Early Prognosis of Metabolic Dysfunction-Associated Fatty Liver Disease Using Deep Learning and Clinical Data Analysis

Jeannie She
jeanshe@mit.edu

Erica Song
ericason@mit.edu

Naiqi Zhang
naiqiz79@mit.edu

Abstract

Metabolic dysfunction-associated fatty liver disease (MAFLD) affects 25% of adults in the United States and affects those with Type 2 diabetes and class III obesity at disproportionately higher rates. We present a deep learning-based framework for early prognosis of MAFLD in adults using structured clinical data from Mass General Brigham. Our approach utilizes binary classification, neural network prediction, linear and logistic regression, and survival modeling, as well as experimentation with addressing class imbalance. The study supports early clinical risk stratification and reveals predictive biomarkers using SHAP interpretation.

1 Introduction

Metabolic dysfunction-associated fatty liver disease (MAFLD), formerly known as nonalcoholic fatty liver disease (NAFLD), is a widespread and progressive condition linked to substantial morbidity and mortality [1]. In its early stages, MAFLD is reversible through lifestyle interventions. However, it frequently advances undetected to metabolic dysfunction-associated steatohepatitis (MASH), a more severe state characterized by hepatic lipid accumulation, inflammation, hepatocellular injury, and fibrosis — potentially leading to cirrhosis or hepatocellular carcinoma [1, 3]. Accurate and early prognosis of MAFLD progression could enable timely intervention and reduce downstream health risks.

We expand upon prior work on the Mass General Brigham (MGB) Research Patient Data Repository (RPDR), where logistic regression and tree-based models were used to predict disease progression in MAFLD patients aged 30 and older [3]. Those models were trained on structured tabular data, including diagnoses, procedures, medications, physical measurements, and laboratory results.

In this project, our contributions are as follows: (1) we develop and evaluate deep learning architectures — including binary neural network classifiers, time-to-event neural network models, and a Survival Net (DeepSurv) model — to improve predictive performance over standard regression baselines [2]; (2) We address the severe class imbalance in progression outcomes experimenting with SMOTE, downsampling, and class-weighted loss; (3) we interpret model outputs with SHAP values and coefficient analysis to highlight key clinical risk factors and compare with important factors (such as diabetes and obesity) identified in previous studies on MAFLD.

2 Related Work

Our work builds directly upon the thesis by Jonathan Li (2024) [3], who trained different machine learning models to predict progression from MAFLD to more advanced outcomes such as MASH, hepatic failure, cirrhosis, and liver cancer using EHR data from the MGB RPDR [3]. His models

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

included logistic regression, decision trees, random forest, and XGBoost, achieving an AUROC of up to 0.80. Despite strong performance on some discrimination metrics, all models suffered from low AUPRC and sensitivity, reflecting challenges related to class imbalance and potential overfitting. Feature inputs included demographics, lab results (using means for LOINC-coded test results), procedures, physical measurements, medications, and diagnostic codes. In his thesis, Li also explored large language models (LLMs) for note-based classification, though performance was limited by the format and quality of the text data [3].

In our study, we followed a similar approach to data preprocessing to construct a comparable tabular dataset. This included filtering for MAFLD-diagnosed patients age 30 and above, limiting to features recorded within 2 years prior to MAFLD diagnosis, and excluding patients with other liver or substance-related conditions. We also experimented with the same class imbalance strategies - specifically SMOTE oversampling and majority-class downsampling - as explored in the thesis.

To extend his work into time-to-event prediction, we adopted DeepSurv, a Cox proportional hazards deep neural network proposed by Katzman et al. (2018) [2]. DeepSurv replaces the linear log-risk assumption in traditional Cox models with a neural network architecture, allowing for more expressive, nonlinear modeling of covariate interactions. The network is trained using a negative partial likelihood loss and supports risk ranking as well as personalized treatment recommendations. DeepSurv has demonstrated competitive performance across real and simulated datasets, outperforming both linear Cox models and random survival forests in multiple settings.

Together, these foundational studies shape our methodological direction, allowing us to evaluate the comparative advantages of deep learning architectures over traditional classifiers and survival models in the context of liver disease prognosis.

3 Methods

3.1 Overview and Problem Formulation

The problem we aim to solve is the early prediction of progression from metabolic dysfunction-associated fatty liver disease (MAFLD) to more advanced liver outcomes such as MASH, cirrhosis, or hepatocellular carcinoma. Since early-stage MAFLD is potentially reversible, this task is clinically meaningful. The dataset is highly imbalanced, with only a small proportion (4.52%) of patients experiencing progression. Our modeling challenge is twofold: (1) learn effective representations of structured clinical data to predict future progression, and (2) ensure that predictive models generalize well to underrepresented outcomes while maintaining interpretability.

We assume that structured features (labs, diagnoses, medications, physical measurements) contain meaningful signals of disease risk. We further assume that incorporating time-to-event data using survival models can add value beyond binary classification.

3.2 Modeling Approaches

We implemented and compared three modeling architectures.

3.2.1 Binary Neural Network Model

This model performs binary classification to predict whether a patient progresses to advanced liver conditions. Figure 1 below performs outcome classification:

- **Inputs**: Vector of 2850 clinical features (e.g., labs, demographics, codes).
- **Outputs**: Logits for categories of progression vs. censorship.
- **Loss function**: Binary Cross Entropy with Logits Loss (`BCEWithLogitsLoss`), which combines a sigmoid layer and `BCELoss` in one function.
- **Optimizer**: Adam optimizer with learning rate $1e{-}3$.
- **Training details**: Model trained on GPU with early stopping based on validation AUROC. Output logits are directly used with the `BCEWithLogitsLoss`.
- **Baselines**: Logistic regression.

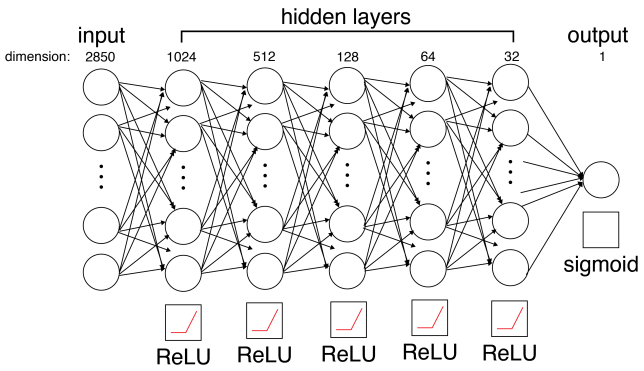


Figure 1: Binary neural network architecture.

3.2.2 Time-to-event Neural Network Model

This regression model is designed to predict the number of days until a patient progresses to advanced liver disease. Figure 2 below performs Time-to-event prediction:

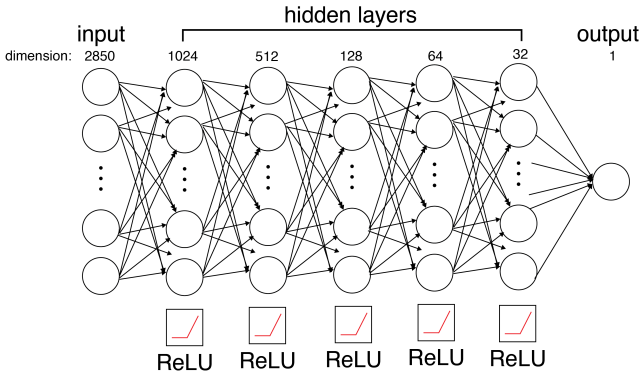


Figure 2: Quantitative neural network architecture.

- **Inputs**: Vector of 2850 clinical features (e.g., labs, medications, diagnoses).
- **Outputs**: Predicted days to progression.
- **Loss function**: Mean Squared Error Loss (`MSELoss`).
- **Optimizer**: Adam optimizer with learning rate $1e{-}3$.
- **Baselines**: Linear regression.
- **Training details**: Model trained using mini-batch gradient descent for 50 epochs, with a train/test split of 70/30. Predictions were further interpreted using SHAP values to quantify the contribution of each clinical feature to the predicted progression timeline.

3.2.3 Survival Net (DeepSurv) Model

We implemented a deep learning-based Cox proportional hazards model, known as DeepSurv, using the `pycox` library to predict time-to-progression. The architecture is a simple feedforward neural network trained with partial likelihood loss and optimized for concordance. Figure 3 below represents the DeepSurv model that predicts time-to-progression:

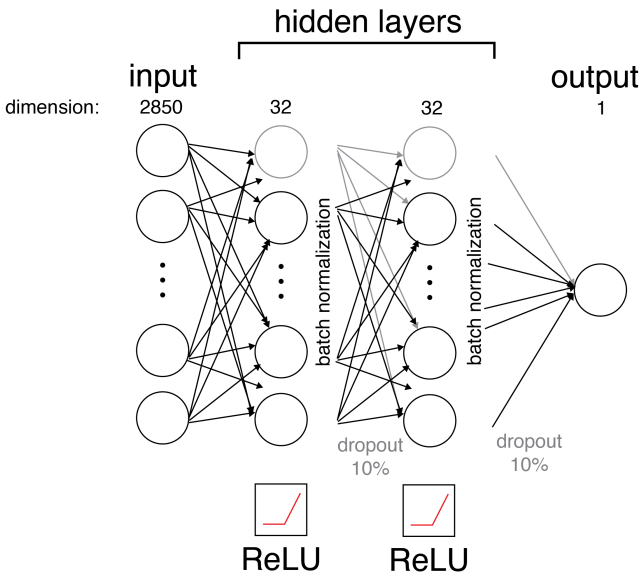


Figure 3: Cox PH DeepSurv architecture.

- **Inputs**: Vector of 2850 clinical features (e.g., labs, medications, diagnoses).
- **Outputs**: Hazard function.
- **Architecture**:
  - Implemented via `torchtuples.practical.MLPVanilla`.
  - Two hidden layers with 32 neurons each.
  - ReLU activation, batch normalization enabled, dropout rate of 0.1.
  - Output layer: 1 neuron with no bias term and linear activation.
- **Loss function**: Negative log Cox partial likelihood.
- **Optimizer**: `tt.optim.Adam`. Learning rate was determined using a learning rate finder based on a previous work by Dr. Leslie N. Smith [6].
- **Training setup**:
  - Batch size: 256
  - Epochs: up to 512
  - Early stopping based on validation loss
- **Visualization**: Model predictions were visualized using mean survival functions across groups split by feature values (such as insulin vs. no insulin), enabling an interpretable assessment of time-to-event separation.

3.2.4 Baseline Models

We trained and evaluated three baseline machine learning models for the three different outcomes being predicted. For the time-to-event prediction, we trained a linear regression model on all of the data used in the quantitative, time-to-event neural network. For binary classification, we trained a logistic regression on the same data used to train the binary neural network. Finally, for survival analysis, we trained a Cox regression model on coefficients selected via Lasso regression and compared it to the DeepSurv model.

The linear, logistic, and Lasso regression models were implemented with the `sklearn` package, while Cox regression was implemented using `lifelines`. We used default hyperparameters provided by the package.

3.3 Combating Class Imbalance

To address the imbalance in progression outcomes of the binary classification model, we implemented and benchmarked four techniques. Ultimately, we did not use any of these techniques because they resulted in very poor performance with in AUROC of around 0.5. However, these attempts are documented because they are worth further experimentation.

SMOTE (Synthetic Minority Over-sampling Technique): We applied SMOTE to address the extreme class imbalance present in the binary classification task (progression vs. non-progression). SMOTE is an oversampling method that generates synthetic samples of the minority class by interpolating between existing samples and their nearest neighbors in feature space. This approach is supposed to duplicate points and improves the generalizability of minority class representations during training.

In our implementation, SMOTE was applied **only to the scaled training set** to prevent information leakage into the test set. SMOTE was performed with `random_state=42` for reproducibility. Figure 4 below shows class distribution before and after SMOTE:

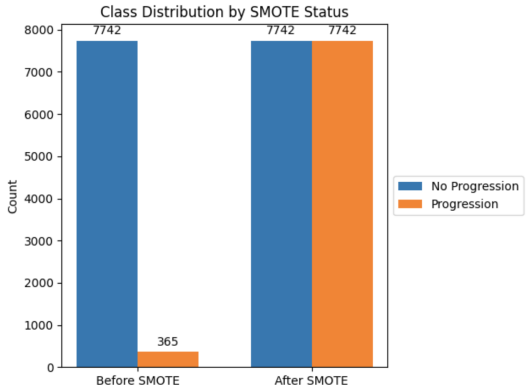


Figure 4: Class Distribution Before and After SMOTE

This resulted in a fully balanced training set with 15,838 samples in total. While SMOTE is designed to mitigate poor recall in the minority class by generating synthetic examples, in our case, the synthetic data points were often clustered too closely in feature space. As a result, the model overfit to the oversampled minority class and began predicting all cases as positive for progression. Consequently, although AUROC improved during training, the test AUROC remains at around 0.5, further indicating lack of generalizability.

Downsampling: To balance the output classes in the training data, we downsampled the majority class of patients who did not progress to match the number of patients in the minority class.

We repeated downsampling 10 times (with random seed 42 for reproducibility) and trained 10 different models on these samples. Performance was measured via two different methods: the average AUROC across the 10 samples and the AUROC resulting from an average of logits of the 10 samples.

Because there were so few patients who progressed in the dataset, this method of downsampling left 730 rows of data in the training set, with 365 in each outcome class.

The model performed very poorly with an average training AUROC of 0.5322 and a testing AUROC of 0.5027 across 10 samples, indicating that the model did not learn relationships between the features and outcome effectively. Averaging across all the model logits to make a single prediction for each patient also resulted in poor performance with training AUROC of 0.5037 and testing AUROC of 0.4980.

On further inspection, we found that all of the predictions were of the original majority class, predicting that the patient does not progress to MASH. The logits were also very close to zero, suggesting

that changing the decision threshold would not help in this case. It is likely the model could not learn patterns well enough from the training data because the dataset was so small after downsampling.

Loss Reweighting: We implemented class-based loss reweighting to improve recall on the minority class (progression). Initially, we computed weights inversely proportional to class frequency - assigning higher weights to underrepresented samples using the formula below:

$$\text{weight} = \frac{\text{total}}{\text{class count}}$$

However, the conservative weighting was not enough to shift the model's decision boundary; as a result, it continued to favor predicting the no-progression outcome, with test AUROC remaining around 0.5 and the recall for class 1 (progression) being 0.

To address this, we manually assigned a more aggressive weight of 50.0 to the minority class and 1.0 to the majority class - heavily penalizing false negatives. These custom weights were applied to the binary cross-entropy loss (`BCEWithLogitsLoss`) by disabling its internal reduction and multiplying loss terms by their respective sample weights. To further prevent overfitting to the dominant class, we integrated early stopping into training. Validation AUROC was monitored at each epoch, and training was halted if performance failed to improve over 5 consecutive epochs. However, with this adjustment, the model still performs poorly with a test AUROC between 0.54 to 0.58. Although it began predicting a higher proportion of class 1 (progression) cases, leading to slightly improved recall score (0.3); but the precision score for class 1 (progression) remained low (0.07), indicating that the model had not truly learned to distinguish between classes but was instead biased toward overpredicting the minority class.

Downsampling and Loss Reweighting: We hypothesized that downsampling or loss reweighting alone may not be sufficient to combat the massive class imbalance in our dataset effectively, so we attempted downsampling and reweighting both on the training data, following both protocols described above for downsampling and loss reweighting (but without early stopping). However, the AUROC did not improve. The average AUROC metrics across 10 samples were 0.5293 for the training set and 0.5029 for the testing set. The AUROC derived from averaging the logits produced by the 10 different models was 0.4999 for the training set and 0.4989 for the testing set.

Again, all the predictions were of the majority class in the original dataset. This is still likely due to the small size of the data that results from downsampling.

3.4 Interpretability

One of the main challenges with deep learning models is lack of interpretability. To address this, we leveraged linear regression, Cox regression, and SHAP scores to identify the most influential features in time-to-event predictions and classification of progression, respectively. We also ran a logistic regression to identify influential features in binary classification, although we did not obtain SHAP scores for the deep learning binary classfier due to poor performance; we decided the features have little meaning as predictors. For each of the linear, logistic, and Cox regression models, we identified 10 features with the most positive coefficients and 10 features with the most negative coefficients (Tables 7, 8, 9). Then, we computed SHAP scores on the time-to-event neural network and to compare results. We considered features ranked highly by SHAP scores to be influential features in our final interpretation.

4 Data and Experiment Setup

We used structured EHR data from the MGB RPDR, specifically focusing on NAFL patients in the January 2025 request of data from RPDR. Data fields spanned multiple domains including demographics, diagnoses, labs, medications, physical measurements, encounters, and more—organized into separate sheets per domain (e.g., `Dem`, `Dia`, `Lab`, etc.).

Cohort Selection: Our cohort includes adult patients (age 30 and up) diagnosed with NAFL or MAFLD. Structured features were extracted from a 2-year window preceding each patient's first

MAFLD diagnosis. We followed Li's original cohort logic with several modifications: we retained unknown sex values (rather than removing them), dropped procedures due to redundancy, and excluded categorical physical features from neural networks [3]. Patients with diagnoses that were considered worsened liver progression, alcohol related disorders, drug disorders other than nicotine and caffeine, and other liver diseases prior to MAFLD were censored. These codes are listed in Table 6.

Variables and Feature Construction: We extracted 1836 diagnoses, 89 lab values, 920 medications, and physical measurements including BMI. Labs were aggregated using means of their quantitative values. Physical measurements included both mean BMI and most recent BMI before MAFLD diagnosis. Medications and diagnoses were one-hot encoded. After inner joining on `StudyID`, the final dataset included 11,582 rows and 2,850 features. Ultimately, we did not choose to include data from the radiology, allergy, procedures, or encounters datasets, as those were sparse, irrelevant to this prediction task, and/or repetitive of the information in the data we already included.

Preprocessing: Categorical variables were one-hot encoded. Numerical variables, including age, BMI and labs, were standardized using an instance of `StandardScaler` fit on the training set only. This same transformation was then applied to both training and test sets to avoid data leakage. Synthetic oversampling (e.g., SMOTE), downsampling, and class reweighting were applied only to the training data.

Train/Test Split and Benchmarks: We performed an 70/30 train-test split using `train_test_split` with a fixed `random_state=42` to ensure reproducibility.

5 Results

5.1 Binary Neural Network

The neural network trained for binary classification achieved a test AUROC of 0.5155 and a training AUROC of 1. This indicates extreme overfitting to the training set, and the model has little generalizable knowledge of the effects of input features on the output category.

5.2 Time-to-event Neural Network

After evaluating the model on the test set, we observed a test mean squared error (MSE) of 253,614.92 days^2, corresponding to a root mean squared error (RMSE) of approximately 503.6 days. The mean absolute error (MAE) was 375.85 days, or roughly one year. Given that the target variable — days until progression — ranges from 0 to 2,247 days (about 6.15 years), and has a mean of 830.75 days, we consider this to be a reasonably strong model performance.

5.3 DeepSurv

The Survival Net model can be evaluated using a number of different metrics, listed in table 5. The concordance statistic, identifying how well the model discriminates between individuals who experience and don't experience the progression outcome suggests that the model may have overfit on the training data. The Brier score plotted against time in Figure 7 demonstrates promising performance of the model for the duration of most patients; the sharp increase near the end of the time range is expected and is often disregarded. The integrated negative binomial log likelihood (NBLL) evaluates the model's ability to predict full survival distributions over time, while and partial log likehood (PLL) measures how well the model can predict individual's risks without estimating the baseline hazard function.

Table 1: Cox PH DeepSurv Metrics

	Concordance Index	**Integrated Brier Score**	**Integrated NBLL**	**PLL**
Train	0.858	0.0461	0.177	-7.65
Test	0.550	0.0549	0.232	-7.39

With the predicted survival curves of this model, we subset the patients based on relevant features and plot the average predicted survival probabilities. Figure 5 and 6 below display survival curve prediction of the population stratified on relevant features; Figure 7 displays Brier score plotted over time for two Cox PH models:

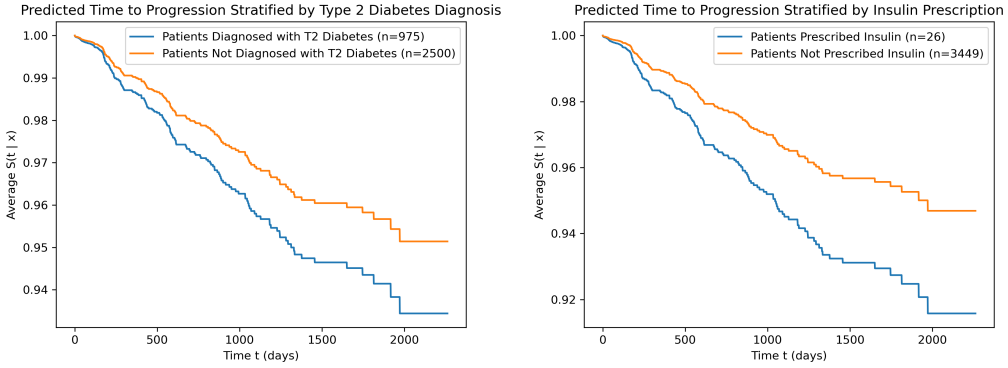


Figure 5: Time to progression survival prediction using DeepSurv stratified by Type 2 diabetes without complications diagnosis (E11.9) and Lispro insulin medication prescription (119019).

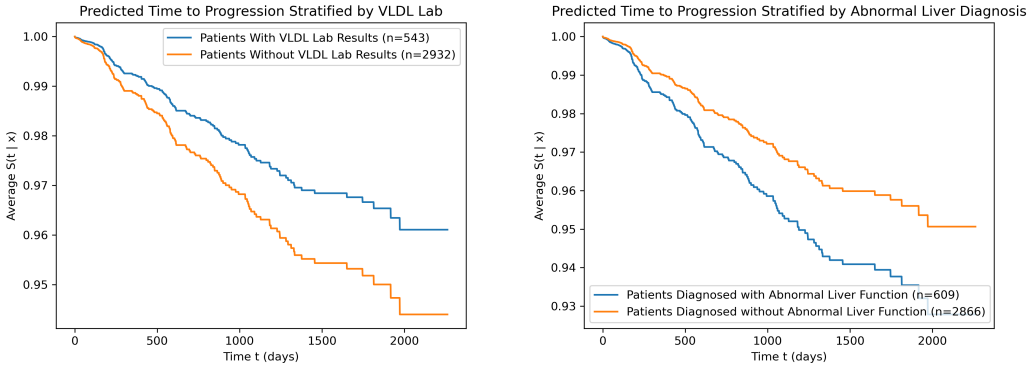
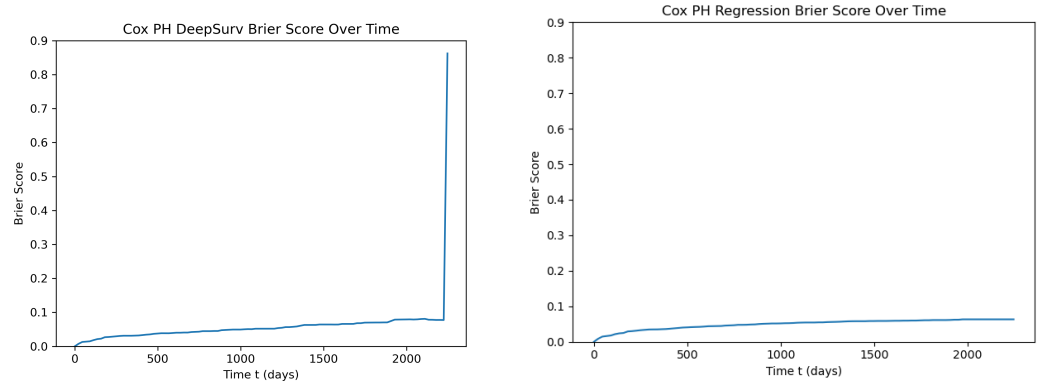


Figure 6: Time to progression survival prediction using DeepSurv stratified by VLDL lab test ordered (2091-7) and abnormal liver diagnosis (R94.5).



(a) Cox PH DeepSurv model.

(b) Cox PH regression model.

Figure 7: Brier score plotted over time for two Cox PH models predicting on a test dataset.

5.4 Baseline Models

For logistic regression, tables 2 and 3 below describe AUROC, AUPRC, and other metrics from the classification report:

Table 2: Logistic Regression Metrics

	AUROC	**AUPRC**
Train	0.918	0.840
Test	0.534	0.0573

Table 3: Classification Report for Logistic Regression Model on Training and Testing Sets

Set	**Label**	**Precision**	**Recall**	**F1-Score**
Train	No Progression	0.99	1.00	1.00
	Progression	1.00	0.84	0.91
	Accuracy			0.99
	Macro Avg	0.99	0.92	0.95
	Weighted Avg	0.99	0.99	0.99
Test	No Progression	0.96	0.98	0.97
	Progression	0.18	0.09	0.12
	Accuracy			0.94
	Macro Avg	0.57	0.53	0.54
	Weighted Avg	0.92	0.94	0.93

Table 4 below summarizes metrics for the linear regression model:

Table 4: Linear Regression Metrics

	Mean Squared Error (days)	**Mean Absolute Error (days)**
Train	136,825.480	292.486
Test	385,958.424	487.322

Lastly, the Cox PH regression model performance is summarized in table 5 below and in figure 7b:

Table 5: Cox PH Regression Metrics

	Concordance Index	**Integrated Brier Score**
Train	0.852	0.0314
Test	0.594	0.0496

5.5 Interpretability

The SHAP analysis and regression analysis found conflicting influential features listed in the subsections below. For the traditional machine learning methods, all the features are listed in descending order of coefficient magnitude. The corresponding human-readable descriptions were derived from the original data files.

5.5.1 Binary Progression Classification

For logistic regression, the top three features most negatively correlated with time-to-event are: family history of diabetes mellitus, myalgia, and osteoarthritis of the knee. The top three features most

positively correlated with time-to-event are use of meperidine, use of insulin, and unspecified allergy. SHAP scores were not included due to the poor performance of this model; we concluded that even if we discovered relatively influential features, they would not be good predictors of progression.

5.5.2 Time-to-Event

Based on SHAP scores, the most influential features in the time-to-event neural network are shown in Figure 8. Each dot represents an individual patient and the color indicates the feature value.

Reticulocyte Blood Test, Prealbumin Blood Test, and Vitamin B12 Blood Test are among the most important predictors. These blood tests suggest relationships with liver health. The spread in the negative portion of the SHAP values indicate that low reticulocyte count tends to push the model output to a shorter prediction of time-to-event, which is inconsistent with research where high reticulocyte count has been identified as a risk factor for the onset of MAFLD [4]. Our SHAP scores suggest that deficiencies in these markers may be associated with poorer prognosis.

Features like Cholesterol Test indicate a consistent directional effect — the higher the cholesterol value for patients, the more it tends to push the model output to a shorter prediction of time-to-event. The fact that cholesterol, LDL, and triglyceride lab tests are present in top 10 influential features in the model indicate that the model's learned feature relationships correlate with clinical knowledge of the relationship between obesity, diabetes, and MAFLD.

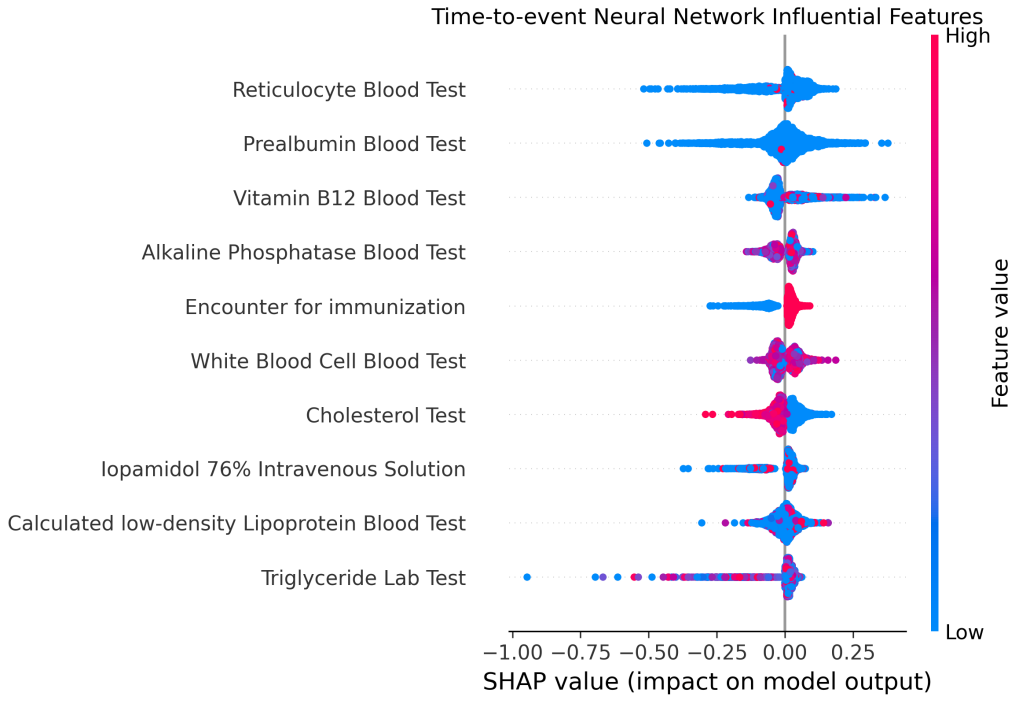


Figure 8: Beeswarm plots indicating the most influential features and their SHAP values on time-to-event neural network outcome.

For linear regression, usage of Leucovorin, Azithtomycin, and Pantaprazole were all positively correlated with time to progression. Additionally, mean corpuscular hemoglobin concentration (MCHC), kit for preparation of Tc-99m-Medronate Sodium, and Vancomycin all correlated positively with time to progression. Sheng et al. (2022) has shown that high MCHC levels is associated with worse prognosis is severe forms of liver cirrhosis, indicating a potential learned relationship between an outstanding feature and longer time-to-event prognosis [5]. A hypothesis for why this feature appears as a positively impacting feature may be due to a cohort-specific pattern or other confounding effect.

5.5.3 Survival Analysis

The SHAP scores for DeepSurv on this dataset are displayed in figure 9 below:

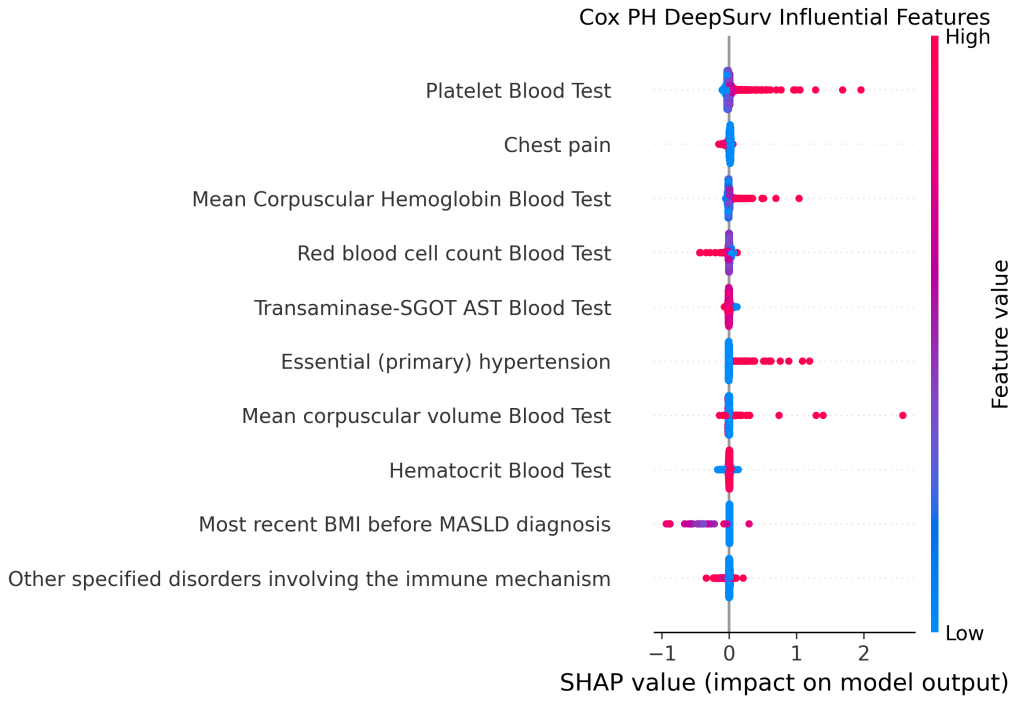


Figure 9: Beeswarm plots indicating the most influential features and their SHAP values on Cox PH DeepSurv outcome.

In the DeepSurv survival model, the most influential features include MCHC blood test, transaminase-SGOT AST blood test, and most recent BMI before MAFLD diagnosis. A higher MCHC value was associated with a longer time-to-event prediction, which matches the aforementioned relationship in the linear regression model as well. This may also be an artifact of Simpson's paradox, particularly due to the fact that we did not explore further stratification of our patient population for model training. The decision for a clinician to order the transaminase-SGOT AST may be indicative of potential liver health inquiry; although the small variance of the values does not necessarily indicate strong impact on the model output, the fact that it has high influence on the model is encouraging. Finally, the relationship between high values of the last BMI measured before the MAFLD diagnosis and a smaller model output correlates with the expectation of obesity potentially serving as a risk factor for MAFLD [1].

For Cox PH regression, the most negatively correlated features to outcome are the alanine aminotransferase (ALT) blood test, Azithromycin prescription, and Carvedilol prescription. However, Vitamin D3 capsule, Tizanidine, and kit preparation of Tc 99m-sestamibi IV solution being prescribed all correlated positively with time-to-event. Again, the presence of the ALT blood test being ordered by a clinician to monitor a patient's liver health is indicative itself, which suggests that the model has learned meaningful feature-to-outcome relationships.

6 Discussion

In general, we showed that even vanilla deep learning methods tend to perform at least as well as traditional machine learning regression methods. This is likely due to nonlinear activation layers in the neural networks that allow more complicated modelling of nonlinear relationships between the features and the outcome. Although the regression models and corresponding deep learning models

have some overlap in influential predictors, the deep learning models also provide more reasonable results in terms of influential features than traditional machine learning models do.

Our models revealed several clinically meaningful predictors of MAFLD progression. SHAP analysis of the time-to-event neural network highlighted features such as reticulocyte count, cholesterol test, and LDL test as positively associated with progression risk. These results support clinical intuition and reinforce known biomarkers. Similarly, linear and logistic models also ranked ALT/AST labs and medication codes among the top predictive features, suggesting that multiple modeling approaches have identified outstanding relationships between liver-related biomarkers and time-to-event outcome.

Despite these promising signals, the binary model exhibited a low test precision even when recall improved under class reweighting. The best PyCox survival model achieved a concordance score of 0.550 and an integrated Brier score of 0.055, showing reasonable discrimination and calibration. For regression-based progression timing, the time-to-event neural net achieved a mean absolute error of 375.85 days and a mean squared error of 253,614.92 days^2.

Limitations. First, we exclusively used structured data from the RPDR and did not incorporate unstructured notes (e.g., discharge summaries, imaging reports), which may contain high-yield predictive signals. Second, our dataset suffers from class imbalance: only a small fraction of patients progressed to advanced liver disease. While we tested SMOTE, downsampling, and loss reweighting, none of these methods substantially improved generalization. This highlights the lack of a robust strategy for addressing sample imbalance in real-world medical data.

Additionally, our models were trained only on adult cohorts and lack pediatric generalization. External validation on other EHR systems and demographic subgroups remains a key next step.

Finally, due to restrictions on the MGB data, we could not access clinical notes or event dates. This limited our ability to subset based on ICD-10 code adoption dates or verify timing of events.

Future Directions. We recommend continued tuning of hyperparameters and model architecture for the deep learning models. In addition, the dataset contains a wealth of unstructured clinical text, which can be leveraged through large language models (LLMs) to augment signal extraction. This text can be used to validate ICD-10 billing codes as well to produce data that more accurately reflects the patients' clinical state. Another important avenue of exploration is more methods of addressing class imbalance, such as generating synthetic data.

This study focused on adult patients age 30 and above, but little investigation has been made into MAFLD progression among pediatric cohorts. Future works can assess age-dependent differences in progression. Moreover, we recommend broadening outcomes beyond liver-specific endpoints to investigate systemic disease burden and mortality.

7 Code Availability

The code used for analysis is available at `https://github.com/jeannieshe/mlhc-masld`.

8 Acknowledgments & Member Contributions

We thank Dr. Amy Tsurumi for clinical mentorship. Contributions:

- Jeannie She: Jeannie cleaned physical measurement data, which includes calculating mean BMI and most recent BMI before first MAFLD diagnosis (both continuous and categorical, although the categorical BMI was not used in this model). Additionally, Jeannie setup the neural network model architecture for all of the aforementioned models, which include the time-to-event and binary neural networks and the DeepSurv survival net model. Finally, her work included calculating and interpreting the SHAP scores for the deep learning models.
- Erica Song: Erica cleaned lab test and medication data. For lab test data, Erica chose to use the mean value calculated over all entries for a patient to represent that lab for that patient; for medication, Erica created one-hot encoded vectors for each patient where the entry was 1 if the patient had received that prescription during any hospital visit. Erica ran linear,

logistic, and Lasso regression baselines to serve as comparisons against our deep learning models. Erica experimented with downsampling the data and identified the features with the greatest magnitude coefficients in these models.

- Naiqi Zhang: Naiqi cleaned diagnosis and demographic data, which involved one-hot encoding diagnoses and sex of each patient. Additionally, Naiqi implemented SMOTE and reweighting for the binary classification neural net to address the class imbalance. Along with these methods for addressing class imbalance, Naiqi evaluated and interpreted their outcomes and offered avenues for further improvement.

Appendix

Table 6: Table of codes and conditions we used as exclusion criteria.

Code	Description
A06.4	Amebic liver abscess
B15	Acute hepatitis A
B15.0	Hepatitis A with hepatic coma
B15.9	Hepatitis A without hepatic coma
B16	Acute hepatitis B
B16.0	Acute hepatitis B with delta-agent with hepatic coma
B16.0	Acute hepatitis B with delta-agent without hepatic coma
B16.2	Acute hepatitis B without delta-agent with hepatic coma
B16.9	Acute hepatitis B without delta-agent and without hepatic coma
B17	Other acute viral hepatitis
B17.0	Acute delta-(super) infection of hepatitis B carrier
B17.1	Acute hepatitis C
B17.10	Acute hepatitis C without hepatic coma
B17.11	Acute hepatitis C with hepatic coma
B17.2	Acute hepatitis E
B17.8	Other specified acute viral hepatitis
B17.9	Acute viral hepatitis, unspecified
B18	Chronic viral hepatitis
B18.0	Chronic viral hepatitis B with delta-agent
B18.1	Chronic viral hepatitis B without delta-agent
B18.2	Chronic viral hepatitis C
B18.8	Other chronic viral hepatitis
B18.9	Chronic viral hepatitis, unspecified
B19	Unspecified viral hepatitis
B19.0	Unspecified viral hepatitis with hepatic coma
B19.1	Unspecified viral hepatitis B
B19.10	Unspecified viral hepatitis B without hepatic coma
B19.11	Unspecified viral hepatitis B with hepatic coma
B19.2	Unspecified viral hepatitis C
B19.20	Unspecified viral hepatitis C without hepatic coma
B19.21	Unspecified viral hepatitis C with hepatic coma
B19.9	Unspecified viral hepatitis without hepatic coma
B25.1	Cytomegaloviral hepatitis
B67.0	Echinococcus granulosus infection of liver
B67.5	Echinococcus multilocularis infection of liver
B67.8	Echinococcosis, unspecified, of liver
B94.2	Sequelae of viral hepatitis
E24.4	Alcohol-induced pseudo-Cushing's syndrome
E83.01	Wilson's disease
E83.1	Disorders of iron metabolism
E83.10	Disorder of iron metabolism, unspecified
E83.11	Hemochromatosis
E83.110	Hereditary hemochromatosis
E83.111	Hemochromatosis due to repeated red blood cell transfusions
E83.118	Other hemochromatosis
E83.119	Hemochromatosis, unspecified
E83.19	Other disorders of iron metabolism
E88	Other and unspecified metabolic disorders
E88.0	Disorders of plasma-protein metabolism, not elsewhere classified
E88.01	Alpha-1-antitrypsin deficiency
E88.02	Plasminogen deficiency
E88.09	Other disorders of plasma-protein metabolism, not elsewhere classified
E88.0A	Disorders of plasma-protein metabolism, not elsewhere classified

Code	Description
E88.0B	Disorders of plasma-protein metabolism, not elsewhere classified
F10	Alcohol related disorders
F10.1	Alcohol abuse
F10.10	Alcohol abuse, uncomplicated
F10.11	Alcohol abuse, in remission
F10.12	Alcohol abuse with intoxication
F10.120	Alcohol abuse with intoxication, uncomplicated
F10.121	Alcohol abuse with intoxication delirium
F10.129	Alcohol abuse with intoxication, unspecified
F10.13	Alcohol abuse, with withdrawal
F10.130	Alcohol abuse with withdrawal, uncomplicated
F10.131	Alcohol abuse with withdrawal delirium
F10.132	Alcohol abuse with withdrawal with perceptual disturbance
F10.139	Alcohol abuse with withdrawal, unspecified
F10.14	Alcohol abuse with alcohol
F10.15	Alcohol abuse with alcohol
F10.150	Alcohol abuse with alcohol
F10.151	Alcohol abuse with alcohol
F10.159	Alcohol abuse with alcohol
F10.18	Alcohol abuse with other alcohol
F10.180	Alcohol abuse with alcohol
F10.181	Alcohol abuse with alcohol
F10.182	Alcohol abuse with alcohol
F10.188	Alcohol abuse with other alcohol
F10.19	Alcohol abuse with unspecified alcohol
F10.2	Alcohol dependence
F10.20	Alcohol dependence, uncomplicated
F10.21	Alcohol dependence, in remission
F10.22	Alcohol dependence with intoxication
F10.220	Alcohol dependence with intoxication, uncomplicated
F10.221	Alcohol dependence with intoxication delirium
F10.229	Alcohol dependence with intoxication, unspecified
F10.23	Alcohol dependence with withdrawal
F10.230	Alcohol dependence with withdrawal, uncomplicated
F10.231	Alcohol dependence with withdrawal delirium
F10.232	Alcohol dependence with withdrawal with perceptual disturbance
F10.239	Alcohol dependence with withdrawal, unspecified
F10.24	Alcohol dependence with alcohol
F10.25	Alcohol dependence with alcohol
F10.250	Alcohol dependence with alcohol
F10.251	Alcohol dependence with alcohol
F10.259	Alcohol dependence with alcohol
F10.26	Alcohol dependence with alcohol
F10.27	Alcohol dependence with alcohol
F10.28	Alcohol dependence with other alcohol
F10.280	Alcohol dependence with alcohol
F10.281	Alcohol dependence with alcohol
F10.282	Alcohol dependence with alcohol
F10.288	Alcohol dependence with other alcohol
F10.29	Alcohol dependence with unspecified alcohol
F10.9	Alcohol use, unspecified
F10.90	Alcohol use, unspecified, uncomplicated
F10.91	Alcohol use, unspecified, in remission
F10.92	Alcohol use, unspecified with intoxication
F10.920	Alcohol use, unspecified with intoxication, uncomplicated
F10.921	Alcohol use, unspecified with intoxication delirium
F10.929	Alcohol use, unspecified with intoxication, unspecified

Code	Description
F10.93	Alcohol use, unspecified with withdrawal
F10.930	Alcohol use, unspecified with withdrawal, uncomplicated
F10.931	Alcohol use, unspecified with withdrawal delirium
F10.932	Alcohol use, unspecified with withdrawal with perceptual disturbance
F10.939	Alcohol use, unspecified with withdrawal, unspecified
F10.94	Alcohol use, unspecified with alcohol
F10.95	Alcohol use, unspecified with alcohol
F10.950	Alcohol use, unspecified with alcohol
F10.951	Alcohol use, unspecified with alcohol
F10.959	Alcohol use, unspecified with alcohol
F10.96	Alcohol use, unspecified with alcohol
F10.97	Alcohol use, unspecified with alcohol
F10.98	Alcohol use, unspecified with other alcohol
F10.980	Alcohol use, unspecified with alcohol
F10.981	Alcohol use, unspecified with alcohol
F10.982	Alcohol use, unspecified with alcohol
F10.988	Alcohol use, unspecified with other alcohol
F10.99	Alcohol use, unspecified with unspecified alcohol
F11	Opioid related disorders
F11.1	Opioid abuse
F11.10	Opioid abuse, uncomplicated
F11.12	Opioid abuse with intoxication
F11.120	Opioid abuse with intoxication, uncomplicated
F11.121	Opioid abuse with intoxication delirium
F11.122	Opioid abuse with intoxication with perceptual disturbance
F11.129	Opioid abuse with intoxication, unspecified
F11.14	Opioid abuse with opioid-induced mood disorder
F11.15	Opioid abuse with opioid-induced psychotic disorder
F11.150	Opioid abuse with opioid-induced psychotic disorder with delusions
F11.151	Opioid abuse with opioid-induced psychotic disorder with hallucinations
F11.159	Opioid abuse with opioid-induced psychotic disorder, unspecified
F11.18	Opioid abuse with other opioid-induced disorder
F11.181	Opioid abuse with opioid-induced sexual dysfunction
F11.182	Opioid abuse with opioid-induced sleep disorder
F11.188	Opioid abuse with other opioid-induced disorder
F11.19	Opioid abuse with unspecified opioid-induced disorder
F11.2	Opioid dependence
F11.20	Opioid dependence, uncomplicated
F11.21	Opioid dependence, in remission
F11.22	Opioid dependence with intoxication
F11.220	Opioid dependence with intoxication, uncomplicated
F11.221	Opioid dependence with intoxication delirium
F11.222	Opioid dependence with intoxication with perceptual disturbance
F11.229	Opioid dependence with intoxication, unspecified
F11.23	Opioid dependence with withdrawal
F11.24	Opioid dependence with opioid-induced mood disorder
F11.25	Opioid dependence with opioid-induced psychotic disorder
F11.250	Opioid dependence with opioid-induced psychotic disorder with delusions
F11.251	Opioid dependence with opioid-induced psychotic disorder with hallucinations
F11.259	Opioid dependence with opioid-induced psychotic disorder, unspecified
F11.28	Opioid dependence with other opioid-induced disorder
F11.281	Opioid dependence with opioid-induced sexual dysfunction
F11.282	Opioid dependence with opioid-induced sleep disorder
F11.288	Opioid dependence with other opioid-induced disorder
F11.29	Opioid dependence with unspecified opioid-induced disorder
F11.9	Opioid use, unspecified
F11.90	Opioid use, unspecified, uncomplicated

Code	Description
F11.92	Opioid use, unspecified with intoxication
F11.920	Opioid use, unspecified with intoxication, uncomplicated
F11.921	Opioid use, unspecified with intoxication delirium
F11.922	Opioid use, unspecified with intoxication with perceptual disturbance
F11.929	Opioid use, unspecified with intoxication, unspecified
F11.93	Opioid use, unspecified with withdrawal
F11.94	Opioid use, unspecified with opioid-induced mood disorder
F11.95	Opioid use, unspecified with opioid-induced psychotic disorder
F11.950	Opioid use, unspecified with opioid-induced psychotic disorder with delusions
F11.951	Opioid use, unspecified with opioid-induced psychotic disorder with hallucinations
F11.959	Opioid use, unspecified with opioid-induced psychotic disorder, unspecified
F11.98	Opioid use, unspecified with other specified opioid-induced disorder
F11.981	Opioid use, unspecified with opioid-induced sexual dysfunction
F11.982	Opioid use, unspecified with opioid-induced sleep disorder
F11.988	Opioid use, unspecified with other opioid-induced disorder
F11.99	Opioid use, unspecified with unspecified opioid-induced disorder
F12	Cannabis related disorders
F12.1	Cannabis abuse
F12.10	Cannabis abuse, uncomplicated
F12.11	Cannabis abuse, in remission
F12.12	Cannabis abuse with intoxication
F12.120	Cannabis abuse with intoxication, uncomplicated
F12.121	Cannabis abuse with intoxication delirium
F12.122	Cannabis abuse with intoxication with perceptual disturbance
F12.129	Cannabis abuse with intoxication, unspecified
F12.13	Cannabis abuse with withdrawal
F12.15	Cannabis abuse with psychotic disorder
F12.150	Cannabis abuse with psychotic disorder with delusions
F12.151	Cannabis abuse with psychotic disorder with hallucinations
F12.159	Cannabis abuse with psychotic disorder, unspecified
F12.18	Cannabis abuse with other cannabis
F12.180	Cannabis abuse with cannabis
F12.188	Cannabis abuse with other cannabis
F12.19	Cannabis abuse with unspecified cannabis
F12.2	Cannabis dependence
F12.20	Cannabis dependence, uncomplicated
F12.21	Cannabis dependence, in remission
F12.22	Cannabis dependence with intoxication
F12.220	Cannabis dependence with intoxication, uncomplicated
F12.221	Cannabis dependence with intoxication delirium
F12.222	Cannabis dependence with intoxication with perceptual disturbance
F12.229	Cannabis dependence with intoxication, unspecified
F12.25	Cannabis dependence with psychotic disorder
F12.250	Cannabis dependence with psychotic disorder with delusions
F12.251	Cannabis dependence with psychotic disorder with hallucinations
F12.259	Cannabis dependence with psychotic disorder, unspecified
F12.28	Cannabis dependence with other cannabis
F12.280	Cannabis dependence with cannabis
F12.288	Cannabis dependence with other cannabis
F12.29	Cannabis dependence with unspecified cannabis
F12.9	Cannabis use, unspecified
F12.90	Cannabis use, unspecified, uncomplicated
F12.92	Cannabis use, unspecified with intoxication
F12.921	Cannabis use, unspecified with intoxication delirium
F12.95	Cannabis use, unspecified with psychotic disorder
F12.98	Cannabis use, unspecified with other cannabis
F12.99	Cannabis use, unspecified with unspecified cannabis

Code	Description
F13	Sedative, hypnotic, or anxiolytic-related disorders
F13.1	Sedative, hypnotic or anxiolytic-related abuse
F13.10	Sedative, hypnotic or anxiolytic abuse, uncomplicated
F13.11	Sedative, hypnotic or anxiolytic abuse, in remission
F13.12	Sedative, hypnotic or anxiolytic abuse with intoxication
F13.120	Sedative, hypnotic or anxiolytic abuse with intoxication, uncomplicated
F13.121	Sedative, hypnotic or anxiolytic abuse with intoxication delirium
F13.129	Sedative, hypnotic or anxiolytic abuse with intoxication, unspecified
F13.13	Sedative, hypnotic or anxiolytic abuse with withdrawal
F13.130	Sedative, hypnotic or anxiolytic abuse with withdrawal, uncomplicated
F13.131	Sedative, hypnotic or anxiolytic abuse with withdrawal delirium
F13.132	Sedative, hypnotic or anxiolytic abuse with withdrawal with perceptual disturbance
F13.139	Sedative, hypnotic or anxiolytic abuse with withdrawal, unspecified
F13.14	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced mood disorder
F13.15	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced psychotic disorder
F13.150	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced psychotic disorder with delusions
F13.151	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced psychotic disorder with hallucination
F13.159	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced psychotic disorder, unspecified
F13.18	Sedative, hypnotic or anxiolytic abuse with other sedative, hypnotic or anxiolytic-induced disorders
F13.180	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced anxiety disorder
F13.181	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced sexual dysfunction
F13.182	Sedative, hypnotic or anxiolytic abuse with sedative, hypnotic or anxiolytic-induced sleep disorder
F13.188	Sedative, hypnotic or anxiolytic abuse with other sedative, hypnotic or anxiolytic-induced disorder
F13.19	Sedative, hypnotic or anxiolytic abuse with unspecified sedative, hypnotic or anxiolytic-induced disorder
F13.2	Sedative, hypnotic or anxiolytic-related dependance
F13.20	Sedative, hypnotic or anxiolytic dependence, uncomplicated
F13.21	Sedative, hypnotic or anxiolytic dependence, in remission
F13.22	Sedative, hypnotic or anxiolytic dependence with intoxication
F13.220	Sedative, hypnotic or anxiolytic dependence with intoxication, uncomplicated
F13.221	Sedative, hypnotic or anxiolytic dependence with intoxication delirium
F13.229	Sedative, hypnotic or anxiolytic dependence with intoxication, unspecified
F13.23	Sedative, hypnotic or anxiolytic dependence with withdrawal
F13.230	Sedative, hypnotic or anxiolytic dependence with withdrawal, uncomplicated
F13.231	Sedative, hypnotic or anxiolytic dependence with withdrawal delirium
F13.232	Sedative, hypnotic or anxiolytic dependence with withdrawal with perceptual disturbance
F13.239	Sedative, hypnotic or anxiolytic dependence with withdrawal, unspecified
F13.24	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced mood disorder
F13.25	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced psychotic disorder
F13.250	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced psychotic disorder with delusi
F13.251	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced psychotic disorder with halluc
F13.259	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced psychotic disorder, unspecified

Code	Description
F13.26	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced persisting amnestic disorder
F13.27	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced persisting dementia
F13.28	Sedative, hypnotic or anxiolytic dependence with other sedative, hypnotic or anxiolytic-induced disorders
F13.280	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced anxiety disorder
F13.281	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced sexual dysfunction
F13.282	Sedative, hypnotic or anxiolytic dependence with sedative, hypnotic or anxiolytic-induced sleep disorder
F13.288	Sedative, hypnotic or anxiolytic dependence with other sedative, hypnotic or anxiolytic-induced disorder
F13.29	Sedative, hypnotic or anxiolytic dependence with unspecified sedative, hypnotic or anxiolytic-induced disorder
F13.9	Sedative, hypnotic or anxiolytic-related use, unspecified
F13.90	Sedative, hypnotic, or anxiolytic use, unspecified, uncomplicated
F13.91	Sedative, hypnotic or anxiolytic use, unspecified, in remission
F13.92	Sedative, hypnotic or anxiolytic use, unspecified with intoxication
F13.920	Sedative, hypnotic or anxiolytic use, unspecified with intoxication, uncomplicated
F13.921	Sedative, hypnotic or anxiolytic use, unspecified with intoxication delirium
F13.929	Sedative, hypnotic or anxiolytic use, unspecified with intoxication, unspecified
F13.93	Sedative, hypnotic or anxiolytic use, unspecified with withdrawal
F13.930	Sedative, hypnotic or anxiolytic use, unspecified with withdrawal, uncomplicated
F13.931	Sedative, hypnotic or anxiolytic use, unspecified with withdrawal delirium
F13.932	Sedative, hypnotic or anxiolytic use, unspecified with withdrawal with perceptual disturbances
F13.939	Sedative, hypnotic or anxiolytic use, unspecified with withdrawal, unspecified
F13.94	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced mood disorder
F13.95	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced psychotic disorder
F13.950	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced psychotic disorder with halluci
F13.951	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced psychotic disorder with halluci
F13.959	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced psychotic disorder, unspecified
F13.96	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced persisting amnestic disorder
F13.97	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced persisting dementia
F13.98	Sedative, hypnotic or anxiolytic use, unspecified with other sedative, hypnotic or anxiolytic-induced disorders
F13.980	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced anxiety disorder
F13.981	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced sexual dysfunction
F13.982	Sedative, hypnotic or anxiolytic use, unspecified with sedative, hypnotic or anxiolytic-induced sleep disorder
F13.988	Sedative, hypnotic or anxiolytic use, unspecified with other sedative, hypnotic or anxiolytic-induced disorder
F13.99	Sedative, hypnotic or anxiolytic use, unspecified with unspecified sedative, hypnotic or anxiolytic-induced disorder
F14	Cocaine related disorders
F14.1	Cocaine abuse

Code	Description
F14.10	Cocaine abuse, uncomplicated
F14.11	Cocaine abuse, in remission
F14.12	Cocaine abuse with intoxication
F14.120	Cocaine abuse with intoxication, uncomplicated
F14.121	Cocaine abuse with intoxication with delirium
F14.122	Cocaine abuse with intoxication with perceptual disturbance
F14.129	Cocaine abuse with intoxication, unspecified
F14.13	Cocaine abuse, unspecified with withdrawal
F14.14	Cocaine abuse with cocaine-induced mood disorder
F14.15	Cocaine abuse with cocaine-induced psychotic disorder
F14.150	Cocaine abuse with cocaine-induced psychotic disorder with delusions
F14.151	Cocaine abuse with cocaine-induced psychotic disorder with hallucinations
F14.159	Cocaine abuse with cocaine-induced psychotic disorder, unspecified
F14.18	Cocaine abuse with other cocaine-induced disorder
F14.180	Cocaine abuse with cocaine-induced anxiety disorder
F14.181	Cocaine abuse with cocaine-induced sexual dysfunction
F14.182	Cocaine abuse with cocaine-induced sleep disorder
F14.188	Cocaine abuse with other cocaine-induced disorder
F14.19	Cocaine abuse with unspecified cocaine-induced disorder
F14.2	Cocaine dependence
F14.20	Cocaine dependence, uncomplicated
F14.21	Cocaine dependence, in remission
F14.22	Cocaine dependence with intoxication
F14.220	Cocaine dependence with intoxication, uncomplicated
F14.221	Cocaine dependence with intoxication delirium
F14.222	Cocaine dependence with intoxication with perceptual disturbance
F14.229	Cocaine dependence with intoxication, unspecified
F14.23	Cocaine dependence with withdrawal
F14.24	Cocaine dependence with cocaine-induced mood disorder
F14.25	Cocaine dependence with cocaine-induced psychotic disorder
F14.250	Cocaine dependence with cocaine-induced psychotic disorder with delusions
F14.251	Cocaine dependence with cocaine-induced psychotic disorder with hallucinations
F14.259	Cocaine dependence with cocaine-induced psychotic disorder, unspecified
F14.28	Cocaine dependence with other cocaine-induced disorder
F14.280	Cocaine dependence with cocaine-induced anxiety disorder
F14.281	Cocaine dependence with cocaine-induced sexual dysfunction
F14.282	Cocaine dependence with cocaine-induced sleep disorder
F14.288	Cocaine dependence with other cocaine-induced disorder
F14.29	Cocaine dependence with unspecified cocaine-induced disorder
F14.9	Cocaine use, unspecified
F14.90	Cocaine use, unspecified, uncomplicated
F14.91	Cocaine use, unspecified, in remission
F14.92	Cocaine use, unspecified with intoxication
F14.920	Cocaine use, unspecified with intoxication, uncomplicated
F14.921	Cocaine use, unspecified with intoxication delirium
F14.922	Cocaine use, unspecified with intoxication with perceptual disturbance
F14.929	Cocaine use, unspecified with intoxication, unspecified
F14.93	Cocaine use, unspecified with withdrawal
F14.94	Cocaine use, unspecified with cocaine-induced mood disorder
F14.95	Cocaine use, unspecified with cocaine-induced psychotic disorder
F14.950	Cocaine use, unspecified with cocaine-induced psychotic disorder with delusions
F14.951	Cocaine use, unspecified with cocaine-induced psychotic disorder with hallucinations
F14.959	Cocaine use, unspecified with cocaine-induced psychotic disorder, unspecified
F14.98	Cocaine use, unspecified with other specified cocaine-induced disorder
F14.980	Cocaine use, unspecified with cocaine-induced anxiety disorder
F14.981	Cocaine use, unspecified with cocaine-induced sexual dysfunction
F14.982	Cocaine use, unspecified with cocaine-induced sleep disorder

Code	Description
F14.988	Cocaine use, unspecified with other cocaine-induced disorder
F14.99	Cocaine use, unspecified with unspecified cocaine-induced disorder
F16	Hallucinogen related disorders
F16.1	Hallucinogen abuse
F16.10	Hallucinogen abuse, uncomplicated
F16.11	Hallucinogen abuse, in remission
F16.12	Hallucinogen abuse with intoxication
F16.120	Hallucinogen abuse with intoxication, uncomplicated
F16.121	Hallucinogen abuse with intoxication with delirium
F16.122	Hallucinogen abuse with intoxication with perceptual disturbance
F16.129	Hallucinogen abuse with intoxication, unspecified
F16.14	Hallucinogen abuse with hallucinogen
F16.15	Hallucinogen abuse with hallucinogen
F16.150	Hallucinogen abuse with hallucinogen
F16.151	Hallucinogen abuse with hallucinogen
F16.159	Hallucinogen abuse with hallucinogen
F16.18	Hallucinogen abuse with other hallucinogen
F16.180	Hallucinogen abuse with hallucinogen
F16.183	Hallucinogen abuse with hallucinogen persisting perception disorder (flashbacks)
F16.188	Hallucinogen abuse with other hallucinogen
F16.19	Hallucinogen abuse with unspecified hallucinogen
F16.2	Hallucinogen dependence
F16.20	Hallucinogen dependence, uncomplicated
F16.21	Hallucinogen dependence, in remission
F16.22	Hallucinogen dependence with intoxication
F16.220	Hallucinogen dependence with intoxication, uncomplicated
F16.221	Hallucinogen dependence with intoxication with delirium
F16.229	Hallucinogen dependence with intoxication, unspecified
F16.24	Hallucinogen dependence with hallucinogen
F16.25	Hallucinogen dependence with hallucinogen
F16.250	Hallucinogen dependence with hallucinogen
F16.251	Hallucinogen dependence with hallucinogen
F16.259	Hallucinogen dependence with hallucinogen
F16.28	Hallucinogen dependence with other hallucinogen
F16.280	Hallucinogen dependence with hallucinogen
F16.283	Hallucinogen dependence with hallucinogen persisting perception disorder (flashbacks)
F16.288	Hallucinogen dependence with other hallucinogen
F16.29	Hallucinogen dependence with unspecified hallucinogen
F16.9	Hallucinogen use, unspecified
F16.90	Hallucinogen use, unspecified, uncomplicated
F16.91	Hallucinogen use, unspecified, in remission
F16.92	Hallucinogen use, unspecified with intoxication
F16.920	Hallucinogen use, unspecified with intoxication, uncomplicated
F16.921	Hallucinogen use, unspecified with intoxication with delirium
F16.929	Hallucinogen use, unspecified with intoxication, unspecified
F16.94	Hallucinogen use, unspecified with hallucinogen
F16.95	Hallucinogen use, unspecified with hallucinogen
F16.950	Hallucinogen use, unspecified with hallucinogen
F16.951	Hallucinogen use, unspecified with hallucinogen
F16.959	Hallucinogen use, unspecified with hallucinogen
F16.98	Hallucinogen use, unspecified with other specified hallucinogen
F16.980	Hallucinogen use, unspecified with hallucinogen
F16.983	Hallucinogen use, unspecified with hallucinogen persisting perception disorder (flashbacks)
F16.988	Hallucinogen use, unspecified with other hallucinogen
F16.99	Hallucinogen use, unspecified with unspecified hallucinogen

Code	Description
F18	Inhalant-related disorders
F18.1	Inhalant abuse
F18.10	Inhalant abuse, uncomplicated
F18.11	Inhalant abuse, in remission
F18.12	Inhalant abuse with intoxication
F18.120	Inhalant abuse with intoxication, uncomplicated
F18.121	Inhalant abuse with intoxication delirium
F18.129	Inhalant abuse with intoxication, unspecified
F18.14	Inhalant abuse with inhalant-induced mood disorder
F18.15	Inhalant abuse with inhalant-induced psychotic disorder
F18.150	Inhalant abuse with inhalant-induced psychotic disorder with delusions
F18.151	Inhalant abuse with inhalant-induced psychotic disorder with hallucinations
F18.159	Inhalant abuse with inhalant-induced psychotic disorder, unspecified
F18.17	Inhalant abuse with inhalant-induced dementia
F18.18	Inhalant abuse with other inhalant-induced disorders
F18.180	Inhalant abuse with inhalant-induced anxiety disorder
F18.188	Inhalant abuse with other inhalant-induced disorder
F18.19	Inhalant abuse with unspecified inhalant-induced disorder
F18.2	Inhalant dependence
F18.20	Inhalant dependence, uncomplicated
F18.21	Inhalant dependence, in remission
F18.22	Inhalant dependence with intoxication
F18.220	Inhalant dependence with intoxication, uncomplicated
F18.221	Inhalant dependence with intoxication delirium
F18.229	Inhalant dependence with intoxication, unspecified
F18.24	Inhalant dependence with inhalant-induced mood disorder
F18.25	Inhalant dependence with inhalant-induced psychotic disorder
F18.250	Inhalant dependence with inhalant-induced psychotic disorder with delusions
F18.251	Inhalant dependence with inhalant-induced psychotic disorder with hallucinations
F18.259	Inhalant dependence with inhalant-induced psychotic disorder, unspecified
F18.27	Inhalant dependence with inhalant-induced dementia
F18.28	Inhalant dependence with other inhalant-induced disorders
F18.280	Inhalant dependence with inhalant-induced anxiety disorder
F18.288	Inhalant dependence with other inhalant-induced disorder
F18.29	Inhalant dependence with unspecified inhalant-induced disorder
F18.9	Inhalant use, unspecified
F18.90	Inhalant use, unspecified, uncomplicated
F18.91	Inhalant use, unspecified, in remission
F18.92	Inhalant use, unspecified with intoxication
F18.920	Inhalant use, unspecified with intoxication, uncomplicated
F18.921	Inhalant use, unspecified with intoxication with delirium
F18.929	Inhalant use, unspecified with intoxication, unspecified
F18.94	Inhalant use, unspecified with inhalant-induced mood disorder
F18.95	Inhalant use, unspecified with inhalant-induced psychotic disorder
F18.950	Inhalant use, unspecified with inhalant-induced psychotic disorder with delusions
F18.951	Inhalant use, unspecified with inhalant-induced psychotic disorder with hallucinations
F18.959	Inhalant use, unspecified with inhalant-induced psychotic disorder, unspecified
F18.97	Inhalant use, unspecified with inhalant-induced persisting dementia
F18.98	Inhalant use, unspecified with other inhalant-induced disorders
F18.980	Inhalant use, unspecified with inhalant-induced anxiety disorder
F18.988	Inhalant use, unspecified with other inhalant-induced disorder
F18.99	Inhalant use, unspecified with unspecified inhalant-induced disorder
F19	Other psychoactive substance related disorders
F19.1	Other psychoactive substance abuse
F19.10	Other psychoactive substance abuse, uncomplicated
F19.12	Other psychoactive substance abuse with intoxication
F19.120	Other psychoactive substance abuse with intoxication, uncomplicated

Code	Description
F19.121	Other psychoactive substance abuse with intoxication delirium
F19.122	Other psychoactive substance abuse with intoxication with perceptual disturbances
F19.129	Other psychoactive substance abuse with intoxication, unspecified
F19.14	Other psychoactive substance abuse with psychoactive substance-induced mood disorder
F19.15	Other psychoactive substance abuse with psychoactive substance-induced psychotic disorder
F19.150	Other psychoactive substance abuse with psychoactive substance-induced psychotic disorder with delusions
F19.151	Other psychoactive substance abuse with psychoactive substance-induced psychotic disorder with hallucinations
F19.159	Other psychoactive substance abuse with psychoactive substance-induced psychotic disorder, unspecified
F19.16	Other psychoactive substance abuse with psychoactive substance-induced persisting amnestic disorder
F19.17	Other psychoactive substance abuse with psychoactive substance-induced persisting dementia
F19.18	Other psychoactive substance abuse with other psychoactive substance-induced disorders
F19.180	Other psychoactive substance abuse with psychoactive substance-induced anxiety disorder
F19.181	Other psychoactive substance abuse with psychoactive substance-induced sexual dysfunction
F19.182	Other psychoactive substance abuse with psychoactive substance-induced sleep disorder
F19.188	Other psychoactive substance abuse with other psychoactive substance-induced disorder
F19.19	Other psychoactive substance abuse with unspecified psychoactive substance-induced disorder
F19.2	Other psychoactive substance dependence
F19.20	Other psychoactive substance dependence, uncomplicated
F19.21	Other psychoactive substance dependence, in remission
F19.22	Other psychoactive substance dependence with intoxication
F19.220	Other psychoactive substance dependence with intoxication, uncomplicated
F19.221	Other psychoactive substance dependence with intoxication delirium
F19.222	Other psychoactive substance dependence with intoxication with perceptual disturbance
F19.229	Other psychoactive substance dependence with intoxication, unspecified
F19.23	Other psychoactive substance dependence with withdrawal
F19.230	Other psychoactive substance dependence with withdrawal, uncomplicated
F19.231	Other psychoactive substance dependence with withdrawal delirium
F19.232	Other psychoactive substance dependence with withdrawal with perceptual disturbance
F19.239	Other psychoactive substance dependence with withdrawal, unspecified
F19.24	Other psychoactive substance dependence with psychoactive substance-induced mood disorder
F19.25	Other psychoactive substance dependence with psychoactive substance-induced psychotic disorder
F19.250	Other psychoactive substance dependence with psychoactive substance-induced psychotic disorder with delusions
F19.251	Other psychoactive substance dependence with psychoactive substance-induced psychotic disorder with hallucinations
F19.259	Other psychoactive substance dependence with psychoactive substance-induced psychotic disorder, unspecified
F19.26	Other psychoactive substance dependence with psychoactive substance-induced persisting amnestic disorder
F19.27	Other psychoactive substance dependence with psychoactive substance-induced persisting dementia

Code	Description
F19.28	Other psychoactive substance dependence with other psychoactive substance-induced disorders
F19.280	Other psychoactive substance dependence with psychoactive substance-induced anxiety disorder
F19.281	Other psychoactive substance dependence with psychoactive substance-induced sexual dysfunction
F19.282	Other psychoactive substance dependence with psychoactive substance-induced sleep disorder
F19.288	Other psychoactive substance dependence with other psychoactive substance-induced disorder
F19.29	Other psychoactive substance dependence with unspecified psychoactive substance-induced disorder
F19.9	Other psychoactive substance use, unspecified
F19.90	Other psychoactive substance use, unspecified, uncomplicated
F19.92	Other psychoactive substance use, unspecified with intoxication
F19.920	Other psychoactive substance use, unspecified with intoxication, uncomplicated
F19.921	Other psychoactive substance use, unspecified with intoxication with delirium
F19.922	Other psychoactive substance use, unspecified with intoxication with perceptual disturbance
F19.929	Other psychoactive substance use, unspecified with intoxication, unspecified
F19.93	Other psychoactive substance use, unspecified with withdrawal
F19.930	Other psychoactive substance use, unspecified with withdrawal, uncomplicated
F19.931	Other psychoactive substance use, unspecified with withdrawal delirium
F19.932	Other psychoactive substance use, unspecified with withdrawal with perceptual disturbance
F19.939	Other psychoactive substance use, unspecified with withdrawal, unspecified
F19.94	Other psychoactive substance use, unspecified with psychoactive substance-induced mood disorder
F19.95	Other psychoactive substance use, unspecified with psychoactive substance-induced psychotic disorder
F19.950	Other psychoactive substance use, unspecified with psychoactive substance-induced psychotic disorder with delusions
F19.951	Other psychoactive substance use, unspecified with psychoactive substance-induced psychotic disorder with hallucinations
F19.959	Other psychoactive substance use, unspecified with psychoactive substance-induced psychotic disorder, unspecified
F19.96	Other psychoactive substance use, unspecified with psychoactive substance-induced persisting amnestic disorder
F19.97	Other psychoactive substance use, unspecified with psychoactive substance-induced persisting dementia
F19.98	Other psychoactive substance use, unspecified with other psychoactive substance-induced disorders
F19.980	Other psychoactive substance use, unspecified with psychoactive substance-induced anxiety disorder
F19.981	Other psychoactive substance use, unspecified with psychoactive substance-induced sexual dysfunction
F19.982	Other psychoactive substance use, unspecified with psychoactive substance-induced sleep disorder
F19.988	Other psychoactive substance use, unspecified with other psychoactive substance-induced disorder
F19.99	Other psychoactive substance use, unspecified with unspecified psychoactive substance-induced disorder
G31.2	Degeneration of nervous system due to alcohol
G62.1	Alcoholic polyneuropathy
G72.1	Alcoholic myopathy
I42.6	Alcoholic cardiomyopathy
I82.0	Budd Chiari syndrome

Code	Description
K29.2	Alcoholic gastritis
K29.20	Alcoholic gastritis without bleeding
K29.21	Alcoholic gastritis with bleeding
K70	Alcoholic liver disease
K70.0	Alcoholic fatty liver
K70.1	Alcoholic hepatitis
K70.10	Alcoholic hepatitis without ascites
K70.11	Alcoholic hepatitis with ascites
K70.2	Alcoholic fibrosis and sclerosis of liver
K70.3	Alcoholic cirrhosis of liver
K70.30	Alcoholic cirrhosis of liver without ascites
K70.31	Alcoholic cirrhosis of liver with ascites
K70.4	Alcoholic hepatic failure
K70.40	Alcoholic hepatic failure without coma
K70.41	Alcoholic hepatic failure with coma
K70.9	Alcoholic liver disease, unspecified
K71.0	Toxic liver disease with cholestasis
K71.10	Toxic liver disease with hepatic necrosis, without coma
K71.2	Toxic liver disease with acute hepatitis
K71.6	Toxic liver disease with hepatitis, not elsewhere classified
K71.7	Toxic liver disease with fibrosis and cirrhosis of liver
K71.8	Toxic liver disease with other disorders of liver
K71.9	Toxic liver disease, unspecified
K73.0	Chronic persistent hepatitis, not elsewhere classified
K73.1	Chronic lobular hepatitis, not elsewhere classified
K73.2	Chronic active hepatitis, not elsewhere classified
K73.8	Other chronic hepatitis, not elsewhere classified
K73.9	Chronic hepatitis, unspecified
K74.3	Primary biliary cirrhosis
K74.4	Secondary biliary cirrhosis
K74.5	Biliary cirrhosis, unspecified
K75.0	Abscess of liver
K75.1	Phlebitis of portal vein
K75.2	Nonspecific reactive hepatitis
K75.3	Granulomatous hepatitis, not elsewhere classified
K75.4	Autoimmune hepatitis
K75.9	Inflammatory liver disease, unspecified
K76.5	Hepatic veno-occlusive disease
K83.0	Cholangitis
K83.01	Primary sclerosing cholangitis
K83.09	Other cholangitis
K83.0A	Cholangitis
K83.0F	Cholangitis
K85.2	Alcohol induced acute pancreatitis
K85.20	Alcohol induced acute pancreatitis without necrosis or infection
K85.21	Alcohol induced acute pancreatitis with uninfected necrosis
K85.22	Alcohol induced acute pancreatitis with infected necrosis
K86.0	Alcohol-induced chronic pancreatitis
O35.4	Maternal care for (suspected) damage to fetus from alcohol
O98.4	Viral hepatitis complicating pregnancy, childbirth and the puerperium
O98.41	Viral hepatitis complicating pregnancy
O98.411	Viral hepatitis complicating pregnancy, first trimester
O98.412	Viral hepatitis complicating pregnancy, second trimester
O98.413	Viral hepatitis complicating pregnancy, third trimester
O98.419	Viral hepatitis complicating pregnancy, unspecified trimester
O98.42	Viral hepatitis complicating childbirth
O98.43	Viral hepatitis complicating the puerperium

Code	Description
O99.31	Alcohol use complicating pregnancy, childbirth, and the puerperium
O99.310	Alcohol use complicating pregnancy, unspecified trimester
O99.311	Alcohol use complicating pregnancy, first trimester
O99.312	Alcohol use complicating pregnancy, second trimester
O99.313	Alcohol use complicating pregnancy, third trimester
O99.314	Alcohol use complicating childbirth
P04.3	Newborn affected by maternal use of alcohol
P78.81	Congenital cirrhosis of liver
Q44.6	Cystic disease of liver
Q86.0	Fetal alcohol syndrome
R78	Findings of drugs and other substances, not normally found in blood
R78.0	Finding of alcohol in blood
R78.1	Finding of opiate drug in blood
R78.2	Finding of cocaine drug in blood
R78.3	Finding of hallucinogen in blood
R78.4	Finding of other drugs with addictive potential in blood
R78.5	Finding of other psychotropic drug in blood
T40	Poisoning by, adverse effect of and underdosing of narcotics and psychodysleptics hallucinogens
T40.4	Poisoning by, adverse effect of and underdosing of other synthetic narcotics
T40.49	Poisoning by, adverse effect of and underdosing of other synthetic narcotics
T40.491	Poisoning by other synthetic narcotics, accidental (unintentional)
T40.492	Poisoning by other synthetic narcotics, intentional self
T40.493	Poisoning by other synthetic narcotics, assault
T40.494	Poisoning by other synthetic narcotics, undetermined
T40.495	Adverse effect of other synthetic narcotics
T40.496	Underdosing of other synthetic narcotics
T40.5	Poisoning by, adverse effect of and underdosing of cocaine
T40.6	Poisoning by, adverse effect of and underdosing of other and unspecified narcotics
T40.60	Poisoning by, adverse effect of and underdosing of unspecified narcotics
T40.601	Poisoning by unspecified narcotics, accidental (unintentional)
T40.602	Poisoning by unspecified narcotics, intentional self
T40.603	Poisoning by unspecified narcotics, assault
T40.604	Poisoning by unspecified narcotics, undetermined
T40.605	Adverse effect of unspecified narcotics
T40.606	Underdosing of unspecified narcotics
T40.69	Poisoning by adverse effect of and underdosing of other narcotics
T40.691	Poisoning by other narcotics, accidental (unintentional)
T40.692	Poisoning by other narcotics, intentional self
T40.693	Poisoning by other narcotics, assault
T40.694	Poisoning by other narcotics, undetermined
T40.695	Adverse effect of other narcotics
T40.696	Underdosing of other narcotics
T40.7	Poisoning by, adverse effect of and underdosing of cannabis (derivatives)
T40.71	Poisoning by, adverse effect of and underdosing of cannabis (derivatives)
T40.711	Poisoning by cannabis, accidental (unintentional)
T40.712	Poisoning by cannabis, intentional self
T40.713	Poisoning by cannabis, assault
T40.714	Poisoning by cannabis, undetermined
T40.715	Adverse effect of cannabis
T40.716	Underdosing of cannabis
T50.7	Poisoning by, adverse effect of and underdosing of analeptics and opioid receptor antagonists
T51	Toxic effect of alcohol
T51.0	Toxic effect of ethanol
T51.1	Toxic effect of methanol
T51.2	Toxic effect of 2-Propanol

Code	Description
T51.8	Toxic effect of other alcohols
T51.9	Toxic effect of unspecified alcohol
T51.91	Toxic effect of unspecified alcohol, accidental (unintentional)
T51.92	Toxic effect of unspecified alcohol, intentional self-harm
T51.93	Toxic effect of unspecified alcohol, assault
T51.94	Toxic effect of unspecified alcohol, undetermined
X45	Accidental poisoning by and exposure to alcohol
X65	Intentional self-poisoning by and exposure to alcohol, undetermined intent
Y15	Poisoning by and exposure to alcohol, undetermined intent
Y57.3	intoxication with medicine against alcohol abuse
Y90	Evidence of alcohol involvement determined by blood alcohol level
Y90.0	Blood alcohol level of less than 20 mg/100 ml
Y90.1	Blood alcohol level of 20-39 mg/100 ml
Y90.2	Blood alcohol level of 40-59 mg/100 ml
Y90.3	Blood alcohol level of 60-79 mg/100 ml
Y90.4	Blood alcohol level of 80-99 mg/100 ml
Y90.5	Blood alcohol level of 100-119 mg/100 ml
Y90.6	Blood alcohol level of 120-199 mg/100 ml
Y90.7	Blood alcohol level of 200-239 mg/100 ml
Y90.8	Blood alcohol level of 240 mg/100 ml or more
Y91	Evidence of alcohol involvement determined by level of intoxication
Y91.0	Mild alcohol intoxication
Y91.1	Moderate alcohol intoxication
Y91.2	Severe alcohol intoxication
Y91.3	Very severe alcohol intoxication
Y91.9	Alcohol involvement, not otherwise specified
Z02.83	Encounter for blood-alcohol and blood-drug test
Z20.5	Contact with and (suspected) exposure to viral hepatitis
Z22.51	Carrier of viral hepatitis
Z22.52	Carrier of viral hepatitis C
Z50.2	Alcohol rehabilitation
Z71.4	advice and check-up for alcohol abuse
Z71.41	Alcohol abuse counseling and surveillance of alcoholic
Z72.1	lifestyle problems due to use of alcohol

Table 7: Top 10 Features Most Negatively and Positively Correlated with Progression Classification (Logistic Regression)

Rank	Negative Correlations	Positive Correlations
1	ICD-10 Z83.3: Family history of diabetes mellitus	EPIC-MED 123225: Meperidine (Pf) 100 mg/ml Injection Solution
2	ICD-10 M79.18: Myalgia, other site	EPIC-MED 119019: Insulin Lispro 100 Unit/ml Subcutaneous Pen
3	ICD-10 M17.9: Osteoarthritis of knee, unspecified	ICD-10 T78.40XA: Allergy, unspecified, initial encounter
4	ICD-10 R79.9: Abnormal finding of blood chemistry, unspecified	ICD-10 L70.9: Acne, unspecified
5	ICD-10 L72.3: Sebaceous cyst	ICD-9 260: Major Pancreas, Liver And Shunt Procedures (APR v30)
6	ICD-10 J01.90: Acute sinusitis, unspecified	ICD-10 D89.89: Other specified disorders involving the immune mechanism, not elsewhere classified
7	ICD-10 H90.5: Unspecified sensorineural hearing loss	EPIC-MED 2405: Diazepam 5 mg Tablet
8	EPIC-MED 4573: Lorazepam 1 mg Tablet	ICD-10 G47.10: Hypersomnia, unspecified
9	EPIC-MED 89422: Omeprazole 20 mg Tablet, Delayed Release	EPIC-PRC 63601006: Pneumococcal Polysac Vaccine 23-V 2 Yrs/¡subq/Im HC
10	ICD-10 E87.1: Hypo-osmolality and hyponatremia	ICD-10 R93.2: Abnormal findings on diagnostic imaging of liver and biliary tract

Table 8: Top 10 Features Most Negatively and Positively Correlated with Time-to-Progression (Linear Regression)

Rank	Negative Correlation	Positive Correlation
1	EPIC-MED 3004080138: Leucovorin IVPB in D5w 250 mL (250 mg Vial)	Lab 786-4: MCHC Blood
2	HCPCS J0456: Azithromycin 500mg inj	EPIC-MED 126282: Kit For Preparation of Tc-99m-Medronate Sodium 20 mg IV Solution
3	EPIC-MED 26226: Pantoprazole 40 mg Intravenous Solution	EPIC-MED 4004080184: Vancomycin IVPB 2000 mg (500 ml)
4	EPIC-MED 7312: Sodium Bicarbonate 650 mg Tablet	EPIC-MED 122799: Oxytocin 30 Unit/500 ml In 0.9% Sodium Chloride Intravenous Solution
5	EPIC-MED 700594: Maalox-Benadryl-Lidocaine (Miracle Mouthwash-Mbl) Suspension Cmpd	ICD-10 R40.2362: Coma scale, best motor response, obeys command
6	HCPCS J2590: Injection, oxytocin, up to 10 units	EPIC-MED 77195: Levetiracetam 500 mg/5 ml Intravenous Solution
7	ICD-9 660: Major Hematologic/Immunologic Diagnosis Exc Sickle Cell Crisis and Coagul (APR v30)	HCPCS J2248: Injection, micafungin sodium, 1 mg
8	HCPCS J1030: Injection, Methylprednisolone Acetate, 40 mg	EPIC-MED 4393: Leucovorin Calcium 350 mg Solution For Injection
9	EPIC-MED 701190: Norepinephrine Infusion Syringe In Swfi 80 mcg/ml Cmpd MGH	HCPCS C9113: Injection, pantoprazole sodium, per vial
10	CPT 91300: COVID-19 Vaccine (SARS-CoV-2, mRNA-LNP) intramuscular use	EPIC-MED 21063: Azithromycin 500 mg Intravenous Solution

Table 9: Top 10 Features Most Negatively and Positively Correlated with Duration and Event (Cox PH Regression)

Rank	Negative Correlations	Positive Correlations
1	LOINC 1742-6: ALT	EPIC-MED 15636: Cholecalciferol (Vitamin D3) 5,000 Unit Capsule
2	EPIC-MED 20943: Azithromycin 250 mg Tablet	EPIC-MED 14793: Tizanidine 4 mg Tablet
3	EPIC-MED 15747: Carvedilol 6.25 mg Tablet	EPIC-MED 126684: Kit preparation of Tc 99m-sestamibi Combo No.1 IV Solution
4	EPIC-MED 119654: Dextrose 50% In Water (D50w) Intravenous Syringe	EPIC-MED 19696: Benzocaine 20% Mucosal Aerosol Spray
5	EPIC-MED 19882: Sertraline 25 mg Tablet	EPIC-MED 11351: Sertraline 50 mg Tablet
6	EPIC-MED 10839: Peg 3350-Electrolytes 236 Gram-22.74 Gram-6.74 Gram-5.86 Gram Solution	EPIC-MED 169388: Tbo-Filgrastim 480 mcg/0.8 ml Subcutaneous Syringe
7	EPIC-MED 12652: Succinylcholine Chloride 100 mg/5 ml (20 mg/ml) Intravenous Syringe	EPIC-MED 106783: Fosaprepitant 150 mg Intravenous Solution
8	EPIC-MED 10323: Iohexol 350 mg Iodine/ml Intravenous Solution	EPIC-MED 120153: Insulin Nph Human Recombinant 100 Unit/ml (3 ml) Subcutaneous Pen
9	LOINC XC5-9: VLDL	EPIC-MED 105633: Bupivacaine-Epinephrine (Pf) 0.25%-1:200,000 Injection Solution
10	EPIC-MED 11442: Sucralfate 1 Gram Tablet	EPIC-MED 106364: Bupivacaine-Epinephrine (Pf) 0.5%-1:200,000 Injection Solution

References

[1] Metabolic Dysfunction-Associated Steatohepatitis: What It Is, Causes.

[2] Jared L. Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Medical Research Methodology*, 18(1):24, February 2018.

[3] Jonathan Li. Predicting Progression of Metabolic Dysfunction-associated Steatotic Liver Disease. Technical report, Massachusetts Institute of Technology, February 2024.

[4] Peng-Cheng Ma, Qi-Mei Li, Rui-Ning Li, Chang Hong, Hao Cui, Zi-Yong Zhang, Yan Li, Lu-Shan Xiao, Hong Zhu, Lin Zeng, Jun Xu, Wei-Nan Lai, and Li Liu. A high reticulocyte count is a risk factor for the onset of metabolic dysfunction-associated steatotic liver disease: Cross-sectional and prospective studies of data of 310,091 individuals from the UK Biobank. *Frontiers in Pharmacology*, 15, July 2024. Publisher: Frontiers.

[5] Xinyu Sheng, Wenyi Chen, Yanping Xu, Feiyan Lin, and Hongcui Cao. Higher Mean Corpuscular Hemoglobin Concentration is Associated with Worse Prognosis of Hepatorenal Syndrome: A Multicenter Retrospective Study. *The American Journal of the Medical Sciences*, 363(1):25–32, January 2022.

[6] Leslie N. Smith. Cyclical Learning Rates for Training Neural Networks, April 2017. arXiv:1506.01186 [cs].