Original Paper

Predicting Cirrhotic Patients' Transition From Compensated to Decompensated Stage: A Graph-Based Multitask Deep Learning Method

Mohammadsadegh Abolhasani[1], BSc; Da Xu[2], PhD; Paul Jen-Hwa Hu[1], PhD; Ting-Shuo Huang[3,4,5], PhD

[1]Department of Operations and Information Systems, David Eccles School of Business, University of Utah, Salt Lake City, UT, United States
[2]Department of Marketing, Analytics, and Professional Sales, School of Business Administration, University of Mississippi, University, MS, United States
[3]Division of General Surgery, Department of Surgery, Jen-Ai Hospital, Taichung, Taiwan
[4]Department of Surgery, Keelung Chang Gung Memorial Hospital, Keelung, Taiwan
[5]Department of Chinese Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan

Corresponding Author:

Paul Jen-Hwa Hu, PhD
Department of Operations and Information Systems
David Eccles School of Business, University of Utah
1655 East Campus Center Drive
Salt Lake City, UT 84112
United States
Phone: 1 801-587-7785
Email: paul.hu@eccles.utah.edu

Abstract

Background: Cirrhosis is a chronic liver disease, manifested by fibrosis and regenerative nodules. It can be classified clinically as compensated or decompensated. Accurate predictions of patients' transition from compensated to decompensated cirrhosis are important but difficult because of the multifactorial nature of disease progression, comorbid conditions, and patient heterogeneity.

Objective: To effectively identify at-risk patients with cirrhosis and facilitate timely interventions across chronic diseases and patient populations, this study aimed to develop a deep learning–based method to estimate patient transition from compensated to decompensated cirrhosis. The proposed method seeks to provide superior predictive utilities and help physicians derive actionable clinical insights by modeling comorbidity relationships and symptom interdependencies that influence the progression of cirrhosis among patients.

Methods: The proposed method was guided by clinical knowledge, considers patient demographics and medications, and leverages a graph structure to capture important comorbidity interplays. This method was augmented with post hoc interpretability to help clinicians understand its predictions for clinical decision-making and patient management. Using clinical data from 29,915 patients with cirrhosis, we evaluated the proposed method in comparison with 12 prevalent methods.

Results: The proposed method attained an average area under the curve (AUC) of 0.871 (SD 0.004), *F*-measure of 0.678 (SD 0.010), recall of 0.741 (SD 0.022), and precision of 0.663 (SD 0.005). It showed a 6.09% improvement in AUC and a 13.95% improvement in *F*-measure over the best-performing benchmark method. Supported by post hoc interpretation analyses, this method can provide physicians with interpretable visualizations of the underlying comorbidity and symptom dynamics to better support their decision-making and adoption.

Conclusions: Empirical results underscore the value of demographic similarity, comorbidity graphs, and extended loss functions for identifying patients at higher risk of decompensation by capturing finer-grained information from essential patient features that tend to be overlooked by most existing methods. The proposed method not only supports physicians' clinical decision-making and patient management but can also enhance health care organizations' resource allocation efficiency and cost containment.

JMIR Med Inform 2026;14:e91693; doi: 10.2196/91693

Keywords: cirrhotic patient transition; decompensation symptoms; cirrhosis; multitask learning; graph networks; deep learning; post hoc interpretation

Introduction

Background

Cirrhosis is a chronic liver disease that features fibrosis and regenerative nodules [1]. It can be classified clinically as compensated or decompensated [2]. The former entails an early disease stage characterized by a significantly scarred liver that still functions sufficiently to perform its essential responsibilities, without overt symptoms [3]. The latter denotes an advanced stage and is manifested by substantially decreased functional capacity that often leads to serious prognoses [4]. Ascites, varices, hepatic encephalopathy (HE), and hepatorenal syndrome (HRS) are important symptoms that mark the onset of liver decompensation [5]. When patients deteriorate from the compensated to the decompensated stage, they face a much higher mortality rate and need substantial clinical resources. Volk et al [6] stressed the criticality of identifying patients at higher risk of decompensation, so that they can be monitored closely and receive timely interventions [7].

However, accurately estimating the risk that patients with cirrhosis likely deteriorate from the compensated to the decompensated stage is difficult, largely due to the inherently multifactorial and heterogeneous nature of liver cirrhosis progression. Effective predictions are further complicated by interrelated variables that pertain to liver function measures, lifestyle factors, comorbid conditions, and temporal disease dynamics [8]. Physicians tend to depend on invasive hepatic venous pressure gradient measures, which, however, are not conveniently available in many health care organizations [9,10]. Noninvasive tests and serum-based biomarkers are also accessible to physicians. For example, transient elastography (eg, FibroScan; Echosens) and other modalities (eg, acoustic radiation using impulse imaging and magnetic resonance elastography) can assess liver stiffness to estimate decompensation risk indirectly [9,11,12]. Most physicians often use score-based methods, such as the Child-Pugh score, to estimate a patient's likelihood of deterioration. These methods are usually restricted by their exclusion of essential etiologies and can benefit from being complemented by other methods or additional factors [13,14]. For example, other clinical measures that consider multiple recurrent events of different types in the patient history, such as disease burden score, can be incorporated as an outcome variable in regression models [15].

Data-driven predictive analytics is particularly promising and can complement score-based methods [16]. Toward that end, machine learning–based methods integrate structured, longitudinal clinical variables (eg, laboratory results and elastography measures) to capture complex relationships and interplays among different conditions and medications, which are not attainable by fixed-point, score-based methods [17]. Similarly, deep learning–based methods are suited for estimating patient outcomes and risks because they can model complex relationships and temporal dependencies for individual-level predictions, using clinical, biomedical, and time series patient data available in electronic health records (EHRs) [18]. A recent review of deep learning shows that these methods have become increasingly applied in clinical settings because of their inclusion of temporal intervals, multiple feature types, and larger sample sizes for better predictive performance [19]. However, deep learning–based methods also have their own challenges; for example, predictive models built from EHR data require bias mitigation and fairness evaluation to ensure equitable performance across patient subgroups. These challenges underscore the need for evaluation beyond prediction accuracy [20]. Overall, predictive analytics methods can leverage detailed patient data and harness algorithmic power to complement score-based methods.

Different data-driven predictive methods can be applied to estimate decompensation among patients with cirrhosis [21], but many appear limited in prediction effectiveness [22]. For example, most existing methods target a single binary outcome, such as compensated versus decompensated or presence versus absence of a specific symptom, and often ignore the multifaceted nature of disease progression that involves multiple, distinct decompensation symptoms [23]. This implies the clinical value of multitask learning for modeling multiple, interrelated patient outcomes simultaneously to better support downstream tasks and predictions, as exemplified by methods incorporating attention mechanisms [24]. In addition, many methods overlook demographic similarity, which can inform decompensation symptom estimates by leveraging individual risk profiles across different patient groups [25,26]. Nor do prevalent methods consider symptom interplays; for example, the onset of ascites might be associated with a patient's subsequent risk of HE or HRS. Modeling such interplays could provide a means to exploit valuable intersymptom relationships for more accurate patient risk estimates and richer clinical insights [27]. In a related sense, modeling symptom interplays can address the "black box" nature of deep learning and provide post hoc interpretability. To illustrate, patient-level comorbidity graphs allow for visualizations of the relationships among diseases to facilitate clinicians' comprehension, reasoning, and adoption.

Objective

The main prediction task was symptom-level decompensation risk. We estimated 4 distinct decompensation symptoms among patients within a 3-year horizon after their cirrhosis diagnosis: ascites, varices, HE, and HRS. A patient's transition from the compensated to the decompensated stage is manifested by the onset of 1 or more of these symptoms. Guided by clinical knowledge about cirrhosis and its progression, we developed a graph-based multitask deep learning method that considers patient demographics, etiological characteristics, medications, and interdisease interactions. This method also accounts for the relationship

of decompensation symptoms with mortality, the inclusion of which in the predictive model construction can augment the main prediction task because mortality is a critical patient outcome resulting from decompensation [28].

The proposed method includes mortality prediction as an auxiliary task during model training, which helps the method learn meaningful information of disease severity and adverse progression. Clinical occurrence of decompensation symptoms is highly associated with worsening prognosis, making mortality a useful complementary signal for training a predictive model, although it is not needed when applying the method to new patients (for decompensation symptom prediction). Incorporating this important patient outcome helps the model encode shared representations of risk across related clinical end points and can strengthen the identification of patients who are likely to develop decompensation symptoms. In this way, the auxiliary task enriches the model learning process by leveraging an essential and closely related outcome to reflect the disease progression more broadly.

Moreover, patients are stratified into cohorts based on their condition, and a comorbidity graph is constructed to encode clinically plausible interdisease relationships. Compared with existing methods, the proposed method is novel in 3 important aspects. First, it adopts a multitask framework and incorporates mortality prediction as an auxiliary task to better estimate distinct symptoms simultaneously using a graph neural network (GNN). Second, this method incorporates patient similarity to create different patient groups and constructs group-level graphs to model comorbidity interplays for predictive efficacy, in line with cirrhosis literature recognizing that similar patients tend to have more comparable outcomes than dissimilar patients [25]. Third, it integrates signals from individual- and group-level embeddings to account for both individual heterogeneity and cohort characteristics, so it can distinguish disease chronicity and further enhance predictions by considering temporal dynamics of disease. Furthermore, the proposed method offers post hoc interpretability to help clinicians understand its predictions for clinical decision-making and use.

This study underscores the advantages of graph-based modeling and multitask learning for capturing comorbidity interplays, patient demographics, and symptoms to better identify patients with cirrhosis at higher risk of decompensation in a timely manner. Evaluation results affirm the proposed method's superior predictive utilities over many prevalent methods, thus demonstrating the clinical value and feasibility of predicting multiple decompensation symptoms simultaneously. This method adds to predictive analytics research beyond a simplistic, binary view and enables nuanced risk stratification and personalized management of individual patients with a specific decompensation trajectory. The proposed method is generalizable and can be adapted to other chronic diseases featuring similar multifactorial progression and important symptom interplays that challenge analytics methods. Supported by timely and accurate decompensation estimates at the individual level, physicians can identify higher-risk patients for proactive management, therapeutic planning, and efficient resource allocation, with cost containment implications.

Related Work

Overview of Cirrhosis and Compensated Versus Decompensated Stage

Marked by irreversible liver scarring and progressive hepatic dysfunction, cirrhosis is a formidable challenge to patients with chronic liver disease [7,29]. It can be attributed to etiological factors that influence clinical conditions, disease progression or manifestation, and prognosis [30]. The main causes of cirrhosis are chronic infections of hepatitis B virus (HBV) and hepatitis C virus (HCV). In the United States, HCV-related cirrhosis has declined due to antiviral treatments, yet it still remains prevalent and accounts for 34% of cirrhosis cases [31]. Alarmingly, HBV is the leading cause of cirrhosis in Asia, accounting for approximately 60% of cases in China [32]. Cirrhosis requires timely identification and has important morbidity and mortality implications. In particular, patients with higher decompensation risk need close monitoring and early interventions to mitigate further deterioration and avoid extensive, costly therapeutic interventions such as liver transplantation [33].

Cirrhosis can be categorized as compensated or decompensated, which refer to a significantly scarred liver that retains essential functions without overt symptoms [34] and substantial declines of the liver capacity, respectively. At the decompensated stage, patients need extensive treatments and potential liver transplantation [35]. Beyond etiologies, patient demographics are also important. For example, male and female patients seem to exhibit distinct progression patterns [36], with men showing fast-paced decompensation and higher risk. Additionally, patients aged >60 years tend to have worse cirrhosis-related conditions than their younger counterparts [37].

Ascites, variceal bleeding, HE, and HRS represent important symptoms of decompensation [35,38]. These different symptoms manifest decompensation that requires specific interventions for patient care and management [21]. For example, ascites, which features a pathological accumulation of fluid within the peritoneal cavity [39], arises among approximately 50% of patients with cirrhosis in the compensated stage within 10 years of their cirrhosis diagnosis [39]. Varices entail dilated veins, typically in the esophagus, and often result from the increased portal hypertension associated with cirrhosis [40]. They significantly increase the risk of life-threatening hemorrhage, and such bleeding episodes lead to a high mortality rate [40]. In addition, HE features impaired brain function [41], although it is potentially reversible. Its clinical manifestation involves a spectrum of neurologic and psychiatric complications that range from minimal cognitive deficits to overt personality changes, disorientation, and coma [41]. While relatively unusual, HRS reflects renal impairment in patients with cirrhosis, characterized by substantial reduction in renal blood flow and glomerular filtration rates, without clinical evidence of intrinsic kidney disease [42]. As a result of functional

changes in renal circulation, HRS can potentially be reversed by interventions such as vasoconstrictor therapy and liver transplantation [42]. In summary, patients with cirrhosis can develop distinct symptoms over time, without any systematic, particular order [14].

Accurately estimating the risk that patients with cirrhosis likely develop distinct decompensation symptoms is crucial because each symptom has unique prognostic implications and requires differential care [43]. For example, ascites needs fluid and hemodynamic controls such as diuretics and paracentesis, while HE requires neurological and metabolic treatments such as lactulose and rifaximin [44-46]. Identifying patients at higher risk of a specific symptom, early in the disease progression, allows for timely intervention, efficient resource allocation, and improved patient outcomes [47]. Toward that end, Schneider et al [14] indicated that a 3-year time frame for deterioration estimates is clinically meaningful and appropriate for timely stratification that separates higher-risk and other patients, suggesting that higher-risk classification can preserve more than 95% sensitivity for decompensation. Interview results from 3 hepatologists confirmed the clinical value of a 3-year prediction time frame.

One hepatologist noted that "Liver decompensation marks a critical 'point of no return' in the disease trajectory, where liver transplantation often becomes the only curative option. A three-year prediction window provides the necessary lead time for the complex process of transplant evaluation, donor searching, and physiological optimization."

Another hepatologist agreed and indicated that "Predicting distinct decompensation events within a three-year window among compensated cirrhotic patients enables early identification of high-risk patients and facilitates preventive strategies to delay disease progression. It also supports personalized monitoring and management plans."

Data-Driven Analytics for Predicting Risk and Outcomes in Patients With Liver-Related Disease

The increasing availability of voluminous, detailed patient data fuels the development and use of data-driven predictive analytics for managing patients with liver-related diseases. Estimating the likelihood that each patient with cirrhosis will deteriorate from the compensated to the decompensated stage is critical. Supervised learning supports such predictions, typically viewing them as a classification task. According to their modeling specifics, existing methods can be classified as regression-based, tree-based, or neural network–based methods [48]. A regression-based method estimates patient risk numerically or dichotomously by specifying its relationships with select predictors, often in the form of a linear or logistic model [49]. For example, Cox regression models can predict the mortality of patients who have entered the terminal stage of their liver disease [50]. In general, regression-based methods are simplistic and interpretable, but they cannot handle complex relationships or patterns in patient data and are sensitive to outliers [51-53]. Tree-based methods use feature values to split patients into subsets, with the objective of classification according to the target variable value [54]. Exemplary methods include, but are not limited to, random forest (RF), AdaBoost (ADAB), and XGBoost (XGB). Wu et al [55] applied random forests to estimate fatty liver disease, and Li et al [56] used AdaBoost to detect liver disease among patients. Afreen et al [57] and Dalal et al [58] designed tree-based boosting methods for liver disease predictions. Despite their computational efficiency and interpretability, tree-based methods are not effective for nonlinear relationships and are prone to overfitting and biases in multilevel features [59,60].

Finally, neural network–based methods use feedforward, recurrent, or deep architectures to simulate interconnected neurons in the human brain. These methods are sophisticated and can capture complex relationships, patterns, and mechanisms in patient data [59]. For example, a fully connected feedforward neural network can estimate patient risks and outcomes because of its ability to model nonlinear relationships and flexible feature representations [61]. Multitask artificial neural networks have also been designed for clinical predictions [62]. In general, feedforward neural network–based methods require substantial computational processing, need large-scale data to train, and have an intrinsic "black box" nature [63,64]. Recurrent neural networks, such as long short-term memory (LSTM) and its variants, extend these methods for sequential and time series data analyses and can capture temporal dependencies [65]. As Liu et al [66] reported, LSTM is effective for estimating delirium onset in hospitalized patients. Shuvo et al [67] also used a multitask LSTM to estimate blood glucose concentrations. Moreover, deep neural networks (DNNs), exemplified by convolutional neural networks (CNNs) or temporal convolutional networks (TCNs), contain multiple hidden layers capable of learning hierarchical feature abstractions, so they can model complex relationships for estimating disease progression [68]. Singh et al [69] used DenseNet, a CNN variant, to estimate hepatic decompensation in patients with primary sclerosing cholangitis. Kok et al [70] used a TCN to estimate sepsis. Li et al [71] used a multitask CNN to achieve fairness in health care predictions, while Ramesh et al [72] designed a multitask TCN to recognize surgical phases in gastric bypass surgeries. Other methods, such as a multimodal spatiotemporal GNN [73], fuse longitudinal patient data and chest x-rays to model patient admission similarity in a graph and thereby estimate hospital readmissions. Multitask deep learning can integrate clinical variables and multiparametric magnetic resonance imaging scans to simultaneously estimate cancer recurrence (metastasis) versus disease-free survival among patients with rectal cancer [74]. Graph-based neural methods appear to outperform traditional tree-based methods in patient risk predictions, with interpretable insights into the relationships among risk factors [75]. Overall, DNNs offer effective predictions, but they need substantial data for training, require significant computing resources, and depend on regularization to avoid overfitting [76].

Gap Analysis

A review of existing analytics methods and related literature indicates 3 limitations that hinder effective estimates of the deterioration of patients with cirrhosis from the compensated to the decompensated stage. First, most existing methods consider a single, binary target variable, which restricts their use for stratifying patient risk for multiple, distinct symptoms to support personalized care and management [77-79]. Similarly, patients who likely develop decompensation symptoms have a markedly higher risk of death [30], suggesting that mortality could provide a valuable (severity) signal to train a model for decompensation predictions. Learning this auxiliary task in parallel, through shared representations, should enhance an analytics method's performance for the (main) symptom prediction task [80]. That is, interdependent complications and mortality can be considered jointly to better predict patient outcomes [81]. Second, relatively few methods consider demographic similarities among patients, which restricts their ability to capture essential patterns and temporal dynamics embedded in comorbidities and symptoms. Conceivably, patients sharing similar demographics and characteristics are more likely to have comparable prognoses and outcomes than dissimilar patients [36,37]. By considering such similarities, a predictive method can capture essential patterns to inform downstream decompensation estimates. Third, existing methods do not model complex relationships among comorbid conditions, differential comorbidity importance, or temporal chronicity. Consequently, they are not suited to the multifaceted nature of liver cirrhosis. Advanced frameworks that consider the underlying clinical mechanisms can enhance downstream predictions by recognizing and capturing how comorbidities and risk factors uniquely relate to decompensation risk [82-85]. However, existing methods tend to overlook comorbidity interplays that could increase patient decompensation risk. They also do not account for time-lagged, nonlinear impacts of preexisting comorbidities and conditions to inform the subsequent decompensation estimates [86].

To address these gaps and provide efficacious risk predictions, we developed a graph-based multitask deep learning method that adopts a multitask framework to estimate distinct outcomes simultaneously. As described previously, 4 decompensation symptoms constitute the main target variables, with mortality being an auxiliary task. With a GNN, our method learns representations across distinct symptoms and uses dedicated output heads for each symptom to capture both shared and task-specific features, thereby exploiting essential relationships of different symptoms for downstream predictions. In general, graphs can incorporate clinical insight into predictive models to increase diagnostic predictions and generate contextually relevant clinical insights using patient data in EHRs [87]. The proposed method constructs patient graphs based on demographic similarity and accounts for comorbidity interplays with respect to distinct decompensation symptoms. Learnable influence factors (IFs) can then be constructed from patient graphs to assess the impact of each impactful comorbidity on the onset of a specific symptom. Next, we used these factor values to compute condition impact scores (CIS) that quantify the severity of comorbidity interplays to inform decompensation risk estimates at both the group and individual levels. Moreover, the proposed method generates individual-level temporal graph embeddings to capture temporal dynamics of disease progression, which can enrich the node representation with essential progression information for efficacious predictions and clinical value. Table 1 compares representative methods for predicting hematology patient risks and outcomes, in comparison with the proposed method, to convey its differences and pertinent novelties.

Table 1. Representative learning methods included as benchmarks for patient risk predictions, relative to the proposed method.

Method	Approach	Focal task	Patient similarity	Disease relationship	Temporal graph embeddings	Comorbidity interplays
XGBoost [58]	Tree-based	Single task	No	No	No	No
AdaBoost [56]	Tree-based	Single task	No	No	No	No
Random forest [55]	Tree-based	Single task	No	No	No	No
Linear regression [49]	Regression based	Single task	No	No	No	No
Multi-FCN[a] [62]	Neural network	Multitask	No	Yes	No	No
Multi-CNN[b] [71]	Neural network	Multitask	No	Yes	No	No
Multi-TCN[c] [72]	Neural network	Multitask	No	Yes	No	No
Multi-LSTM[d] [67]	Neural network	Multitask	No	Yes	No	No
Single-FCN [61]	Neural network	Single task	No	Yes	No	No
Single-CNN [69]	Neural network	Single task	No	Yes	No	No
Single-TCN [70]	Neural network	Single task	No	Yes	No	No
Single-LSTM [66]	Neural network	Single task	No	Yes	No	No
Parallel CNN+Feature Nets [74]	Multimodal deep neural network	Multitask	No	Yes	No	No
MM-STGNN[e] [73]	Multimodal graph neural network	Single task	Yes	Yes	Yes	No

Method	Approach	Focal task	Patient similarity	Disease relationship	Temporal graph embeddings	Comorbidity interplays
Proposed method	Graph- and attention-based DNN[f]	Multitask with specialized loss	Yes	Yes	Yes, individual level	Yes

[a]FCN: fully connected network.
[b]CNN: convolutional neural network.
[c]TCN: temporal convolutional network.
[d]LSTM: long short-term memory.
[e]MM-STGNN: multimodal spatiotemporal graph neural network.
[f]DNN: deep neural network.

Methods

Problem Formulation

Let $P=\{P_1, P_2, ..., P_i\}$ be a set of patients and $D=\{ < d_{i,1}, \delta_{i,1} > , < d_{i,2}, \delta_{i,2} > , ..., < d_{i,D}, \delta_{i,D} > \}$ be a set of diseases (comorbidities), where $d_{i,j}$ is the binary indicator of morbidity j for patient P_i, and $\delta_{i,j}$ is the temporal difference (time interval) between the onset of that comorbidity and the cirrhosis diagnosis. $M=\{ < med_{i,1}, dosage_{i,1}, dur_{i,1}, S_{i,1} > , < med_{i,2}, dosage_{i,2}, dur_{i,2}, S_{i,2} > , ..., < med_{i,M}, dosage_{i,M}, dur_{i,M}, S_{i,M} > \}$ represents medication information, where $med_{i,m} \in \{0,1\}$ indicates whether a medication m is taken by patient P_i. If $med_{i,m}=1$, we record dosage $dosage_{i,m}$, duration $dur_{i,m}$, and the starting date $S_{i,m}$; otherwise, $dosage_{i,m}$, $dur_{i,m}$, and $S_{i,m}$ are set to 0. AGE_i and SEX_i denote patient age and gender, respectively. $Y_{i,k}, k \in \{1...5\}$ are binary indicators of decompensation symptoms ($k=1, 2, 3, 4$) and death ($k=5$) for patient P_i.

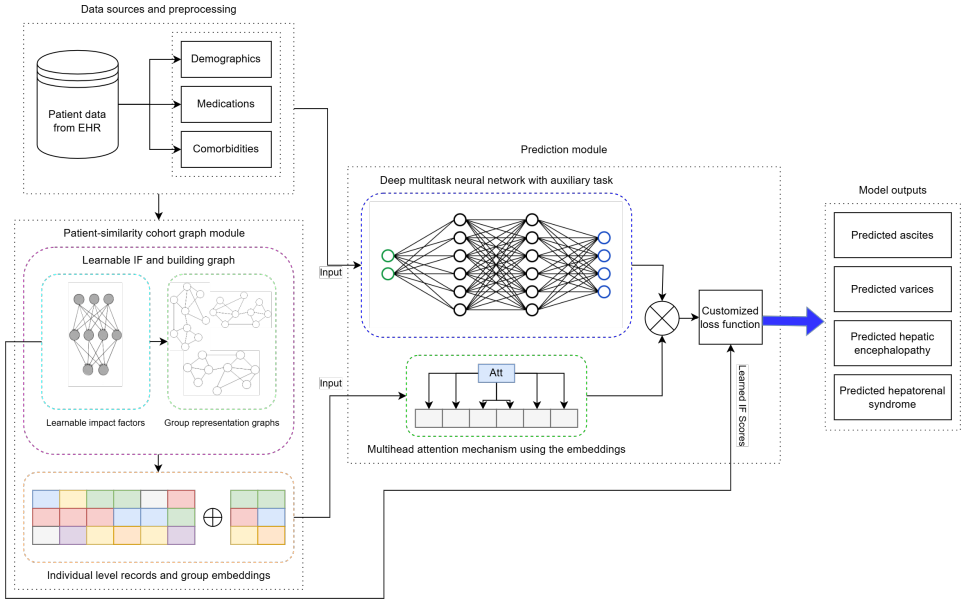
The main prediction task is formally defined as follows: *Given patient comorbidities D, medications M, and demographics (AGE and SEX), a predictive method learns patient embeddings to estimate symptom-specific decompensation risks (ascites, varices, HE, and HRS) within a 3-year window after the cirrhosis diagnosis*. For each patient, a prediction is made on the date of cirrhosis diagnosis. That is, all predictor values (demographics, comorbidities, and medication variables) are extracted from patient data on or before the cirrhosis diagnosis date. Outcomes indicate whether distinct decompensation events (ascites, varices, HE, and HRS) occurred within a 3-year window after the cirrhosis diagnosis.

The proposed method makes use of an auxiliary task, *mortality estimate*, for symptom predictions. Ascites, varices, HE, and HRS mark the onset of decompensation; death is a critical event in a patient's trajectory. Consideration of mortality as an auxiliary task provides a severity signal that regularizes the shared representation. It also allows the method to learn essential feature patterns and symptom constellations for both timely complication and lethality estimates, thereby increasing its ability to distinguish milder from rapidly progressive deteriorations.

Architectural Framework

Figure 1 presents the architectural framework of the proposed method, which has 3 important components: group-specific comorbidity graph construction, group- and individual-level graph representation learning, and multitask estimates. The group-specific comorbidity graph construction stratifies patients by demographics, such as age, gender, and liver condition indicators. For each patient group, it builds a disease relationship graph in which edge weights reflect the group's mutual comorbidity co-inference. Patient-level graph representation learning obtains graph embeddings from each group-specific graph and personalizes them at the individual level by integrating each patient's temporal comorbidity. The multitask estimates component then uses an attention mechanism to feed the personalized embeddings as key and value vectors into a multihead attention to perform the predictions. We describe each key component next.

Figure 1. Architectural framework of the proposed method. EHR: electronic health record; IF: influence factor.



Group-Specific Comorbidity Graph Construction

This graph construction component creates distinct graphs for different demographic cohorts to capture how comorbidities co-occur and interact within each patient group. To create group-specific comorbidity graphs, patients are first stratified into different demographic cohorts, which could enhance downstream predictions because demographic similarity relates to decompensation trajectories [36,37]. For each cohort h, we created a comorbidity graph $G_h=(V_h, E_h)$, where V_h indicates nodes (comorbidities), and E_h denotes edges between 2 comorbidities that coexist in at least one patient within the cohort. For a patient group G_h, we quantified the edge weight between 2 comorbidities p and q based on a CIS, calculated as follows:

$$CIS_k(p,q,G_h) = IF(p,y_k) \times IF(q,y_k) \times \alpha_{p,q}^{(h)} \quad (1)$$

$$Z_{i,k} = \sum_{j=1}^{B} d_{i,j} IF(l_j, y_k \mid G_h(i)) \quad (2)$$

where $\alpha_{p,q}^{(h)}$ is the fraction of patients who have both comorbidities p and q in cohort h, and y_k denotes the k th target. B denotes a prespecified set of clinically prevalent and etiologically relevant comorbidities that are used to define the graph nodes for parsimony and interpretability. $\left[IF(l_j, y_k | G_{h(i)})\right]$ calculates a parametrized function, where l_j denotes the j-th comorbidity node in the select set of B comorbidities. The parameters in $\left[IF(l_j, y_k | G_{h(i)})\right]$ were learned by minimizing $\sum_i \sum_k (z_{i,k}\text{-}y_{i,k})^2$ over the training set. Once learned, a cohort-conditional matrix $\left[IF(l_j, y_k | G_h)\right] \in R^{B \times K}$ is constructed, where each entry $IF(l_j, y_k | G_h)$ represents an IF score (a scalar) for comorbidity l_j and target y_k, calculated for cohort h. This cohort-conditional matrix offers an interpretable schema of group-level comorbidity-symptom relationships; IF scores indicate subtle, clinically relevant comorbidity and reflect the complex, gradual nature of disease progression. Next, we calculated the edge weight by collapsing target-specific CIS (ascites, varices, HE, and HRS) as follows:

$$w_{p,q}^{(h)} = \sum_k CIS_k(p,q,G_h) \quad (3)$$

where each edge in G_h reflects the aggregate, multitask synergy of a specific disease pair.

Individual-Level Graph Representation Learning

With graph G_h, we applied Node2Vec [88] to represent nodes and edges in a way that preserves both local and global disease relationships. For each disease l_j in graph G_h, Node2Vec outputs an embedding vector $\phi_h(l_j) \in R^v$. To account for disease temporal dynamics at the patient level, we constructed individual-level embeddings:

$$v_{i,j} = \phi_h(l_j) \times \delta_{i,j} \times d_{i,j} \quad (4)$$

where $v_{i,j}$ is the individual-level embedding of comorbidity j for patient P_i, and $\delta_{i,j}$ denotes the time since diagnosis of the comorbidity, which serves as a proxy for disease chronicity. The individual-level embeddings were constructed using the group-level embeddings by incorporating disease duration to distinguish long-standing risk factors from acute conditions, thereby enhancing individual-level risk stratification for decompensation predictions.

As broad demographic trends, such as age-gender strata and overall disease impact, modulate decompensation risk beyond any single comorbidity, we endowed each cohort h with a compact "group fingerprint" $g_{emb}(G_h) \in R^u$, obtained

by applying principal component analysis (PCA) to the V-dimensional Node2Vec space:

$$g_{emb}(G_h)=U^{\top}\underline{\phi_h}, g_{emb}(G_h) \in R^{u} \quad (5)$$

$$\underline{\phi_h}=\frac{1}{B_h}\sum_{j=1}^{B_h}\phi_h(l_j), \underline{\phi_h} \in R^{v} \quad (6)$$

where $U \in R^{v \times u}$ is the matrix of top-u PCA eigenvectors, $\underline{\phi_h}$ is cohort h's mean node representation, and B_h is the total number of important comorbidities in cohort h. $g_{emb}(G_h)$ offers a low-rank projection that can reduce noise, curb overparameterization, and provide the network with a succinct summary of cohort-wide factors [89], which individual-level embeddings alone cannot encode (eg, shared lifestyle and care access patterns).

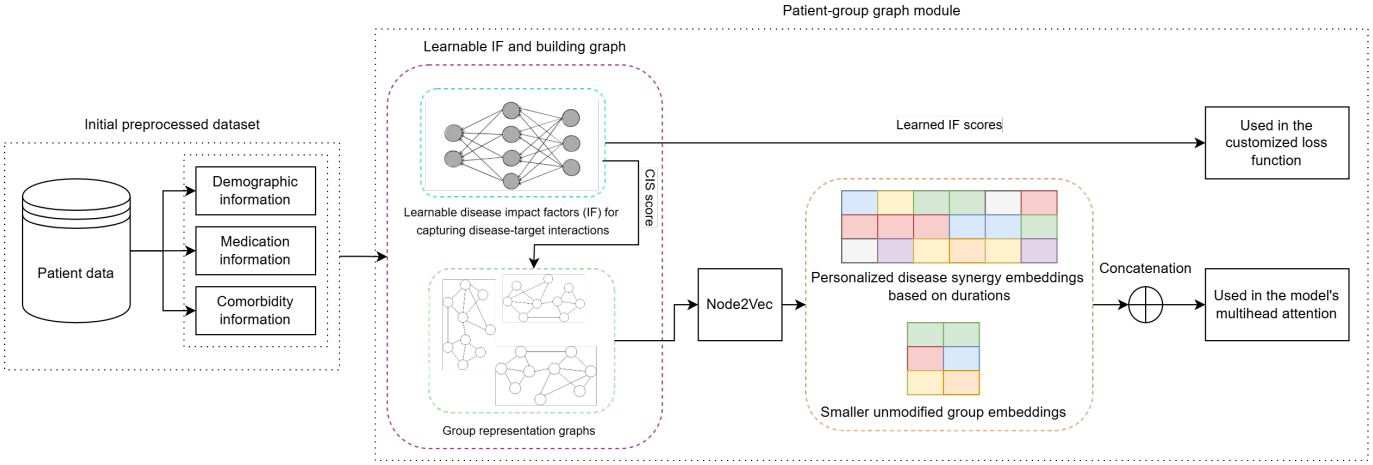
We then concatenated patients' individual-level embedding $v_{i,j}$ with a group fingerprint $g_{emb}(G_h)$ to form a combined context tensor.

$$v_{i,j}=\left[v_{i,j}; g_{\text{emb}}(G_h)\right], \tilde{v}_{i,j} \in \mathbb{R}^{v+u}, \tilde{V}_i=\{\tilde{v}_{i,j}\}_{j=1}^{B} \quad (7)$$

where $\tilde{V}_i$ represents the concatenated context tensor for patient P_i, and B denotes the total number of comorbidities in $G_{h(i)}$. This effort integrates both individual disease dynamics and group-level characteristics, thus providing a more comprehensive representation of each patient's profile.

As group parameters, PCA was exclusively fit on the training-set embeddings; testing-set embeddings were transformed using the fixed, training-fitted parameters (centering and projection) without refitting. Figure 2 presents the pipeline of patient embedding learning and use, which comprises IF score learning, subgroup graph constructions, and individual-level embedding generation, which jointly support downstream prediction tasks.

Figure 2. Infographic depiction of patient embedding learning and use module, which takes patient data as inputs and transforms them to embeddings using a cohort-level graph. CIS: condition impact score; IF: influence factor.



Multitask Learning

A binary mortality indicator was included as an auxiliary task for accurate decompensation symptom predictions because mortality is an important clinical event associated with decompensation [28]. The proposed method adopts a multitask architecture to predict 4 decompensation symptoms and mortality within a 3-year window. Mortality can inform (severe) prognosis representation learning, thereby improving decompensation predictions. Inclusion of this auxiliary task is aligned with clinical knowledge that recognizes the relationship of decompensation progression and mortality. By incorporating mortality in the modeling, we have an additional means to distinguish patient trajectories that vary in diseases [90,91]. We construct an input $X_i=\left[AGE_i, SEX_i, \{d_{i,j}, \delta_{i,j}\}_{j=1}^{D}, \{med_{i,m}, dosage_{i,m}, dur_{i,m}, S_{i,m}\}_{m=1}^{M}, \hat{v}_i\right]$ for the multitask prediction component, which contains patient P_i's demographics, comorbidities, medications, and the learned patient context vector. X_i is passed to a fully connected layer and converted to a latent vector E:

$$E_i = ReLU(WX_i+b) \quad (8)$$

The multitask prediction component fuses the tabular latent vector with the personalized graph representation into a single context vector c_i and then applies multihead self-attention to obtain an attended representation for prediction.

$$c_i = \left[E_i; f_{\text{graph}}\left(\hat{V}_i\right)\right],$$
$$E_i' = \text{MultiHeadAttention}\left(W^{Q}c_i, W^{K}c_i, W^{V}c_i\right) \quad (9)$$

where $f_{\text{graph}}(\cdot)$ is the graph representation module mapping the per-disease context $\tilde{V}_i$ to a fixed-length vector, and E_i' is input to a fully connected layer to produce a shared representation (task-specific heads):

$$E_{i,refined} = ReLU(Linear(E_i')) \quad (10)$$

A linear head predicts mortality as an auxiliary task. A shared multilayer perceptron head takes the concatenated vector of $E_{i,refined}$ and predicted death probability $\hat{y}_{i,\text{DEADT}}$ as input and projects it to logits for ascites, variceal bleeding, HE, and HRS. For patient i and task k, we computed the following:

$$\begin{gathered}\hat{y}_{i,\text{DEADT}} = w_{\text{death}}^{\top} E_{i,\text{refined}} + b_{\text{death}},\\ P_{i,\text{DEADT}} = \sigma(\hat{y}_{i,\text{DEADT}}),\\ (\hat{y}_{i,\text{ASCT}}, \hat{y}_{i,\text{VART}}, \hat{y}_{i,\text{ENCET}}, \hat{y}_{i,\text{HEPT}}) = f_{\text{main}}([E_{i,\text{refined}}; p_{i,\text{DEADT}}])\end{gathered} \tag{11}$$

ASCT, VART, ENCET, and HEPT represent the 4 decompensation symptoms, ascites, varices, HE, and HRS, respectively, while DEADT represents death. Next, we applied a sigmoid activation on $\hat{y}_{i,k}$ to produce an estimated probability for each symptom. The auxiliary task regularizes model learning by steering the network toward features and trajectories associated with severe clinical outcomes. The auxiliary task helps the method learn patterns indicative of disease progression and thereby differentiates high- and low-risk trajectories more effectively for the main (decompensation) prediction task.

The proposed method was trained with binary cross-entropy (BCE) on each target y_k:

$$\begin{gathered}l_{i,\text{total}} = \sum_{k \in \{\text{ASCT,VART,ENCET,HEPT}\}} l_{i,k} + l_{i,\text{DEADT}},\\ l_{i,k} = l_{\text{BCE},\omega_k}(\hat{y}_{i,k}, y_{i,k})\left[1 + \lambda \sum_{j=1}^{B} d_{i,j} IF(l_j, y_k \mid G_{h(i)})\right],\\ l_{i,\text{DEADT}} = l_{\text{BCE}}(\hat{y}_{i,\text{DEADT}}, y_{i,\text{DEADT}})\end{gathered} \tag{12}$$

where $l_{BCE}(\hat{y}_{i,k}, y_{i,k})$ is a BCE loss, $IF(l_j, y_k)$ is the IF score measuring the influence of patient P_i's comorbidity to target y_k, and $d_{i,j}$ is a binary indicator of each comorbidity. $\lambda \sum_{j=1}^{B}(d_{i,j} \cdot IF(.,.))$ serves as a scaling factor, adjusted by the influence of comorbidity, with λ controlling the strength of the scaling. Among the targets, the mortality head uses the standard BCE without IF modulation because its risk is already modeled through the symptom heads implicitly. In addition, we applied standard positive-class weights [92] to address data imbalances (ie, assigning a higher weight to the infrequent HRS label) and improve the model performance among less frequent classes. We present key hyperparameters in Multimedia Appendix 1.

Data and Evaluation Design

From a major health care organization in Taiwan, we obtained demographic and clinical data for 29,915 patients diagnosed with cirrhosis between 2007 and 2013. Each patient was described by variables pertaining to demographics, comorbidity, and medication (Multimedia Appendix 2). We considered 4 outcome classes (ascites, varices, HE, and HRS) and an auxiliary task (mortality). To ensure comparability across patients, we applied z-score standardization to all disease durations [93]. Table 2 presents patient demographics and descriptive statistics, as well as the outcome class distribution. As shown, patients with HRS were slightly older (mean 57.62, SD 13.017 years) than those with other decompensation symptoms, with the varices group being the youngest (mean 55.88, SD 12.58 years). In all 4 outcome classes, approximately two-thirds of patients were men, suggesting a relative gender imbalance in the sample. Among the symptoms, patients with HRS had the most lingering conditions in their deterioration path to decompensation (239/273, 87.56%) and patients with varices the least (4010/6339, 63.26%). In the sample, 17,464 (58.38%) patients were deceased, and 12,451 (41.62%) were alive.

Table 2. Descriptive statistics of the sample (N=29,915) by outcome class, with symptom rows (ascites, varices, HE, and HRS) indicating patients with distinct, not mutually exclusive symptoms.

Outcome class	Age (y), mean (SD)	Sex, n (%)		Prior liver condition, n (%)	
		Male	Female	Yes	No
Ascites	56.72 (12.72)	3398 (66.65)	1700 (33.35)	3864 (75.79)	1234 (24.21)
Varices	55.88 (12.58)	4382 (69.13)	1957 (30.87)	4010 (63.26)	2329 (36.74)
HE[a]	56.84 (12.95)	2916 (67.69)	1392 (32.31)	3391 (78.71)	917 (21.29)
HRS[b]	57.62 (13.02)	183 (67.03)	90 (32.97)	239 (87.56)	34 (12.44)
No complication[c]	53.12 (14.43)	14,207 (68.91)	6409 (31.09)	9559 (46.37)	11,057 (53.63)

[a]HE: hepatic encephalopathy.
[b]HRS: hepatorenal syndrome.
[c]Includes patients without any decompensation symptoms.

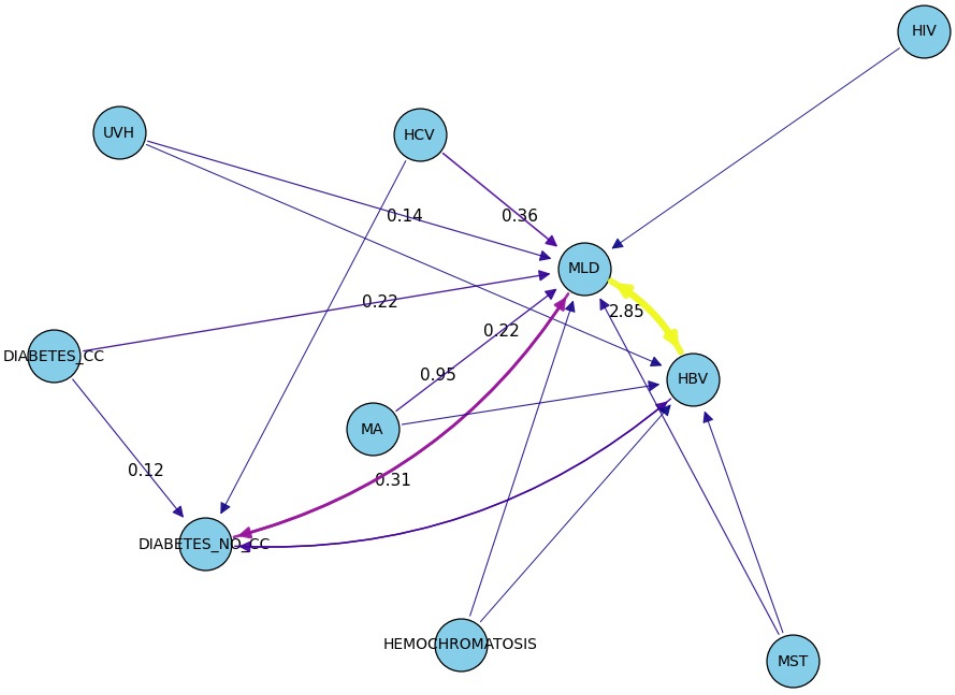
For each patient, variable values on or before the cirrhosis diagnosis date were used for model training, fine-tuning, and testing. Each outcome class was coded as a binary variable, with 1 indicating its occurrence within the 3-year window and 0 otherwise. To evaluate the proposed method, we split the sample into training (80%) and held-out testing (20%), with stratified datasets preserving class proportions across different outcome classes [94]. We repeated this procedure 10 times, using a different random seed each time. We report each investigated method's performance averaged across the 10 trials. We also conducted 2-tailed t tests to examine whether the observed performance differentials were statistically significant. The evaluation relies on recall, precision, F-measure, and area under the curve (AUC). We report macro-averaged values because they reflect overall performance across different outcome classes, with the consideration of the imbalanced distributions in the sample. By and large, F-measure and AUC are more indicative of a method's predictive utilities than recall and precision [95]. For probability-to-label conversion, we adopted appropriate

symptom-specific decision thresholds. Because HRS is highly imbalanced, we increased the weight of the positive class 10-fold for the proposed method and all benchmarks. For this outcome class, the decision threshold was selected separately for each method that used the 5-fold cross-validation on the training data. We evaluated thresholds from 0.05 to 0.95 in increments of 0.05 and selected the threshold that produces the highest F_1-score. This procedure suggests a threshold of 0.85 for the proposed method. For ascites, varices, and HE, we applied a threshold of 0.50. The held-out test data were not used for threshold selection or class calibration.

We considered 12 prevalent methods for benchmarks. Some methods can be applied for both single-task and multitask predictions, so we used "multimethod" to refer to the multitask setting. These benchmarks include 4 conventional machine learning methods (XGB, ADAB, RF, and linear regression), 2 deep learning–based methods (single-task Single-fully connected neural network [FCN] and multitask Multi-FCN), 2 CNNs (single-task Single-CNN and multitask Multi-CNN), 2 TCNs (single-task Single-TCN and multitask Multi-TCN), and 2 LSTM networks (single-task Single-LSTM and multitask Multi-LSTM). These methods represent different analytics approaches and exemplify prevalent methods. All benchmarks used the same input variables and adopted the identical training (80%) and testing (20%) split to ensure consistency in the comparison. For each method, 5-fold cross-validation was applied on training data for parameter tuning. We describe model configurations, key hyperparameter settings, tuning ranges, and procedure in Multimedia Appendix 3.

For the proposed method, we created different demographic patient cohorts. In line with Liu et al [37], we stratified patients by gender and age (≤40, >40 and ≤60, >60 years) to reveal age distributions in the sample. We also designed a binary liver condition indicator that incorporates clinical and lifestyle factors (eg, biliary disease, Wilson disease, lack of medication use, overweight, smoking, and prior mild or severe liver disease or transplantation). This binary indicator consolidates important risk factors into a single feature to reduce sparsity and improve statistical power [96]. Figure 3 presents an exemplary graph to illustrate different diseases and their co-occurrence for a specific patient cohort (group). As shown, the edge between HBV and mild liver disease (MLD; liver damage and coma from hepatitis and unspecified viral hepatitis) has a weight of 2.85, reflecting a strong relationship between these comorbidities for this patient group, which influences decompensation risk. Parameters of the cohort-specific graph were learned from training data and were fixed at testing time. We summarize the different cohorts together with their size in Multimedia Appendix 2. To reduce the effects of varying cohort sizes, the cohort-level embeddings were combined with shared multitask representations and individual-level features to enhance prediction.

Figure 3. A partial, group-specific comorbidity graph for a patient subgroup, with graphs generated from training data and displayed links with weights >0.05 for readability. CC: chronic complications; HBV: hepatitis B virus; HCV: hepatitis C virus; MA: any malignancy, excluding skin cancer; MLD: mild liver disease; MST: metastatic solid tumor; UVH: unidentified viral hepatitis.



Ethical Considerations

This study was approved by the Chang Gung Medical Foundation Institutional Review Board (201800270B0). Data were deidentified prior to any analyses, in compliance with established ethical standards for health care research. All research procedures were performed in accordance with relevant guidelines and regulations. No consent was required from patients in the sample. Physician participation in the interview was voluntary; each physician was explicitly

informed of the voluntary nature, with his or her right to withdraw at any time preserved. Physicians received no financial compensation for their participation in the interview.

Results

Predictive Performance of Each Investigated Method

Table 3 presents the average performance of each method across 10 trials, together with the 95% CIs. As shown, the proposed method outperformed all benchmarks in most performance measures, exhibiting superior capability to estimate patients' risk of decompensation symptoms. It attained an AUC value of 0.871 (SD 0.004), showing a 6.09% improvement over the best-performing benchmark (XGB). The proposed method also achieved an *F*-measure value of 0.678 (SD 0.010), which exhibits a 13.95% improvement over the best-performing benchmark (Single-FCN). It had a recall value significantly higher than that of the best-performing benchmark, a 15.42% improvement over Single-FCN, although its precision was marginally lower than that of the best-performing benchmark (0.663, SD 0.005 vs 0.675, SD 0.131 by ADAB). These comparative results affirm the utilities of this graph-based, multitask, deep learning method for predicting patients' deterioration from the compensated to the decompensated stage.

Table 3. Each investigated method's aggregate performance across main prediction tasks (varices, ascites, hepatic encephalopathy, and hepatorenal syndrome predictions) over 10 trials using random seeds.

Method	Precision, mean (95% CI)	Recall, mean (95% CI)	*F*-measure, mean (95% CI)	AUC[a], mean (95% CI)
XGB[b]	0.668 (0.640-0.696)	0.548 (0.540-0.556)	0.546 (0.535-0.556)	0.821 (0.815-0.827)
ADAB[c]	0.675 (0.644-0.706)	0.553 (0.535-0.570)	0.508 (0.502-0.513)	0.813 (0.809-0.818)
RF[d]	0.671 (0.641-0.701)	0.569 (0.546-0.592)	0.510 (0.504-0.516)	0.820 (0.815-0.824)
LR[e]	0.638 (0.614-0.662)	0.529 (0.524-0.534)	0.519 (0.513-0.525)	0.798 (0.791-0.806)
Single-FCN[f]	0.626 (0.601-0.650)	0.642 (0.624-0.660)	0.595 (0.573-0.618)	0.807 (0.800-0.814)
Single-CNN[g]	0.642 (0.617-0.666)	0.567 (0.546-0.587)	0.525 (0.516-0.534)	0.802 (0.795-0.809)
Single-TCN[h]	0.646 (0.618-0.674)	0.551 (0.538-0.563)	0.554 (0.539-0.568)	0.809 (0.802-0.816)
Single-LSTM[i]	0.633 (0.608-0.658)	0.540 (0.530-0.551)	0.514 (0.505-0.523)	0.781 (0.771-0.792)
Multi-FCN	0.538 (0.504-0.573)	0.510 (0.505-0.514)	0.475 (0.466-0.483)	0.777 (0.766-0.788)
Multi-CNN	0.617 (0.591-0.642)	0.522 (0.516-0.527)	0.504 (0.494-0.513)	0.798 (0.790-0.806)
Multi-TCN	0.588 (0.572-0.603)	0.588 (0.569-0.607)	0.576 (0.560-0.591)	0.738 (0.732-0.744)
Multi-LSTM	0.576 (0.544-0.608)	0.515 (0.509-0.520)	0.488 (0.479-0.497)	0.739 (0.729-0.750)
Proposed method	0.663 (0.659-0.665)	0.741 (0.732-0.750)	0.678 (0.676-0.682)	0.871 (0.869-0.874)

[a]AUC: area under the curve.
[b]XGB: extreme gradient boosting.
[c]ADAB: adaptive boosting.
[d]RF: random forest.
[e]LR: linear regression.
[f]FCN: fully connected network.
[g]CNN: convolutional neural network.
[h]TCN: temporal convolutional network.
[i]LSTM: long short-term memory.

Prediction Performance for Each Decompensation Symptom

We present the respective methods' predictive performance for each decompensation symptom in Multimedia Appendix 4. As shown, the proposed method achieved the highest AUC values across all symptoms and had the highest recall and F_1-scores for all symptoms except HRS, manifesting its superior predictive performance. For HRS, it had an AUC value of 0.871 (SD 0.012), which was 3.81% higher than that of the best-performing benchmark (0.839, SD 0.013, Single-TCN). Its *F*-measure value was 0.518 (SD 0.018), comparable to that of the best-performing benchmark (0.533, SD 0.011, Single-LSTM). For varices, the proposed method had an *F*-measure value of 0.744 (SD 0.006), which was 14.29% higher than that of the best-performing benchmark (0.651, SD 0.017, Single-FCN). It had an AUC value of 0.877 (SD 0.003), exhibiting a 9.90% improvement over the best-performing benchmark (0.798, SD 0.006, RF). Its recall was 22.87% higher than the best-performing benchmark (0.656, SD 0.006, Multi-TCN), and its precision was 2.67% higher than the best-performing benchmark (0.712, SD 0.019, RF). For ascites, the proposed method attained an *F*-measure of 0.704 (SD 0.006), showing a 7.98% improvement over the best-performing benchmark (0.652, SD 0.019, Single-FCN). It had an AUC value of 0.843 (SD 0.005), which was 2.55% higher than that of the best-performing benchmark (0.822, SD 0.004, XGB). The proposed method had a recall value of 0.762 (SD 0.006), which was 18.14% higher than that of the best-performing benchmark (0.645, SD 0.030, Single-FCN), although its precision was lower than that of the best-performing benchmark (ADAB). For HE, the proposed method attained an *F*-measure value of 0.744 (SD 0.006), which was 29.62% higher than the best-performing

benchmark (0.574, SD 0.070, Single-FCN). Its AUC value was 0.894 (SD 0.004), showing a 7.32% improvement over the best-performing benchmark (0.833, SD 0.003, XGB). The proposed method's recall (0.818, SD 0.011) was notably 42.51% higher than that of the best-performing benchmark (0.574, SD 0.060, Single-FCN), although its precision was lower than that of the best-performing benchmark (ADAB).

These results further indicated that the proposed method can estimate decompensation symptoms among patients with cirrhosis more effectively than many prevalent methods. It attained the highest *F*-measure values for ascites (0.704, SD 0.006), varices (0.744, SD 0.006), and HE (0.744, SD 0.006). For HRS, which is particularly difficult to predict, the proposed method achieved the best AUC value of 0.871 (SD 0.012) while maintaining an *F*-measure value comparable to the best-performing benchmark. The prediction utilities, overall or for each decompensation symptom, affirm the proposed method's greater efficacy for identifying patients at higher risk of decompensated cirrhosis than existing methods.

Ablation Analysis

We conducted an ablation analysis to examine each important component's contribution to the proposed method's predictive utilities. The ablation considered 5 models: M4T, M4*t*+AUX, M4*t*+AUX+Graph, M4T+AUX+Graph+Personal, and M4T+AUX+Graph+Personal+Loss (the complete proposed method). All models followed the same procedure as the main evaluation for training and evaluation (ie, identical patient cohort, input variables, data preprocessing, training-testing split, and performance measures). M4T served as a baseline, with a multitask deep learning framework for all 4 decompensation symptoms (ascites, varices, HE, and HRS). M4T+AUX extends M4T by including an auxiliary mortality task (DEADT) to provide shared severity supervision. M4T+AUX+Graph further incorporates group-specific comorbidity graph representation (ie, Node2Vec-derived disease embeddings from different demographic patient groups). M4T+AUX+Graph+Personal integrates individual-level comorbidity temporal patterns into the graph embeddings. M4T+AUX+Graph+Personal+Loss denotes the proposed method in its entirety, containing all the key components.

Table 4 presents the respective models' average AUC values across the 4 decompensation symptoms. As shown, the multitask baseline (M4T) had an AUC value of 0.803. With the inclusion of the auxiliary mortality task, M4T+AUX attained 0.811, showing a 1.00% improvement over the baseline model. Adding group-level graph representations, M4T+AUX+Graph increased the AUC value to 0.839, exhibiting a 3.45% improvement over M4T+AUX. Adding patient-level temporal personalization, M4T+AUX+Graph+Personal further increased the AUC value to 0.861, exhibiting a marginal increment of 2.62%. Finally, the complete proposed methods attained an AUC value of 0.871 by incorporating IF-based loss reweighting, showing a marginal increment of 1.16%, relative to M4T+AUX+Graph+Personal. The results indicate that each component contributed to the proposed method's performance in a comparable way, with graph representation and temporal personalization offering greater incremental performance gains, while the auxiliary prediction task and IF-weighted optimization yielded additional improvements.

Table 4. Ablation analysis of different models' AUC values averaged across different main prediction tasks (varices, ascites, hepatic encephalopathy, and hepatorenal syndromes).

Model	AUC[a] value across symptoms
M4T	0.803
M4*t*+AUX	0.811
M4*t*+AUX+Graph	0.839
M4*t*+AUX+Graph+Personal	0.861
M4*t*+AUX+Graph+Personal+Loss (complete method)	0.871

[a]AUC: area under the curve.

Interpretability

Beyond prediction performance, the proposed method's estimates provide post hoc interpretability, unfolding essential patient-, disease-, and task-level factors that explain its estimates. We present 4 post hoc analyses: a contrastive comorbidity graph analysis, a Shapley Additive Explanations (SHAP) analysis [97], a counterfactual disease risk analysis, and a decompensation symptoms interdependence analysis. To enhance probability interpretability, we applied post hoc probability calibration per outcome using Platt scaling (logistic calibration). For each target variable, a calibration model was adjusted to fit a held-out (20%) training subset, thereby mapping raw logits to calibrated probabilities. The calibrated probabilities were then used in the subsequent counterfactual disease risk analysis as well as the interdependence analysis of decompensation symptoms.

Figure 4 presents exemplary visualizations of a patient's comorbidity graph, together with her group-specific comorbidity graph as a reference. As shown, the left panel displays the group-specific comorbidity graph for the reference patient cohort (female, age >40 and ≤60 years, and having prior liver issues), where edges are weighted by group-level disease co-occurrence (IF×group multiplier). The right panel depicts an individual patient's network adjusted for chronicity, according to the temporal order and persistence of each condition. The edge weights reveal how different clinical conditions (hepatitis B, MLD, and nonchronic diabetes) of the focal patient have interacted over time and thereby unfold essential comorbidities. The

focal patient's network has fewer nodes and edges, reflecting that this patient has fewer comorbidities than the cohort. These illustrated networks allow physicians to compare a patient with a reference cohort and assess the chronicity and relationship of comorbidities to interpret the method's estimates for clinical decision-making and patient management.

We also performed a SHAP analysis to scrutinize feature importance for each outcome class (ascites, varices, HE, and HRS). SHAP values were calculated for individual test patients using a training-derived background sample, and features were then ranked by the absolute magnitude of SHAP to identify the most influential contributors for each outcome. That is, SHAP was applied for explanation only instead of modifying model training or performance estimates. Figure 5 illustrates the proposed method's predictions via SHAP-enabled visualizations. For consistency of interpretation, we used the same illustrative patient in Figures 4-7, who was randomly selected from testing data and had at least one positive decompensation symptom. As shown, top predictors pertain to the patient's disease network, which reinforces the consistency between graph-derived variables and clinically important variables. In this example, ascites, varices, and encephalopathy predictions were mainly influenced by chronicity of hepatic deterioration and viral hepatitis, as reflected by MLD, HBV, HCV, diabetes, and hypertension features in Figure 5. The visualization of feature importance reveals influential factors and helps physicians better interpret the proposed method's estimates at the patient level.

Figure 4. An exemplary patient reference (ref) cohort graph of comorbidities with condition impact scores (CIS; left) and an individual patient's graph adjusted by chronicity (chron; right). CC: chronic complications; HBV: hepatitis B virus; HCV: hepatitis C virus; IF: influence factor; MA: any malignancy, excluding skin cancer; MLD: mild liver disease; MST: metastatic solid tumor; ref: reference; UVH: unidentified viral hepatitis.

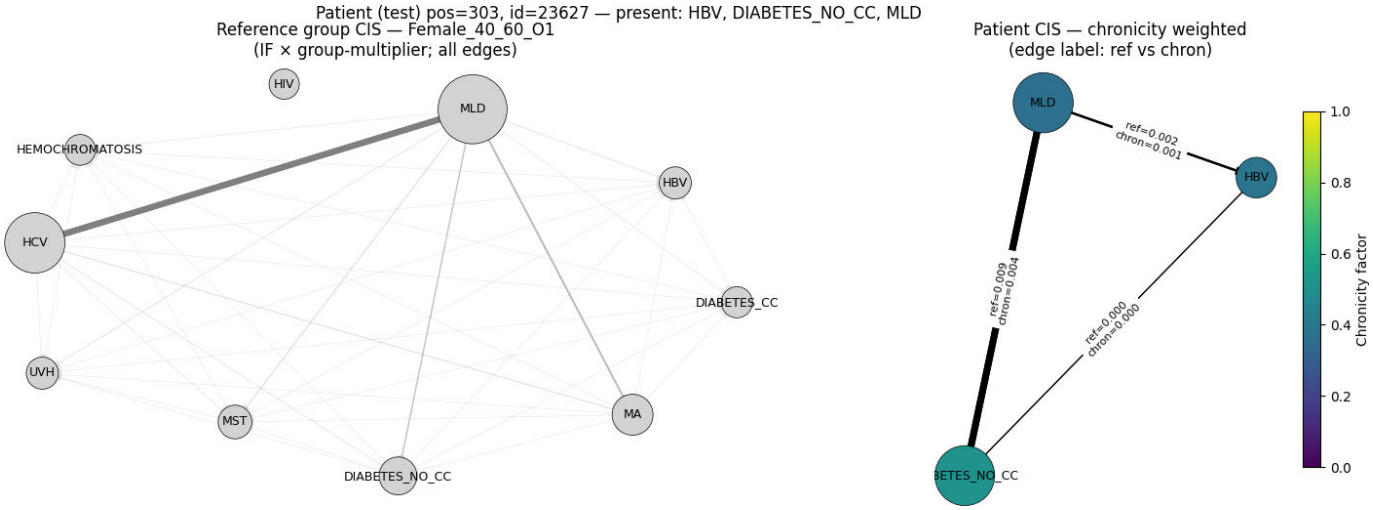


Figure 5. Shapley Additive Explanations (SHAP)–based feature attribution for a patient for each decompensation symptom, which conveys each comorbidity's importance for different decompensation symptom estimates. CC: chronic complications; CHF: congestive heart failure; HBV: hepatitis B virus; HCV: hepatitis C virus; HE: hepatic encephalopathy; HRS: hepatorenal syndrome; MA: any malignancy, excluding skin cancer; MI: myocardial infarction; MLD: mild liver disease; MST: metastatic solid tumor; PVD: peripheral vascular disease.

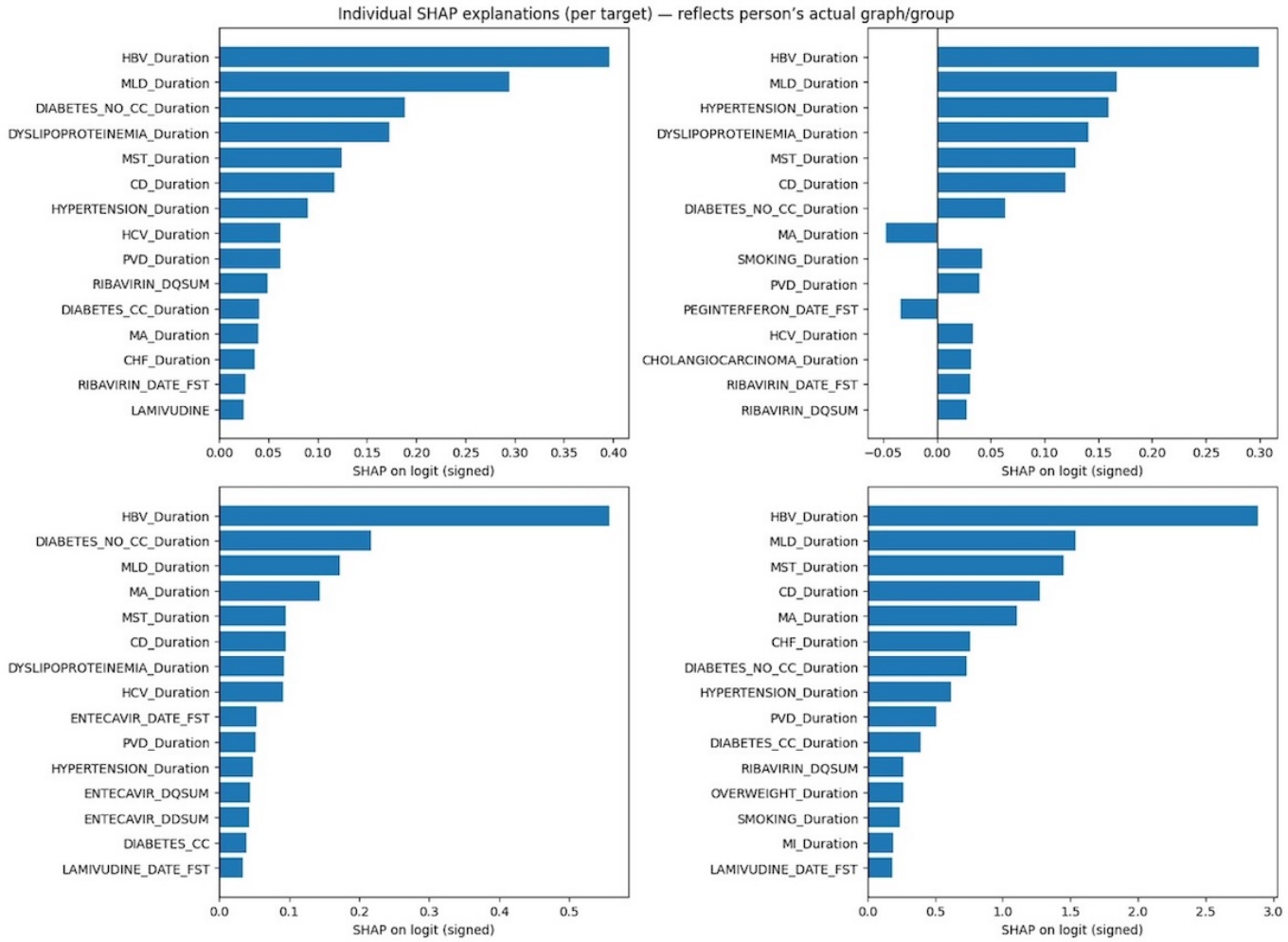


Figure 6. Counterfactual (cf) risk attribution for patient disease trajectories. CC: chronic complications; HBV: hepatitis B virus; HE: hepatic encephalopathy; HRS: hepatorenal syndrome; MLD: mild liver disease.

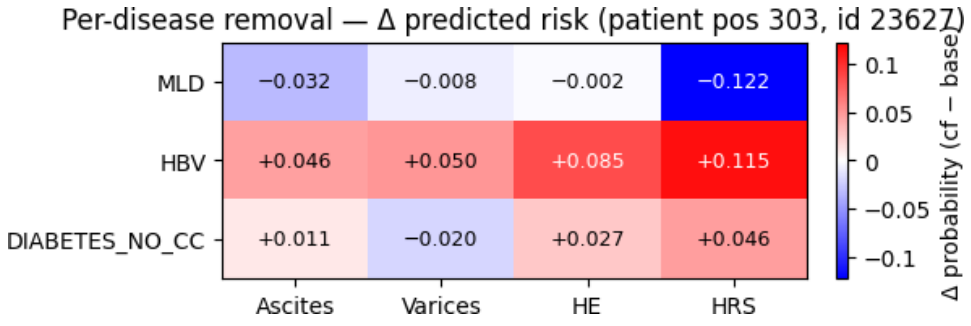
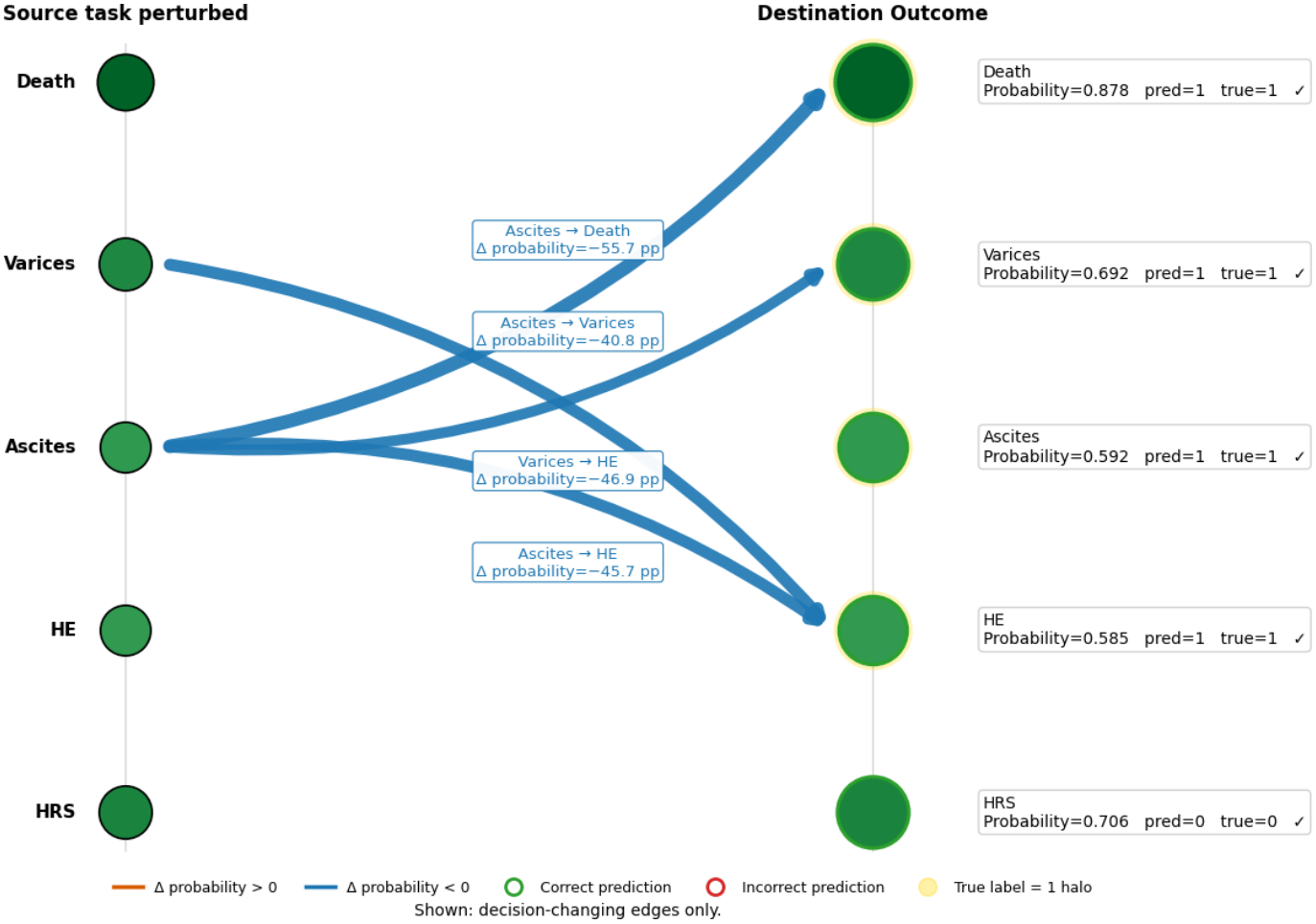


Figure 7. Patient-level intertask association map where nodes denote outcomes (ascites, varices, HE, HRS, and death). A directed edge $A \rightarrow B$ reflects the change $\Delta p_{A \rightarrow B}$ in destination probability when source probability $p(A)$ is replaced by its training-cohort median, with all other outcome probabilities fixed. Edge labels $\Delta p_{A \rightarrow B}$ are shown in percentage points, and edge width is proportional to $|\Delta p_{A \rightarrow B}|$. Blue edges indicate $\Delta p_{A \rightarrow B}<0$, and orange edges signify $\Delta p_{A \rightarrow B}>0$. Only edges with significant (meaningful) impact are displayed. Node outline color indicates prediction correctness (green, correct; red, incorrect), and a yellow halo indicates true label=1. HE: hepatic encephalopathy; HRS: hepatorenal syndrome; pred: predicted label.



To understand the association between comorbidities and decompensation symptom risks, we performed a probabilistic counterfactual risk analysis using patient comorbidity profiles. We first computed baseline predicted probabilities for each decompensation symptom using each patient's observed comorbidities and comorbidity durations. We then generated counterfactual inputs by removing one comorbidity at a time while keeping all other covariates fixed. Removal was proceeded by setting the focal comorbidity's binary indicator, duration feature, and disease-specific node embedding to 0, with all remaining features and embeddings intact. For each removed comorbidity d and outcome y_k, we recomputed risks and calculated the change of risk $\Delta_{d,k}$ as follows.

$$\Delta_{d,k}=p(y_k=1\,|\,\text{comorbidity}d\text{removed})-p(y_k=1\,|\,\text{comorbidity}d\text{present}). \quad \textbf{(13)}$$

A $\Delta_{d,k} > 0$ means that removal of d increases the estimated probability of decompensation symptom k; $\Delta_{d,k} < 0$ indicates a decrease in risk. Figure 6 presents patient-level $\Delta_{d,k}$ values in a comorbidity-by-outcome matrix in which each row refers to a removed comorbidity and each column denotes a decompensation symptom. For the illustrative patient in Figure 6, removing MLD yields negative shifts across all outcomes (ascites: –0.036; varices: –0.008; HE: –0.002; and HRS: –0.073), with the largest magnitude for HRS, implying that, for this patient, the proposed method predicts positive associations of MLD with the estimated risks for different symptoms, particularly HRS. This outcome-specific decomposition is useful and can help physicians understand the predictions for each patient, although it is an associative interpretation of the learned predictive function and should not be interpreted as direct physiological causality or treatment effect. All percentage-point shifts (Δ) in Figure 6 were computed from calibrated probabilities.

We also examined the relationships among patient outcomes, mortality, and decompensation symptoms. Figure 7 presents a patient-level intertask association map that depicts how model-estimated risk for one outcome is linked to the estimated risks of others. These relationships can help clinicians identify outcomes strongly associated with others at the patient level. This analysis perturbs the probability of one outcome and quantifies its directional effect on other outcomes. For each outcome $B \in \{ASCT, VART, ENCET, HEPT, DEADT\}$,

we fit a destination-specific logistic regression stacker with estimated probabilities of all other outcomes, $X_B = \{\hat{p}_A : A \in \{ASCT, VART, ENCET, HEPT, DEADT\} \setminus \{B\}\}$, and response y_B. Each stacker estimates $P(y_B=1 \mid X_B)$ as a calibration and association layer over the proposed method's outputs. For each patient, we used X_B to compute baseline stacked probabilities $\hat{p}_B^{base}$, then replaced an $\hat{p}_A$ with its training-cohort median $\tilde{p}_A$ and kept all remaining predictors fixed to recompute $\hat{p}_B^{cf}$. For each directed perturbation $A \to B$, we defined the edge effect as $\Delta p_{A \to B} = \hat{p}_B^{cf} - \hat{p}_B^{base}$ (in percentage points [pp]). A positive $\Delta p_{A \to B}$ reveals that neutralizing source signal A increases destination risk B, whereas a negative $\Delta p_{A \to B}$ indicates a decrease. In Figure 7, the logistic stackers were fitted on the training data and then applied to the illustrated case drawn from the test data to demonstrate their capabilities under deployment conditions.

Figure 7 illustrates how changes in the prediction probability of source outcome can lead to a reversal of the estimated destination outcome. For this patient, all 5 displayed effects are negative, indicating that replacing the source-outcome probability with its training cohort median decreases the estimated probability of the corresponding destination outcome: $\Delta p_{ASCT \to DEADT} = -50.12$ pp, $\Delta p_{ASCT \to VART} = -42.60$ pp, $\Delta p_{ASCT \to ENCET} = -45.59$ pp, $\Delta p_{VART \to ASCT} = -32.50$ pp, and $\Delta p_{VART \to ENCET} = -45.91$ pp. For example, $\Delta p_{VART \to ENCET} = -45.91$ pp means that replacing the variceal-bleeding probability with its cohort median lowers the stacked encephalopathy probability by 0.459 for this patient. The illustrative patient-level dependency map can support triage and highlight destination risks most sensitive to other outcome signals, with an emphasis on the associations among patient outcomes. All reported percentage-point shifts (Δ) in Figure 7 were computed from calibrated probabilities.

Discussion

Principal Findings

Compared with many prevalent methods, the proposed method exhibited significantly better prediction efficacy, reflected by an average AUC value of 0.871 (SD 0.004) and F-measure of 0.678 (SD 0.010). This method is effective for symptom-level decompensation risk stratification beyond a dichotomous compensation versus decompensation view. An ablation analysis revealed the importance of each key component of the proposed method, perhaps with greater performance gains by graph-based representation and patient-level temporal patterns, which underscores the criticality of capturing comorbidity structure and disease chronicity for downstream predictions. Moreover, the proposed method can provide interpretable analyses and visualizations (such as patient-level comorbidity graphs, SHAP attributions, comorbidity-removal counterfactual maps, and intertask assist maps), which should increase user confidence, trust, and adoption.

Comparison to Prior Work

The proposed method addresses several important limitations of prior work on liver-related risk predictions. For example, most existing methods target a single outcome, such as compensation versus decompensation, without symptom-level consideration. In addition, patient subgroup structure and comorbidity interplays are largely overlooked. The proposed method adds to extant literature by considering distinct decompensation symptoms and their interplays, together with an auxiliary task. It incorporates demographic cohort structure using group-specific comorbidity graphs, as well as patient-level temporal disease information through personalized embeddings. Relative to existing regression-based, tree-based, and deep learning–based methods, our method makes methodological contributions by integrating clinically guided structure into learning via IF- and CIS-informed graph construction and IF-weighted multitask loss. Furthermore, this study provides multiple post hoc analyses for interpretability and visualizations of symptom-specific associative patterns, so the proposed method can serve as a structured and clinically interpretable means to complement score-based methods for assessing decompensation risks among patients with cirrhosis.

Limitations

This study has several limitations that should be considered to interpret the results with caution. First, we used patients from a health care system, who were diagnosed with cirrhosis between 2007 and 2013. This cohort reflects the pre–direct-acting antiviral for HCV, before a major practice change around 2014. Clinical care of patients with chronic liver diseases has evolved over time, such as antiviral access, vaccination coverage, and preventive measures that might affect patients' decompensation risks, disease trajectories, and outcomes. The illustrative post hoc interpretability analyses may offer a partial depiction regarding how etiology shifts may affect predictions. Physicians can perform counterfactual risk analyses to assess how removal or attenuation of specific etiologic factors, such as viral hepatitis–related condition reduction through vaccination and modern antiviral treatment, might alter estimated probabilities for ascites, varices, HE, and HRS. Although such analyses are associative rather than causal, they still can help assess model sensitivity under plausible epidemiologic changes. Future work should enhance robustness and external validation with recalibration using additional patient data, desirable cohorts after 2020. Second, the proposed method formulates outcomes as fixed-horizon binary labels and optimizing them with task-wise binary cross-entropy, rather than using an explicit competing-risks or survival framework. Its inclusion of mortality prediction as an auxiliary task helps the main prediction task but without considering the temporal order of decompensation and mortality. Therefore, estimates should be interpreted as risk within the 3-year window, instead of cause-specific hazard or cumulative-incidence estimates. Future work can extend this method with time-to-event and competing-risks modeling, such as using a cause-specific or subdistribution hazard approach or a multistate survival model. Third,

the proposed method was developed and evaluated using available etiological data in EHRs; it lacks laboratory and blood test data. Future research could extend the method by incorporating additional data for further improved, finer-grained predictions, such as modeling the temporal order of clinical events and adopting a stricter temporal view to formulate the prediction problem.

Future Directions

These limitations in turn suggest promising research directions. We used a particular patient cohort to develop and evaluate the proposed method, and continued research could consider multiple patient cohorts that differ in demographics, ethnic backgrounds, and clinical systems to increase both method generalizability and result robustness. Expanding the scope of inputs by including additional cirrhosis-related conditions such as jaundice [38] or laboratory biomarkers and multimodal data (eg, medical imaging and clinical notes) could further enhance clinical values [98]. Estimating the first-appearing symptom and its timing also is important for clinical decision support and patient management. Moreover, future research can extend the proposed method's interpretability by incorporating intrinsically interpretable methods for increased reasoning and explanations, which can enhance transparency and physicians' trust and clinical adoption [99].

Conclusions

Accurately predicting decompensation among patients with cirrhosis is clinically important and methodologically challenging. We developed and evaluated a graph-based multitask deep learning framework to estimate 4 decompensation symptoms (ascites, varices, HE, and HRS) within a 3-year prediction horizon after the cirrhosis diagnosis. This method incorporates demographic subgroup structure, comorbidity interplays, patient-level temporal information, and IF-weighted multitask optimization for downstream predictions. The proposed method provides more effective predictions than many existing methods, across different performance measures, thereby enabling early at-risk patient identification and risk stratification. The illustrated post hoc interpretability offers patient-level, associative explanations of its estimates for user comprehension and confidence.

Generative AI Use

The authors attest that there was no use of generative artificial intelligence (AI) technology in the ideation, generation of text, figures, or other informational content of this manuscript. The authors declare the use of generative AI technology in coding and code debugging under full human supervision.

Funding

This work has not received any funding or financial support.

Data Availability

The data used in this study cannot be made publicly accessible due to privacy issues. Nevertheless, the authors can arrange data access upon request. The implementation code of the proposed method and used interpretability tools are available in the GitHub repository [100].

Authors' Contributions

MA contributed to writing the original draft and reviewing and editing the manuscript, as well as validation, software, methodology, investigation, formal analysis, and conceptualization. DX contributed to writing the original draft and reviewing and editing the manuscript, as well as validation, methodology, investigation, formal analysis, and conceptualization. PJHH contributed to writing the original draft and reviewing and editing the manuscript, as well as validation, supervision, project administration, methodology, investigation, and conceptualization. TSH contributed to reviewing and editing the manuscript, as well as resources, project administration, investigation, and data curation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Proposed method's hyperparameters and implementation.
[DOCX File (Microsoft Word File), 22 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Variables, descriptive statistics, and coding.
[DOCX File (Microsoft Word File), 33 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Hyperparameter tuning and implementation details.
[DOCX File (Microsoft Word File), 26 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Investigated methods' performance for distinct symptoms.

[DOCX File (Microsoft Word File), 31 KB-Multimedia Appendix 4]

References

1. Schuppan D, Afdhal NH. Liver cirrhosis. The Lancet. Mar 2008;371(9615):838-851. [doi: 10.1016/S0140-6736(08)60383-9]
2. Zipprich A, Garcia-Tsao G, Rogowski S, Fleig WE, Seufferlein T, Dollinger MM. Prognostic indicators of survival in patients with compensated and decompensated cirrhosis. Liver Int. Oct 2012;32(9):1407-1414. [doi: 10.1111/j.1478-3231.2012.02830.x] [Medline: 22679906]
3. Holvast A, Verkade HJ. Complications of liver cirrhosis. In: D'Antiga L, editor. Pediatric Hepatology and Liver Transplantation. Springer International Publishing; 2019:293-297. [doi: 10.1007/978-3-319-96400-3_18]
4. Durand F, Valla D. Assessment of prognosis of cirrhosis. Semin Liver Dis. Feb 2008;28(1):110-122. [doi: 10.1055/s-2008-1040325] [Medline: 18293281]
5. Delik A, Ülger Y. Treatment approach in patients with decompensated liver cirrhosis. In: Advances in Hepatology. IntechOpen; 2021. [doi: 10.5772/intechopen.96155]
6. Volk ML, Tocco RS, Bazick J, Rakoski MO, Lok AS. Hospital readmissions among patients with decompensated cirrhosis. Am J Gastroenterol. Feb 2012;107(2):247-252. [doi: 10.1038/ajg.2011.314] [Medline: 21931378]
7. Ginès P, Krag A, Abraldes JG, Solà E, Fabrellas N, Kamath PS. Liver cirrhosis. The Lancet. Oct 2021;398(10308):1359-1376. [doi: 10.1016/S0140-6736(21)01374-X]
8. Gustot T, Stadlbauer V, Laleman W, Alessandria C, Thursz M. Transition to decompensation and acute-on-chronic liver failure: role of predisposing factors and precipitating events. J Hepatol. Jul 2021;75 Suppl 1:S36-S48. [doi: 10.1016/j.jhep.2020.12.005] [Medline: 34039491]
9. Thiele M, Johansen S, Israelsen M, et al. Noninvasive assessment of hepatic decompensation. Hepatology. Mar 1, 2025;81(3):1019-1037. [doi: 10.1097/HEP.0000000000000618] [Medline: 37801593]
10. D'Amico G, Bernardi M, Angeli P. Towards a new definition of decompensated cirrhosis. J Hepatol. Jan 2022;76(1):202-207. [doi: 10.1016/j.jhep.2021.06.018] [Medline: 34157322]
11. Addissouky TA. Liver stiffness as a dynamic predictor of decompensation and mortality in alcohol-related liver disease. Discov Med. Nov 2024;1(1):96. [doi: 10.1007/s44337-024-00083-x]
12. Chon YE, Jung ES, Park JY, et al. The accuracy of noninvasive methods in predicting the development of hepatocellular carcinoma and hepatic decompensation in patients with chronic hepatitis B. J Clin Gastroenterol. Jul 2012;46(6):518-525. [doi: 10.1097/MCG.0b013e31825079f1] [Medline: 22688146]
13. Assefi M. Comparison of the prognostic value of the Child Pugh and MELD scoring systems in cirrhosis -A systematic review. J Critical Care & Emerg Med. Dec 2022;1(4):1. [doi: 10.47363/JCCEM/2022(1)115]
14. Schneider ARP, Schneider CV, Schneider KM, et al. Early prediction of decompensation (EPOD) score: non-invasive determination of cirrhosis decompensation risk. Liver Int. Mar 2022;42(3):640-650. [doi: 10.1111/liv.15161] [Medline: 35007409]
15. Paukner M, Ladner DP, CAPriCORN Team, Zhao L. Measuring disease burden with individual cumulative incidence in patients with cirrhosis. J Biomed Inform. Sep 2025;169:104883. [doi: 10.1016/j.jbi.2025.104883] [Medline: 40752671]
16. Mostafa F, Hasan E, Williamson M, Khan H. Statistical machine learning approaches to liver disease prediction. Livers. Dec 2021;1(4):294-312. [doi: 10.3390/livers1040023]
17. Rožman M, Kišić A, Oreški D. Comparative analysis of nonlinear models developed using machine learning algorithms. WSEAS TRANSACTIONS ON INFORMATION SCIENCE AND APPLICATIONS. Jan 2, 2024;21:303-307. [doi: 10.37394/23209.2024.21.29]
18. Ravi D, Wong C, Deligianni F, et al. Deep learning for health informatics. IEEE J Biomed Health Inform. Jan 2017;21(1):4-21. [doi: 10.1109/JBHI.2016.2636665] [Medline: 28055930]
19. Hama T, Alsaleh MM, Allery F, et al. Enhancing patient outcome prediction through deep learning with sequential diagnosis codes from structured electronic health record data: systematic review. J Med Internet Res. Mar 18, 2025;27(1):e57358. [doi: 10.2196/57358] [Medline: 40100249]
20. Sasseville M, Ouellet S, Rhéaume C, et al. Bias mitigation in primary health care artificial intelligence models: scoping review. J Med Internet Res. Jan 7, 2025;27(1):e60269. [doi: 10.2196/60269] [Medline: 39773888]
21. Hsu CY, Parikh ND, Huo TI, Tapper EB. Comparison of seven noninvasive models for predicting decompensation and hospitalization in patients with cirrhosis. Dig Dis Sci. Dec 2021;66(12):4508-4517. [doi: 10.1007/s10620-020-06763-9] [Medline: 33387126]
22. Spann A, Yasodhara A, Kang J, et al. Applying machine learning in liver disease and transplantation: a comprehensive review. Hepatology. Mar 2020;71(3):1093-1105. [doi: 10.1002/hep.31103] [Medline: 31907954]
23. Leiby BE, Ten Have TR, Lynch KG, Sammel MD. Bayesian multivariate growth curve latent class models for mixed outcomes. Stat Med. Sep 10, 2014;33(20):3434-3452. [doi: 10.1002/sim.5596] [Medline: 22961883]

24. Yu R, Zheng Y, Zhang R, Jiang Y, Poon CCY. Using a multi-task recurrent neural network with attention mechanisms to predict hospital mortality of patients. IEEE J Biomed Health Inform. Feb 2020;24(2):486-492. [doi: 10.1109/JBHI.2019.2916667] [Medline: 31094697]
25. Pai S, Bader GD. Patient similarity networks for precision medicine. J Mol Biol. Sep 14, 2018;430(18 Pt A):2924-2938. [doi: 10.1016/j.jmb.2018.05.037] [Medline: 29860027]
26. Laferrière-Langlois P, Imrie F, Geraldo MA, et al. Novel preoperative risk stratification using digital phenotyping applying a scalable machine-learning approach. Anesth Analg. Jul 1, 2024;139(1):174-185. [doi: 10.1213/ANE.0000000000006753] [Medline: 38051671]
27. Sheng JQ, Hu PJH, Liu X, Huang TS, Chen YH. Predictive analytics for care and management of patients with acute diseases: deep learning-based method to predict crucial complication phenotypes. J Med Internet Res. Feb 12, 2021;23(2):e18372. [doi: 10.2196/18372] [Medline: 33576744]
28. Heo S, Lee SS, Kim SY, et al. Prediction of decompensation and death in advanced chronic liver disease using deep learning analysis of gadoxetic acid-enhanced MRI. Korean J Radiol. Dec 2022;23(12):1269-1280. [doi: 10.3348/kjr.2022.0494] [Medline: 36447415]
29. Trebicka J, Macnaughtan J, Schnabl B, Shawcross DL, Bajaj JS. The microbiota in cirrhosis and its role in hepatic decompensation. J Hepatol. Jul 2021;75 Suppl 1(Suppl 1):S67-S81. [doi: 10.1016/j.jhep.2020.11.013] [Medline: 34039493]
30. Hsieh MS, Cheng KC, Hsieh ML, Chiang JH, Hsieh VCR. Mortality risk and decompensation in hospitalized patients with non-alcoholic liver cirrhosis: implications for disease management. Int J Environ Res Public Health. Jan 12, 2021;18(2):606. [doi: 10.3390/ijerph18020606] [Medline: 33445719]
31. Idilman R, Aydogan M, Oruncu MB, et al. Natural history of cirrhosis: changing trends in etiology over the years. Dig Dis. 2021;39(4):358-365. [doi: 10.1159/000512746] [Medline: 33142287]
32. Li M, Wei Z, Su J, et al. Changing spectrum and mortality disparities of etiology of liver cirrhosis in Beijing, China. J Med Virol. Jan 2024;96(1):e29405. [doi: 10.1002/jmv.29405] [Medline: 38235623]
33. Tsochatzis EA, Bosch J, Burroughs AK. Liver cirrhosis. The Lancet. May 2014;383(9930):1749-1761. [doi: 10.1016/S0140-6736(14)60121-5]
34. Ginés P, Quintero E, Arroyo V, et al. Compensated cirrhosis: natural history and prognostic factors. Hepatology. 1987;7(1):122-128. [doi: 10.1002/hep.1840070124]
35. Angeli P, Bernardi M, Villanueva C, et al. EASL clinical practice guidelines for the management of patients with decompensated cirrhosis. J Hepatol. Aug 2018;69(2):406-460. [doi: 10.1016/j.jhep.2018.03.024]
36. Rubin JB, Sundaram V, Lai JC. Gender differences among patients hospitalized with cirrhosis in the United States. J Clin Gastroenterol. Jan 2020;54(1):83-89. [doi: 10.1097/MCG.0000000000001192] [Medline: 30807399]
37. Liu S, Li J, Wong Y, et al. Prognostic impact of age on outcomes of hepatic decompensation in patients with compensated cirrhosis (CHESS2102): an international, multicenter cohort study. MedComm (2020). Nov 2024;5(11):e781. [doi: 10.1002/mco2.781] [Medline: 39492833]
38. Mansour D, McPherson S. Management of decompensated cirrhosis. Clinical Medicine. Apr 2018;18(2):s60-s65. [doi: 10.7861/clinmedicine.18-2-s60]
39. Gordon FD. Ascites. Clin Liver Dis. May 2012;16(2):285-299. [doi: 10.1016/j.cld.2012.03.004] [Medline: 22541699]
40. McKay R, Webster NR. Variceal bleeding. Continuing Education in Anaesthesia Critical Care & Pain. Dec 2007;7(6):191-194. [doi: 10.1093/bjaceaccp/mkm038]
41. Ferenci P. Hepatic encephalopathy. Gastroenterol Rep. May 2017;5(2):138-147. [doi: 10.1093/gastro/gox013]
42. Simonetto DA, Gines P, Kamath PS. Hepatorenal syndrome: pathophysiology, diagnosis, and management. BMJ. Sep 14, 2020;370:m2687. [doi: 10.1136/bmj.m2687] [Medline: 32928750]
43. McCoy MS, Angeli P, Trebicka J. Acute and non-acute decompensation of liver cirrhosis. Liver International. Mar 2025;45(3):e15861. [doi: 10.1111/liv.15861]
44. Baiges A, Hernández-Gea V. Management of liver decompensation in advanced chronic liver disease: ascites, hyponatremia, and gastroesophageal variceal bleeding. Clin Drug Investig. Jun 2022;42(Suppl 1):25-31. [doi: 10.1007/s40261-022-01147-5] [Medline: 35476218]
45. Peck-Radosavljevic M, Angeli P, Cordoba J, Farges O, Valla D. Managing complications in cirrhotic patients. UEG Journal. Feb 2015;3(1):80-94. [doi: 10.1177/2050640614560452]
46. Cardenas A, Ginès P. Management of complications of cirrhosis in patients awaiting liver transplantation. J Hepatol. 2005;42 Suppl(1):S124-33. [doi: 10.1016/j.jhep.2004.12.007] [Medline: 15777567]
47. Verma N, Roy A, Valsan A, et al. Liver dysfunction and systemic inflammation drive organ failures in acute decompensation of cirrhosis: a multicentric study. Am J Gastroenterol. 2025;120(1):182-193. [doi: 10.14309/ajg.0000000000003115]

48. Kumar S, Rani P. A comparative survey on machine learning techniques for prediction of liver disease. Presented at: 2023 6th International Conference on Contemporary Computing and Informatics (IC3I); Sep 14-16, 2023:1796-1801; Gautam Buddha Nagar, India. [doi: 10.1109/IC3I59117.2023.10397980]
49. Luo X. A comparison of three estimation methods in linear regression analysis. Presented at: 2016 4th International Conference on Machinery, Materials and Information Technology Applications; Dec 10-11, 2016; Xi'an, China. [doi: 10.2991/icmmita-16.2016.92]
50. Leise MD, Kim WR, Kremers WK, Larson JJ, Benson JT, Therneau TM. A revised model for end-stage liver disease optimizes prediction of mortality among patients awaiting liver transplantation. Gastroenterology. Jun 2011;140(7):1952-1960. [doi: 10.1053/j.gastro.2011.02.017] [Medline: 21334338]
51. Olive DJ. Linear regression analysis. Technometrics. Nov 2003;45(4):362-363. [doi: 10.1198/tech.2003.s163]
52. Huang JC, Ko KM, Shu MH, Hsu BM. Application and comparison of several machine learning algorithms and their integration models in regression problems. Neural Comput & Applic. May 2020;32(10):5461-5469. [doi: 10.1007/s00521-019-04644-5]
53. Clark S, Hyndman RJ, Pagendam D, Ryan LM. Modern strategies for time series regression. arXiv. Preprint posted online on Oct 29, 2020. [doi: 10.48550/arXiv.2010.15997]
54. Andrews PJD, Sleeman DH, Statham PFX, et al. Predicting recovery in patients suffering from traumatic brain injury by using admission variables and physiological data: a comparison between decision tree analysis and logistic regression. J Neurosurg. Aug 2002;97(2):326-336. [doi: 10.3171/jns.2002.97.2.0326] [Medline: 12186460]
55. Wu CC, Yeh WC, Hsu WD, et al. Prediction of fatty liver disease using machine learning algorithms. Comput Methods Programs Biomed. Mar 2019;170:23-29. [doi: 10.1016/j.cmpb.2018.12.032] [Medline: 30712601]
56. Li X, Chen X, Yuan Z. Applicable model of liver disease detection based on the improved CART-AdaBoost algorithm. Presented at: 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA); Jun 28-30, 2021:1177-1181; Dalian, China. [doi: 10.1109/ICAICA52286.2021.9498046]
57. Afreen N, Patel R, Ahmed M, Sameer M. A novel machine learning approach using boosting algorithm for liver disease classification. Presented at: 2021 5th International Conference on Information Systems and Computer Networks (ISCON); Oct 22-23, 2021:1-5; Mathura, India. [doi: 10.1109/ISCON52037.2021.9702488]
58. Dalal S, Onyema EM, Malik A. Hybrid XGBoost model with hyperparameter tuning for prediction of liver disease with better accuracy. World J Gastroenterol. Dec 14, 2022;28(46):6551-6563. [doi: 10.3748/wjg.v28.i46.6551] [Medline: 36569269]
59. Krogh A. What are artificial neural networks? Nat Biotechnol. Feb 2008;26(2):195-197. [doi: 10.1038/nbt1386] [Medline: 18259176]
60. Breiman L. Random forests. Mach Learn. Oct 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
61. Dorffner G, Porenta G. On using feedforward neural networks for clinical diagnostic tasks. Artif Intell Med. Oct 1994;6(5):417-435. [doi: 10.1016/0933-3657(94)90005-1] [Medline: 7842041]
62. He L, Madathil SC, Servis G, Khasawneh MT. Neural network-based multi-task learning for inpatient flow classification and length of stay prediction. Appl Soft Comput. Sep 2021;108:107483. [doi: 10.1016/j.asoc.2021.107483] [Medline: 34658687]
63. Lai Y. Application of machine learning algorithm based on neural network technology. J Phys: Conf Ser. Nov 1, 2021;2066(1):012041. [doi: 10.1088/1742-6596/2066/1/012041]
64. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol. Nov 1996;49(11):1225-1231. [doi: 10.1016/s0895-4356(96)00002-9] [Medline: 8892489]
65. Mao S, Sejdic E. A review of recurrent neural network-based methods in computational physiology. IEEE Trans Neural Netw Learn Syst. Oct 2023;34(10):6983-7003. [doi: 10.1109/TNNLS.2022.3145365] [Medline: 35130174]
66. Liu S, Schlesinger JJ, McCoy AB, et al. New onset delirium prediction using machine learning and long short-term memory (LSTM) in electronic health record. J Am Med Inform Assoc. Dec 13, 2022;30(1):120-131. [doi: 10.1093/jamia/ocac210]
67. Shuvo MMH, Islam SK. Deep multitask learning by stacked long short-term memory for predicting personalized blood glucose concentration. IEEE J Biomed Health Inform. Mar 2023;27(3):1612-1623. [doi: 10.1109/JBHI.2022.3233486] [Medline: 37018303]
68. Subasi A. Chapter 3 - Machine learning techniques. In: Subasi A, editor. Practical Machine Learning for Data Analysis Using Python. Academic Press; 2020:91-202. [doi: 10.1016/B978-0-12-821379-7.00003-5]
69. Singh Y, Faghani S, Eaton JE, Venkatesh SK, Erickson BJ. Deep learning-based prediction of hepatic decompensation in patients with primary sclerosing cholangitis with computed tomography. Mayo Clin Proc Digit Health. Sep 2024;2(3):470-476. [doi: 10.1016/j.mcpdig.2024.07.002] [Medline: 40206109]

70. Kok C, Jahmunah V, Oh SL, et al. Automated prediction of sepsis using temporal convolutional network. Comput Biol Med. Dec 2020;127:103957. [doi: 10.1016/j.compbiomed.2020.103957] [Medline: 32938540]
71. Li C, Ding S, Zou N, Hu X, Jiang X, Zhang K. Multi-task learning with dynamic re-weighting to achieve fairness in healthcare predictive modeling. J Biomed Inform. Jul 2023;143:104399. [doi: 10.1016/j.jbi.2023.104399] [Medline: 37211197]
72. Ramesh S, Dall'Alba D, Gonzalez C, et al. Multi-task temporal convolutional networks for joint recognition of surgical phases and steps in gastric bypass procedures. Int J Comput Assist Radiol Surg. Jul 2021;16(7):1111-1119. [doi: 10.1007/s11548-021-02388-z] [Medline: 34013464]
73. Tang S, Tariq A, Dunnmon JA, et al. Predicting 30-day all-cause hospital readmission using multimodal spatiotemporal graph neural networks. IEEE J Biomed Health Inform. Apr 2023;27(4):2071-2082. [doi: 10.1109/JBHI.2023.3236888] [Medline: 37018684]
74. Ma Q, Meng R, Li R, et al. Multitask deep learning model based on multimodal data for predicting prognosis of rectal cancer: a multicenter retrospective study. BMC Med Inform Decis Mak. Jun 5, 2025;25(1):209. [doi: 10.1186/s12911-025-03050-3] [Medline: 40474195]
75. Hu X, Sun Z, Nian Y, et al. Self-explainable graph neural network for Alzheimer disease and related dementias risk prediction: algorithm development and validation study. JMIR Aging. Jul 8, 2024;7(1):e54748. [doi: 10.2196/54748] [Medline: 38976869]
76. Bejani MM, Ghatee M. A systematic review on overfitting control in shallow and deep neural networks. Artif Intell Rev. Dec 2021;54(8):6391-6438. [doi: 10.1007/s10462-021-09975-1]
77. Zhou K, Huang X, Chen M, et al. Pre-hospital symptom clusters and symptom network analysis in decompensated cirrhotic patients: a cross-sectional study. J Adv Nurs. Jul 2024;80(7):2785-2800. [doi: 10.1111/jan.16044] [Medline: 38197541]
78. Bernardi M, Caraceni P. Novel perspectives in the management of decompensated cirrhosis. Nat Rev Gastroenterol Hepatol. Dec 2018;15(12):753-764. [doi: 10.1038/s41575-018-0045-2] [Medline: 30026556]
79. Fedosina EA, Byeverov AO, Bogomolov PO, Staroverova NP. Some practical issues in the management of patients with decompensated liver cirrhosis. Ter Arkh. Aug 15, 2019;91(8):148-154. [doi: 10.26442/00403660.2019.08.000391] [Medline: 32598767]
80. Caruana R. Multitask learning. Mach Learn. Jul 1997;28(1):41-75. [doi: 10.1023/A:1007379606734]
81. Roy S, Mincu D, Loreaux E, et al. Multitask prediction of organ dysfunction in the intensive care unit using sequential subnetwork routing. J Am Med Inform Assoc. Aug 13, 2021;28(9):1936-1946. [doi: 10.1093/jamia/ocab101] [Medline: 34151965]
82. Asrani SK, Hall L, Reddy V, Ogola G, Izzy M. Comorbid chronic diseases and survival in compensated and decompensated cirrhosis: a population-based study. Am J Gastroenterol. 2022;117(12):2009-2016. [doi: 10.14309/ajg.0000000000001909]
83. Berenguer J, Aldámiz-Echevarría T, Hontañón V, et al. Clinical outcomes and prognostic factors after HCV clearance with DAA in HIV/HCV-coinfected patients with advanced fibrosis/cirrhosis. Hepatology. Jan 1, 2025;81(1):238-253. [doi: 10.1097/HEP.0000000000000838] [Medline: 38452004]
84. Innes H, Walker AJ, Benselin J, et al. Comprehensive comparative analysis of standard validated, genetic, and novel biomarkers to enhance prognostic risk-stratification in patients with hepatitis C virus cirrhosis. Clin Transl Gastroenterol. Feb 9, 2022;13(3):e00462. [doi: 10.14309/ctg.0000000000000462] [Medline: 35142723]
85. Rudler M. Role of comorbidities in the general management of compensated cirrhosis, including malnutrition. In: Portal Hypertension VI. Springer International Publishing; 2016:205-214. [doi: 10.1007/978-3-319-23018-4_22]
86. Konerman MA, Beste LA, Van T, et al. Machine learning models to predict disease progression among veterans with hepatitis C virus. PLOS ONE. 2019;14(1):e0208141. [doi: 10.1371/journal.pone.0208141] [Medline: 30608929]
87. Gao Y, Li R, Croxford E, et al. Leveraging medical knowledge graphs into large language models for diagnosis prediction: design and application study. JMIR AI. Feb 24, 2025;4(1):e58670. [doi: 10.2196/58670] [Medline: 39993309]
88. Xu M. Understanding graph embedding methods and their applications. SIAM Rev. Jan 2021;63(4):825-853. [doi: 10.1137/20M1386062]
89. Abdi H, Williams LJ. Principal component analysis. WIREs Computational Stats. Jul 2010;2(4):433-459. [doi: 10.1002/wics.101]
90. Yang H, Zhang B, Peng C, Yu C, Mou X. Predictive Value of serum lactate levels for mortality in patients with hepatitis B-related decompensated cirrhosis: a retrospective analysis. Br J Hosp Med. Dec 30, 2024;85(12):1-10. [doi: 10.12968/hmed.2024.0531]

91. Salkić N, Jovanović P, Barišić Jaman M, Selimović N, Paštrović F, Grgurević I. Machine learning for short-term mortality in acute decompensation of liver cirrhosis: better than MELD score. Diagnostics (Basel). May 8, 2024;14(10):981. [doi: 10.3390/diagnostics14100981] [Medline: 38786278]
92. Aurelio YS, de Almeida GM, de Castro CL, Braga AP. Learning from imbalanced data sets with weighted cross-entropy function. Neural Process Lett. Oct 2019;50(2):1937-1949. [doi: 10.1007/s11063-018-09977-1]
93. Andrade C. Z scores, standard scores, and composite test scores explained. Indian J Psychol Med. Nov 2021;43(6):555-557. [doi: 10.1177/02537176211046525] [Medline: 35210687]
94. Cervellera C, Maccio D. Distribution-preserving stratified sampling for learning problems. IEEE Trans Neural Netw Learn Syst. Jul 2018;29(7):2886-2895. [doi: 10.1109/TNNLS.2017.2706964] [Medline: 28613186]
95. Soleymani R, Granger E, Fumera G. F-measure curves: a tool to visualize classifier performance under imbalance. Pattern Recognit. Apr 2020;100:107146. [doi: 10.1016/j.patcog.2019.107146]
96. Kasper P, Goeser T. Management of decompensated liver cirrhosis in the surgical intensive care unit. Hepatobiliary Surg Nutr. Feb 28, 2023;12(1):95-98. [doi: 10.21037/hbsn-22-553] [Medline: 36860248]
97. Lundberg S, Lee SI. A unified approach to interpreting model predictions. arXiv. Preprint posted online on Nov 25, 2017. [doi: 10.48550/arXiv.1705.07874]
98. Müller SE, Casper M, Ripoll C, et al. Machine learning models predicting decompensation in cirrhosis. J Gastrointestin Liver Dis. Mar 27, 2025;34(1):71-80. [doi: 10.15403/jgld-5876] [Medline: 40153826]
99. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep. Jan 2019;49(1):15-21. [doi: 10.1002/hast.973] [Medline: 30790315]
100. Abolhasani M. Cirrhosis. GitHub. URL: https://github.com/abolhasani/Cirrhosis [Accessed 2026-10-09]

Abbreviations

ADAB: Adaptive Boosting (AdaBoost)
AUC: area under the curve
BCE: binary cross-entropy
CIS: condition impact score
CNN : convolutional neural network
DNN: deep neural network
EHR: electronic health record
FCN : Fully Connected Neural Network
GNN : graph neural network
HBV : hepatitis B virus
HCV : hepatitis C virus
HE: hepatic encephalopathy
HRS: hepatorenal syndrome
IF : influence factor
LSTM: long short-term memory
M4T: Four-outcome Multitask Baseline Proposed Model
MLD: mild liver disease
PCA: principal component analysis
pp: percentage point
RF: random forest
SHAP: Shapley Additive Explanations
TCN: temporal convolutional neural network
XGB: Extreme Gradient Boosting

Edited by Andrew Coristine; peer-reviewed by Jessica Qiuhua Sheng, Jiahao Meng; submitted 18.Jan.2026; final revised version received 14.Sep.2026; accepted 17.Sep.2026; published 13.Nov.2026

Please cite as:
Abolhasani M, Xu D, Hu PJH, Huang TS
Predicting Cirrhotic Patients' Transition From Compensated to Decompensated Stage: A Graph-Based Multitask Deep Learning Method
JMIR Med Inform 2026;14:e91693
URL: https://medinform.jmir.org/2026/1/e91693
doi: 10.2196/91693

© Mohammadsadegh Abolhasani, Da Xu, Paul Jen-Hwa Hu, Ting-Shuo Huang. Originally published in JMIR Medical Informatics (https://medinform.jmir.org), 13.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on https://medinform.jmir.org/, as well as this copyright and license information must be included.