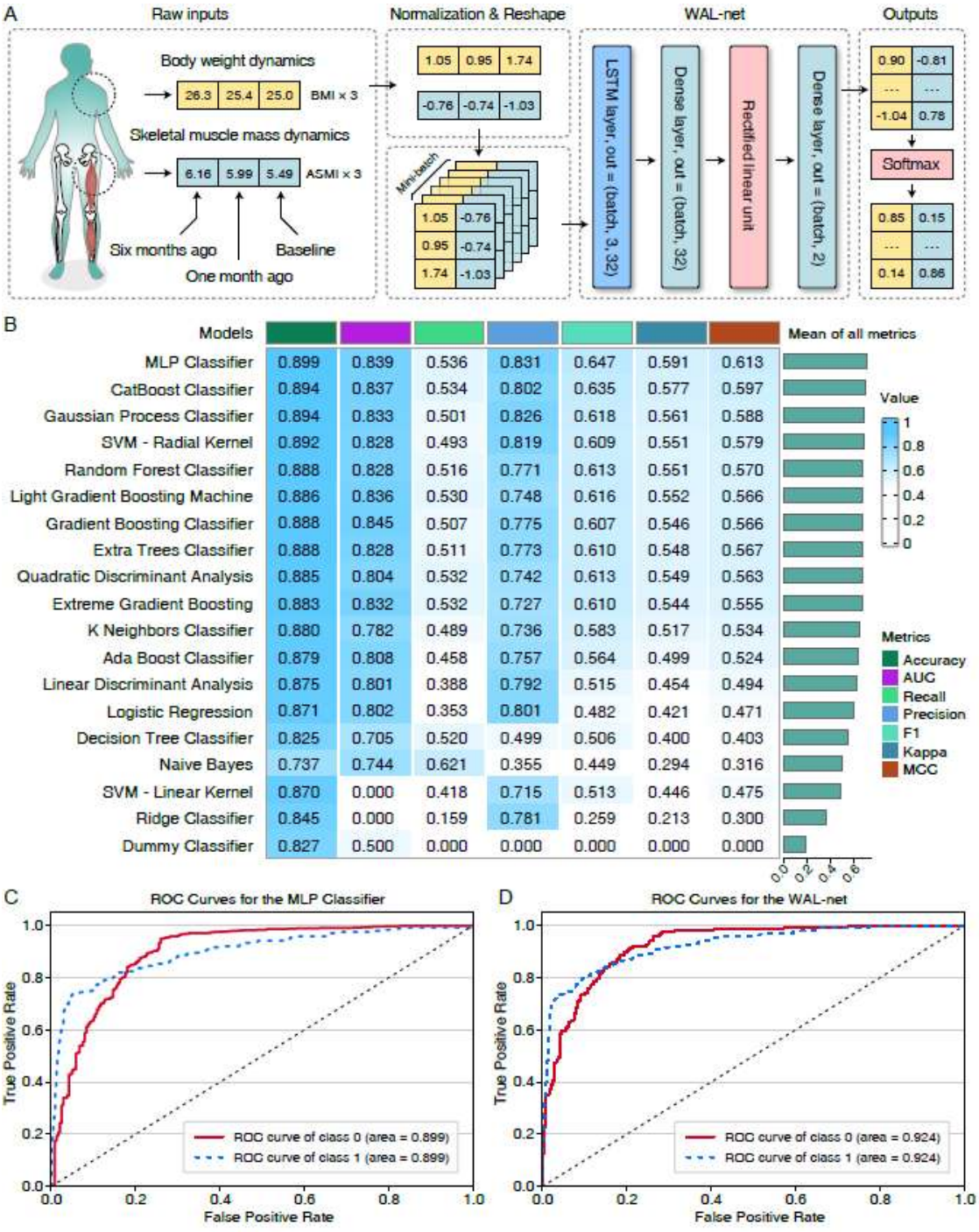

Figure 3. WAL-net and analyses on its superiority. BMI, body mass index; ASMI, appendicular skeletal muscle index; LSTM, long short-term memory; WAL, weight, appendicular skeletal muscle and LSTM; MLP, multilayer perceptron; SVM, supportive vector machine; AUC, area under the curve; MCC, Matthews correlation coefficient; ROC, receiver operating characteristic. (A) A graphical workflow of the WAL-net. (B) Cross-validated results of machine learning models to predict reversible malnutrition in the discovery data. (C) ROC curves for the MLP model in the test data. (D) ROC curves for the WAL-net in the test data.

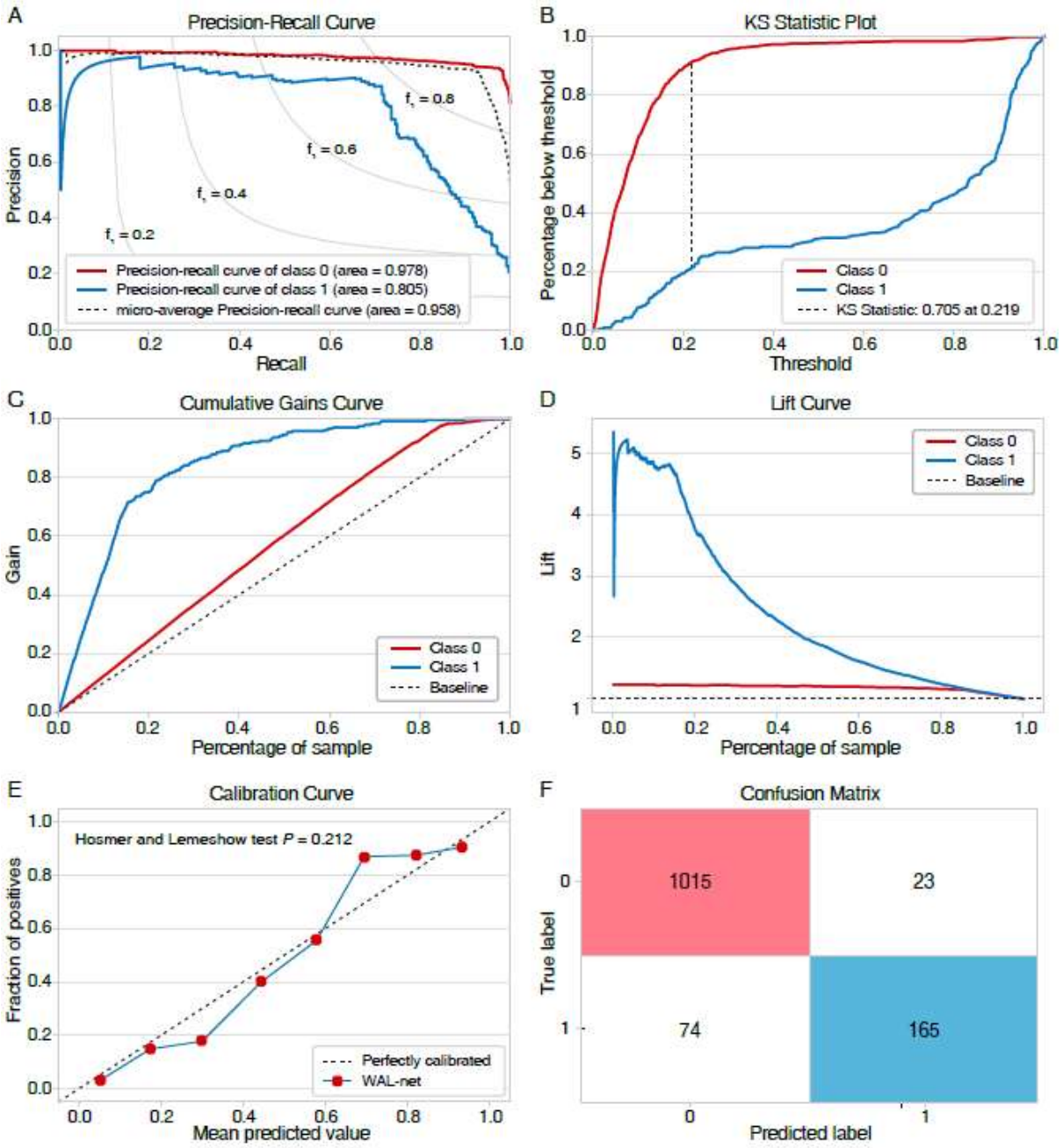

Figure 4. Further evaluation of the WAL-net in holdout test data. (A) Precision-recall curve. (B) Kolmogorov Smirnov statistic plot. (C) Cumulative gains curve. (D) Lift curve. (E) Calibration curve. (F) Confusion matrix.

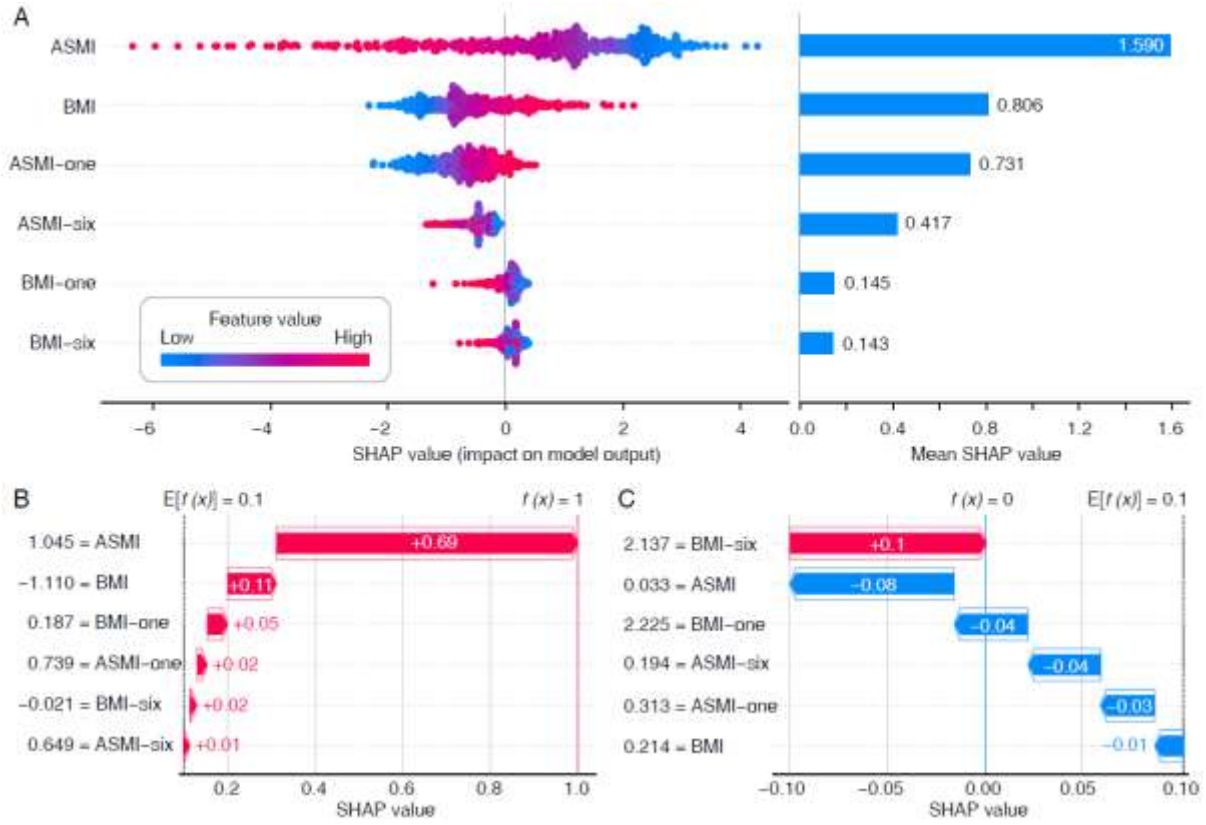

Figure 5. Interpretability analysis of the WAL-net using the shapley additive explanations (SHAP) method. ASMI, appendicular skeletal muscle mass index; BMI, body mass index; −One, one month before baseline; −Six, six months before baseline. (A) Model interpretability at the group level. (B) Model interpretability at the individual level (high probability). (C) Model interpretability at the individual level (low probability).

Table 1. General characteristics of the study population

Characteristics	Overall (n=4254)	Reversible malnutrition			Data split		
		No (n=3500)	Yes (n=754)	*P*	Discovery (n=2977)	Test (n=1277)	*P*
Age, years	61.0 [51.0, 68.0][1]	61.0 [52.0, 69.0]	59.0 [49.0, 68.0]	0.002	60.0 [51.0, 68.0]	61.0 [52.0, 69.0]	0.304
Sex, male	2783 (65.4)[2]	2163 (61.8)	620 (82.2)	<0.001	1960 (65.8)	823 (64.4)	0.402
Smoking, yes	2102 (49.4)	1677 (47.9)	425 (56.4)	<0.001	1472 (49.4)	630 (49.3)	0.974
Alcohol drinking, yes	941 (22.1)	763 (21.8)	178 (23.6)	0.300	644 (21.6)	297 (23.3)	0.258
Tea drinking, yes	1193 (28.0)	951 (27.2)	242 (32.1)	0.007	821 (27.6)	372 (29.1)	0.319
Residency, urban area	2716 (63.8)	2212 (63.2)	504 (66.8)	0.065	1912 (64.2)	804 (63.0)	0.452
Occupation, blue collar	1470 (34.6)	1259 (36.0)	211 (28.0)	<0.001	1034 (34.7)	436 (34.1)	0.737
Comorbidities							
Hypertension	753 (17.7)	607 (17.3)	146 (19.4)	0.206	534 (17.9)	219 (17.1)	0.566
Diabetes	363 (8.5)	284 (8.1)	79 (10.5)	0.042	260 (8.7)	103 (8.1)	0.513
Coronary heart disease	184 (4.3)	150 (4.3)	34 (4.5)	0.861	120 (4.0)	64 (5.0)	0.17

Chronic hepatitis	176 (4.1)	139 (4.0)	37 (4.9)	0.285	128 (4.3)	48 (3.8)	0.464
Anemia	126 (3.0)	102 (2.9)	24 (3.2)	0.782	90 (3.0)	36 (2.8)	0.794
Cancer site				<0.001			0.406
Lung	984 (23.1)	773 (22.1)	211 (28.0)		669 (22.5)	315 (24.7)	
Colorectum	870 (20.5)	695 (19.9)	175 (23.2)		607 (20.4)	263 (20.6)	
Stomach	749 (17.6)	652 (18.6)	97 (12.9)		537 (18.0)	212 (16.6)	
Esophagus	416 (9.8)	381 (10.9)	35 (4.6)		292 (9.8)	124 (9.7)	
Nasopharynx	267 (6.3)	203 (5.8)	64 (8.5)		192 (6.4)	75 (5.9)	
Breast	159 (3.7)	139 (4.0)	20 (2.7)		111 (3.7)	48 (3.8)	
Leukemia	146 (3.4)	111 (3.2)	35 (4.6)		116 (3.9)	30 (2.3)	
Liver	126 (3.0)	97 (2.8)	29 (3.8)		92 (3.1)	34 (2.7)	
Lymphoma	119 (2.8)	92 (2.6)	27 (3.6)		82 (2.8)	37 (2.9)	
Cervix	81 (1.9)	72 (2.1)	9 (1.2)		58 (1.9)	23 (1.8)	
Pancreas	78 (1.8)	69 (2.0)	9 (1.2)		53 (1.8)	25 (2.0)	
Ovary	74 (1.7)	70 (2.0)	4 (0.5)		49 (1.6)	25 (2.0)	
Prostate	46 (1.1)	30 (0.9)	16 (2.1)		29 (1.0)	17 (1.3)	
Biliary tract	38 (0.9)	36 (1.0)	2 (0.3)		26 (0.9)	12 (0.9)	
Bladder	33 (0.8)	26 (0.7)	7 (0.9)		18 (0.6)	15 (1.2)	
Brain	28 (0.7)	21 (0.6)	7 (0.9)		22 (0.7)	6 (0.5)	
Endometrium	27 (0.6)	23 (0.7)	4 (0.5)		16 (0.5)	11 (0.9)	
Multiple myeloma	8 (0.2)	6 (0.2)	2 (0.3)		5 (0.2)	3 (0.2)	

Gastric stroma	5 (0.1)	4 (0.1)	1 (0.1)		3 (0.1)	2 (0.2)	
Clinical stage[3]				0.464			0.176
I	318 (8.0)	261 (7.9)	57 (8.3)		227 (8.2)	91 (7.5)	
II	1291 (32.4)	1068 (32.5)	223 (32.3)		874 (31.5)	417 (34.5)	
III	1241 (31.2)	1012 (30.8)	229 (33.2)		888 (32.0)	353 (29.2)	
IV	1131 (28.4)	950 (28.9)	181 (26.2)		785 (28.3)	346 (28.7)	
Anti-cancer therapy							
Curative surgery	1709 (40.2)	1408 (40.2)	301 (39.9)	0.908	1160 (39.0)	549 (43.0)	0.015
Curative radiotherapy	243 (5.7)	216 (6.2)	27 (3.6)	0.007	167 (5.6)	76 (6.0)	0.713
Curative chemotherapy	827 (19.4)	659 (18.8)	168 (22.3)	0.034	570 (19.1)	257 (20.1)	0.486
Adjuvant chemotherapy	846 (19.9)	688 (19.7)	158 (21.0)	0.448	575 (19.3)	271 (21.2)	0.166
Chemotherapy for metastasis	304 (7.1)	246 (7.0)	58 (7.7)	0.573	214 (7.2)	90 (7.0)	0.922
Targeted therapy	164 (3.9)	134 (3.8)	30 (4.0)	0.928	106 (3.6)	58 (4.5)	0.151
Immunotherapy	113 (2.7)	90 (2.6)	23 (3.1)	0.537	78 (2.6)	35 (2.7)	0.904
Symptomatic therapy	2048 (48.1)	1670 (47.7)	378 (50.1)	0.244	1443 (48.5)	605 (47.4)	0.534
Nutritional support, yes	1571 (36.9)	1334 (38.1)	237 (31.4)	0.001	1099 (36.9)	472 (37.0)	1.00

							0
Nutritional support, type				0.002			0.473
EN only	392 (9.2)	331 (9.5)	61 (8.1)		281 (9.4)	111 (8.7)	
PN only	741 (17.4)	620 (17.7)	121 (16.0)		504 (16.9)	237 (18.6)	
PN+EN	438 (10.3)	383 (10.9)	55 (7.3)		314 (10.5)	124 (9.7)	
No nutrition support	2683 (63.1)	2166 (61.9)	517 (68.6)		1878 (63.1)	805 (63.0)	
Laboratory indices							
Total protein, g/L	67.8 [62.7, 72.8]	67.6 [62.7, 72.6]	68.6 [63.0, 73.5]	0.010	67.7 [62.8, 72.8]	67.9 [62.7, 72.8]	0.961
Creatinine, mmol/L	68.0 [56.5, 81.0]	66.9 [55.6, 80.0]	72.2 [60.9, 85.0]	<0.001	68.0 [56.7, 81.0]	68.0 [56.1, 80.9]	0.896
Albumin, g/L	38.2 [34.5, 41.9]	38.0 [34.2, 41.5]	39.6 [35.6, 43.1]	<0.001	38.2 [34.4, 41.9]	38.4 [34.5, 42.0]	0.529
Urea nitrogen, mmol/L	5.0 [3.9, 6.2]	4.9 [3.8, 6.2]	5.2 [4.1, 6.3]	0.003	5.0 [3.9, 6.2]	5.0 [3.9, 6.2]	0.931
Prealbumin, mg/L	206.2 [160.0, 250.0]	200.0 [154.0, 241.0]	223.0 [180.0, 270.0]	<0.001	207.4 [160.0, 250.0]	203.0 [160.0, 250.0]	0.723
Total bilirubin, μmol/L	10.7 [7.8, 14.6]	10.7 [7.8, 14.6]	10.9 [8.1, 14.7]	0.150	10.6 [7.8, 14.5]	11.1 [7.8, 15.0]	0.158
Direct bilirubin, μmol/L	3.2 [2.2, 4.5]	3.2 [2.2, 4.5]	3.1 [2.3, 4.4]	0.608	3.1 [2.2, 4.4]	3.2 [2.3, 4.5]	0.087
Transferrin, g/L	2.2 [1.8, 2.6]	2.2 [1.8, 2.6]	2.2 [1.9, 2.7]	0.078	2.2 [1.8, 2.6]	2.2 [1.8, 2.6]	0.644
Cholesterol, mmol/L	4.4 [3.8, 5.1]	4.4 [3.8, 5.1]	4.4 [3.8, 5.1]	0.143	4.4 [3.8, 5.1]	4.4 [3.8, 5.0]	0.89

Glucose, mmol/L	5.2 [4.7, 5.8]	5.1 [4.6, 5.8]	5.3 [4.8, 6.0]	<0.001	5.2 [4.7, 5.8]	5.2 [4.7, 5.9]	1 0.675
Triglycerides, mmol/L	1.2 [0.9, 1.6]	1.2 [0.9, 1.6]	1.2 [1.0, 1.7]	<0.001	1.2 [0.9, 1.6]	1.2 [0.9, 1.6]	0.068
ALT, U/L	18.0 [12.0, 28.9]	17.5 [12.0, 28.0]	20.0 [13.0, 31.6]	<0.001	17.8 [12.0, 28.6]	18.0 [12.1, 29.0]	0.256
AST, U/L	21.0 [16.8, 28.5]	21.0 [16.8, 28.5]	21.0 [17.0, 28.1]	0.874	21.0 [16.6, 28.3]	21.5 [17.0, 28.5]	0.107
HDL, mmol/L	1.2 [1.0, 1.4]	1.2 [1.0, 1.4]	1.1 [1.0, 1.3]	0.133	1.2 [1.0, 1.4]	1.2 [1.0, 1.4]	0.241
LDL, mmol/L	2.7 [2.2, 3.2]	2.7 [2.2, 3.2]	2.8 [2.3, 3.3]	0.083	2.7 [2.2, 3.2]	2.7 [2.2, 3.2]	0.889
Hemoglobin, g/L	123.0 [107.0, 137.0]	122.0 [107.0, 136.0]	130.0 [113.0, 144.0]	<0.001	123.0 [107.0, 137.0]	124.0 [108.0, 138.0]	0.258
White blood cells, $\times 10^9$/L	6.0 [4.6, 7.8]	6.0 [4.6, 7.9]	6.0 [4.7, 7.7]	0.940	6.0 [4.6, 7.9]	6.0 [4.6, 7.7]	0.606
Neutrophils, $\times 10^9$/L	3.7 [2.5, 5.5]	3.7 [2.5, 5.6]	3.7 [2.7, 5.3]	0.544	3.7 [2.5, 5.5]	3.8 [2.6, 5.5]	0.379
Lymphocytes, $\times 10^9$/L	1.4 [1.0, 1.9]	1.4 [1.0, 1.9]	1.5 [1.1, 2.0]	0.001	1.4 [1.0, 1.9]	1.4 [1.0, 1.8]	0.001
Red blood cells, $\times 10^{12}$/L	4.2 [3.7, 4.6]	4.1 [3.6, 4.6]	4.4 [3.8, 4.8]	<0.001	4.2 [3.7, 4.6]	4.2 [3.7, 4.6]	0.378
Platelets, $\times 10^9$/L	222.0 [169.0, 289.0]	226.0 [171.0, 292.0]	206.5 [161.0, 267.0]	<0.001	223.0 [169.0, 290.0]	222.0 [168.0, 284.0]	0.569

C-reactive protein, mg/L	4.0 [2.7, 20.7]	4.2 [2.9, 21.8]	3.2 [2.1, 13.4]	0.001	4.0 [2.6, 21.1]	4.0 [3.0, 19.4]	0.866
NLR	2.6 [1.6, 4.4]	2.7 [1.7, 4.5]	2.4 [1.6, 3.9]	0.005	2.6 [1.6, 4.3]	2.8 [1.7, 4.6]	0.004
Anthropometric indices							
Body height, cm	165.0 [159.0, 172.0]	165.0 [158.0, 170.0]	176.0 [168.0, 178.0]	<0.001	166.0 [159.0, 172.0]	165.0 [159.0, 173.0]	0.734
Weight loss, one month, %	5.3 [0.0, 8.0]	5.9 [0.0, 8.5]	0.0 [0.0, 4.5]	<0.001	5.3 [0.0, 8.1]	5.1 [0.0, 7.9]	0.304
Weight loss, six months, %	7.7 [2.5, 12.9]	8.7 [4.3, 13.6]	2.2 [0.0, 6.5]	<0.001	7.7 [2.6, 13.1]	7.7 [2.2, 12.3]	0.161
MAC, cm	25.0 [23.0, 27.0]	24.5 [22.5, 27.0]	27.0 [25.0, 29.0]	<0.001	25.0 [23.0, 27.0]	25.0 [23.0, 27.1]	0.412
TSF, mm	12.0 [8.0, 17.0]	11.0 [8.0, 16.9]	13.2 [10.0, 18.0]	<0.001	12.0 [8.0, 17.0]	12.0 [8.0, 18.0]	0.166
Handgrip strength, kg	23.2 [17.0, 30.3]	22.1 [16.3, 28.9]	28.9 [21.5, 36.9]	<0.001	23.3 [17.1, 30.3]	22.9 [16.8, 30.4]	0.396
MAMC, cm	21.0 [18.9, 23.1]	20.6 [18.7, 22.6]	22.6 [20.5, 24.6]	<0.001	21.0 [18.9, 23.1]	21.0 [18.9, 23.0]	0.978
Calf circumference, cm	32.0 [29.5, 34.0]	31.0 [29.0, 33.6]	34.0 [32.0, 36.7]	<0.001	32.0 [29.5, 34.0]	32.0 [29.2, 34.2]	0.667
BMI−six, kg/m^2	22.6 [20.1, 25.0]	22.2 [19.8, 24.8]	23.5 [21.7, 25.2]	<0.001	22.5 [20.1, 25.0]	22.6 [20.2, 24.9]	0.536
BMI−one, kg/m^2	21.6 [19.1, 24.5]	21.2 [18.7, 24.2]	23.3 [21.5, 25.2]	<0.001	21.6 [19.0, 24.5]	21.9 [19.2, 24.5]	0.146

BMI, kg/m^2	20.9 [19.3, 22.5]	20.9 [19.7, 22.8]	18.4 [18.0, 20.2]	<0.001	20.7 [19.3, 22.5]	20.9 [19.0, 22.5]	0.734
BMI+one, kg/m^2	20.4 [18.4, 23.0]	19.7 [18.0, 22.0]	23.5 [21.8, 25.2]	<0.001	20.3 [18.3, 22.9]	20.7 [18.4, 23.0]	0.148
ASMI−six, kg/m^2	7.1 [6.2, 7.7]	7.0 [6.1, 7.7]	7.6 [7.1, 8.0]	<0.001	7.1 [6.3, 7.7]	7.1 [6.2, 7.7]	0.761
ASMI−one, kg/m^2	6.9 [6.1, 7.7]	6.8 [5.9, 7.5]	7.5 [7.2, 8.0]	<0.001	6.9 [6.1, 7.7]	6.9 [6.1, 7.7]	0.896
ASMI, kg/m^2	6.7 [5.8, 7.4]	6.6 [5.6, 7.2]	7.5 [7.0, 8.0]	<0.001	6.7 [5.8, 7.4]	6.8 [5.7, 7.4]	0.749
ASMI+one, kg/m^2	6.7 [5.8, 7.4]	6.6 [5.6, 7.1]	7.6 [7.2, 8.0]	<0.001	6.7 [5.8, 7.4]	6.7 [5.8, 7.4]	0.949
Food intake				<0.001			0.541
Normal	1449 (34.1)	1072 (30.6)	377 (50.0)		1018 (34.2)	431 (33.8)	
Slightly reduced (25%-50%)	1667 (39.2)	1418 (40.5)	249 (33.0)		1177 (39.5)	490 (38.4)	
Severely reduced (>50%)	1138 (26.8)	1010 (28.9)	128 (17.0)		782 (26.3)	356 (27.9)	
GI symptoms							
Anorexia	1109 (26.1)	990 (28.3)	119 (15.8)	<0.001	772 (25.9)	337 (26.4)	0.784
Nausea	534 (12.6)	482 (13.8)	52 (6.9)	<0.001	354 (11.9)	180 (14.1)	0.053
Vomiting	350 (8.2)	321 (9.2)	29 (3.8)	<0.001	230 (7.7)	120 (9.4)	0.079

Mouth sores	69 (1.6)	56 (1.6)	13 (1.7)	0.932	50 (1.7)	19 (1.5)	0.748
Constipation	507 (11.9)	441 (12.6)	66 (8.8)	0.004	359 (12.1)	148 (11.6)	0.703
Diarrhea	241 (5.7)	202 (5.8)	39 (5.2)	0.576	162 (5.4)	79 (6.2)	0.373
Dry mouth	476 (11.2)	430 (12.3)	46 (6.1)	<0.001	324 (10.9)	152 (11.9)	0.361
Taste changes	329 (7.7)	293 (8.4)	36 (4.8)	0.001	218 (7.3)	111 (8.7)	0.142
Smell changes	162 (3.8)	149 (4.3)	13 (1.7)	0.001	105 (3.5)	57 (4.5)	0.169
Dysphagia	356 (8.4)	326 (9.3)	30 (4.0)	<0.001	239 (8.0)	117 (9.2)	0.245
Early satiety	390 (9.2)	354 (10.1)	36 (4.8)	<0.001	280 (9.4)	110 (8.6)	0.446
Pain	394 (9.3)	337 (9.6)	57 (7.6)	0.088	275 (9.2)	119 (9.3)	0.979
Other	90 (2.1)	79 (2.3)	11 (1.5)	0.214	60 (2.0)	30 (2.3)	0.564
ECOG performance status score				<0.001			0.028
0, fully active	2396 (56.3)	1906 (54.5)	490 (65.0)		1694 (56.9)	702 (55.0)	
1, slightly restricted	1396 (32.8)	1183 (33.8)	213 (28.2)		980 (32.9)	416 (32.6)	
2, moderately restricted	239 (5.6)	208 (5.9)	31 (4.1)		168 (5.6)	71 (5.6)	

3, severely restricted	176 (4.1)	161 (4.6)	15 (2.0)		109 (3.7)	67 (5.2)	
4, completely disabled	47 (1.1)	42 (1.2)	5 (0.7)		26 (0.9)	21 (1.6)	
ECOG performance status, =0	2396 (56.3)	1906 (54.5)	490 (65.0)	<0.001	1694 (56.9)	702 (55.0)	0.259
PG-SGA score, continuous	8.0 [5.0, 11.0]	9.0 [6.0, 12.0]	5.0 [2.0, 8.0]	<0.001	8.0 [5.0, 11.0]	8.0 [5.0, 12.0]	0.369
Global QoL score[4]	66.7 [50.0, 75.0]	66.7 [50.0, 75.0]	66.7 [50.0, 83.3]	<0.001	66.7 [50.0, 75.0]	66.7 [50.0, 75.0]	0.932

Abbreviations: EN, enteral nutrition support; PN, parenteral nutrition support; ALT, Alanine transaminase; AST, Alkaline phosphatase; HDL, high density lipoprotein; LDL low density lipoprotein; NLR, neutrophil to lymphocyte ratio; MAC, mid-arm circumference; TSF, triceps skinfold thickness; MAMC, mid-arm muscle circumference; BMI, body mass index; −six, six months before admission; −one, one month before admission; +one, one month after admission; ASMI, appendicular skeletal muscle mass index; GI, gastrointestinal; ECOG, the Eastern Cooperative Oncology Group; PG-SGA, the patient generated subjective global assessment; QoL, quality of life.

[1]Median [interquartile range], all such values, compared using Wilcoxon's rank-sum test.

[2]Number (percentage), all such values, compared using a Chi-squared test.

[3]Hematological malignancies are not included for clinical staging.

[4]Assessed using the European Organization for Research and Treatment of Cancer QLQ-C30 scale (QLQ-C30). The global QoL scale in QLQ-C30 was used with a higher score indicating a better global QoL.

Table 2. Overall and subgroup model performance in the holdout test set

Group	n	AUC (95%CI)	Accuracy (95%CI)	Kappa	Sensitivity	Specificity	PPV	NPV
Overall, WAL-net[1]	1277	0.924 (0.904, 0.944)	0.924 (0.908, 0.938)	0.728	0.878	0.932	0.690	0.978
Overall, MLP[1]	1277	0.899 (0.874, 0.924)	0.908 (0.890, 0.923)	0.660	0.850	0.917	0.615	0.975
Age, years								
< 60	586	0.918 (0.889, 0.947)	0.915 (0.889, 0.936)	0.710	0.919	0.914	0.648	0.985
≥ 60	691	0.929 (0.900, 0.957)	0.932 (0.911, 0.950)	0.745	0.843	0.947	0.735	0.972
Sex								
Female	454	0.764 (0.684, 0.843)	0.925 (0.897, 0.948)	0.140	1.000	0.925	0.081	1.000
Male	823	0.949 (0.932, 0.966)	0.924 (0.903, 0.941)	0.787	0.876	0.937	0.802	0.963
Cancer site[2]								
Gastrointestinal	672	0.915 (0.884, 0.947)	0.923 (0.900, 0.942)	0.699	0.893	0.927	0.636	0.984
Respiratory	390	0.941 (0.915, 0.966)	0.918 (0.886, 0.943)	0.760	0.852	0.935	0.775	0.960
Hematological	70	0.939 (0.880, 0.999)	0.914 (0.823, 0.968)	0.699	0.818	0.932	0.692	0.965
Other sites	145	0.887 (0.794, 0.980)	0.952 (0.903, 0.980)	0.749	1.000	0.947	0.632	1.000
Clinical stage								
I-II	557	0.890 (0.852, 0.928)	0.910 (0.883, 0.933)	0.658	0.847	0.920	0.610	0.976
III-IV	720	0.947 (0.925, 0.968)	0.935 (0.914, 0.952)	0.776	0.897	0.942	0.748	0.979
Curative surgery								
Yes	549	0.923 (0.892, 0.954)	0.920 (0.894, 0.941)	0.730	0.895	0.924	0.688	0.979
No	728	0.925 (0.899, 0.951)	0.927 (0.906, 0.945)	0.726	0.863	0.938	0.693	0.977
Curative chemotherapy								
Yes	257	0.936 (0.902, 0.969)	0.922 (0.882, 0.952)	0.752	0.870	0.934	0.741	0.970
No	1020	0.921 (0.897, 0.945)	0.925 (0.907, 0.940)	0.720	0.880	0.932	0.676	0.980

Nutritional support								
EN only	111	0.980 (0.954, 1.000)	0.946 (0.886, 0.980)	0.836	0.952	0.944	0.800	0.988
PN only	237	0.906 (0.847, 0.964)	0.916 (0.873, 0.948)	0.694	0.853	0.926	0.659	0.974
PN+EN	124	0.906 (0.813, 1.000)	0.944 (0.887, 0.977)	0.665	0.800	0.956	0.615	0.982
No nutritional support	805	0.922 (0.898, 0.946)	0.921 (0.900, 0.938)	0.724	0.878	0.928	0.688	0.977
C-reactive protein, mg/L								
≤ 10	839	0.936 (0.914, 0.958)	0.927 (0.908, 0.944)	0.760	0.880	0.937	0.740	0.975
> 10	438	0.893 (0.850, 0.936)	0.918 (0.888, 0.942)	0.645	0.870	0.923	0.571	0.984
ECOG performance status score								
= 0, fully active	702	0.938 (0.917, 0.959)	0.919 (0.896, 0.938)	0.755	0.868	0.931	0.752	0.967
> 0, restricted or dead	575	0.896 (0.854, 0.937)	0.930 (0.907, 0.950)	0.664	0.904	0.933	0.573	0.990
PG-SGA score								
0-1	70	0.859 (0.773, 0.944)	0.771 (0.656, 0.863)	0.524	0.791	0.741	0.829	0.690
2-3	143	0.957 (0.921, 0.993)	0.944 (0.893, 0.976)	0.887	0.924	0.961	0.953	0.937
4-8	442	0.893 (0.850, 0.935)	0.905 (0.874, 0.931)	0.655	0.895	0.907	0.586	0.983
≥9	622	0.878 (0.827, 0.929)	0.950 (0.930, 0.966)	0.528	0.864	0.953	0.404	0.995

Abbreviations: AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; WAL, weight + appendicular skeletal muscle mass + long short-term memory; MLP, multilayer perceptron; PG-SGA, the patient generated subjective global assessment.

[1] WAL-net vs. MLP, Delong's test for AUCs, $P = 0.005$.

[2] Gastrointestinal system (colorectum, stomach, esophagus, liver, pancreas, biliary tract and gastric stroma), respiratory system (lung and nasopharynx), hematological system (leukemia, lymphoma, multiple myeloma), other sites (cervix, ovary, endometrium, prostate, bladder, breast and brain).

Table 3. Association of model prediction with short-term clinical outcomes in the holdout test set

Outcomes	Overall (n=1277)	Predicted reversible malnutrition			True reversible malnutrition		
		No (n=1089)	Yes (n=188)	*P*	No (n=1038)	Yes (n=239)	*P*
BMI+1, kg/m^2	20.7 [18.4, 23.0][1]	20.0 [18.3, 22.3]	23.7 [22.0, 25.5]	<0.001	19.8 [18.1, 22.1]	23.5 [21.9, 25.3]	<0.001
BMI+1, underweight	388 (30.4)[2]	383 (35.2)	5 (2.7)	<0.001	388 (37.4)	0 (0.0)	<0.001
ASMI+1, kg/m^2	6.7 [5.8, 7.4]	6.6 [5.6, 7.1]	7.7 [7.4, 8.1]	<0.001	6.6 [5.6, 7.1]	7.6 [7.2, 8.0]	<0.001
ASMI+1, low	575 (45.0)	568 (52.2)	7 (3.7)	<0.001	575 (55.4)	0 (0.0)	<0.001
ECOG performance status score+1				0.019			0.001
0, fully active	274 (21.5)	218 (20.0)	56 (29.8)		202 (19.5)	72 (30.1)	
1, slightly restricted	446 (34.9)	378 (34.7)	68 (36.2)		358 (34.5)	88 (36.8)	
2, moderately restricted	44 (3.4)	38 (3.5)	6 (3.2)		36 (3.5)	8 (3.3)	
3, severely restricted	22 (1.7)	18 (1.7)	4 (2.1)		19 (1.8)	3 (1.3)	
4, completely disabled	4 (0.3)	3 (0.3)	1 (0.5)		2 (0.2)	2 (0.8)	

5, dead	487 (38.1)	434 (39.9)	53 (28.2)		202 (19.5)	72 (30.1)	
ECOG performance status+1, =0	274 (21.5)	218 (20.0)	56 (29.8)	0.004	202 (19.5)	72 (30.1)	<0.001
Increased malnutrition phenotypes[3]	358 (28.0)	351 (32.2)	7 (3.7)	<0.001	358 (34.5)	0 (0.0)	<0.001

Abbreviations: BMI, body mass index; +1, one month after baseline; ASMI, appendicular skeletal muscle mass index; ECOG, the Eastern Cooperative Oncology Group.

[1] Median [interquartile range], all such values, compared using Wilcoxon's rank-sum test.

[2] Number (percentage), all such values, compared using a Chi-squared test.

[3] Defined as an increase in the number of positive malnutrition phenotypes (as diagnosed using the Global Leadership Initiative on Malnutrition framework) from baseline to one month after baseline (might be one to two, or two to three).