Roy *et al. Metab Target Organ Damage.* 2025;5:8
DOI: 10.20517/mtod.2024.84

Metabolism and Target Organ Damage

Review

Open Access

Machine learning models and AI in predicting diagnosis and prognosis in alcohol-related and metabolic dysfunction-associated steatotic liver disease

Akash Roy[1], Nipun Verma[2]

[1]Institute of Gastrosciences and Liver Transplantation, Apollo Multi-speciality Hospitals, Kolkatta 700054, India.
[2]Department of Hepatology, Postgraduate Institute of Medical Education and Research, Chandigarh 160012, India.

Correspondence to: Dr. Nipun Verma, Department of Hepatology, Postgraduate Institute of Medical Education and Research, Sector 12, Chandigarh 160012, India. E-mail: nipun29j@gmail.com

How to cite this article: Roy A, Verma N. Machine learning models and AI in predicting diagnosis and prognosis in alcohol-related and metabolic dysfunction-associated steatotic liver disease. *Metab Target Organ Damage.* 2025;5:8. https://dx.doi.org/10.20517/mtod.2024.84

Received: 21 Sep 2024 **First Decision:** 24 Dec 2024 **Revised:** 8 Feb 2025 **Accepted:** 17 Feb 2025 **Published:** 25 Feb 2025

Academic Editor: Amedeo Lonardo **Copy Editor:** Ting-Ting Hu **Production Editor:** Ting-Ting Hu

Abstract

Steatotic liver disease (SLD) is the most common cause of liver disease globally, with an ever-increasing burden. The two primary components of SLD are metabolic dysfunction-associated steatotic liver disease (MASLD) and Alcohol-Associated Liver Disease (ALD). Both entities have important knowledge gaps in differentiation, diagnosis, risk stratification, and prognosis. Given the enormous burden of both MASLD and ALD and their diverse presentation, they form an ideal ground for the application of artificial intelligence (AI) and machine learning (ML) techniques and algorithms. ML models can aid in disease prediction among large populations and estimate those at the highest risk of disease progression or mortality, while applications with AI technology can aid in better detection and monitored treatment approaches. The use of AI in digital pathology and digital therapeutics are attractive options in moving toward personalized medicine. This review briefly summarizes the knowledge gaps in SLD with emerging literature on the use of ML and AI technologies across domains of disease detection, diagnosis, and prognosis.

Keywords: Steatotic liver disease, machine learning, artificial intelligence, prognosis, metabolic dysfunction, alcohol

INTRODUCTION

Steatotic liver disease (SLD) is the most common cause of liver disease globally, accounting for the majority of patients with cirrhosis and liver-related mortality[1]. The two broad overarching categories in SLD are metabolic dysfunction-associated steatotic liver disease (MASLD) and alcohol-associated liver disease (ALD), which, taken together, are the future drivers of liver disease[2]. While initially thought of as dichotomous entities, given the global increase in alcohol consumption, it is no longer prudent to consider them as binaries[2,3]. Instead, they should be looked at as a spectrum with metabolic dysfunction being the predominant driver on one end and significant alcohol on the other, with interim grey zones depending upon the increasing severity of alcohol intake [Figure 1][3,4]. While different based on nomenclature, ALD and MASLD share stark similarities. The basic pathology starts with increased intrahepatic lipid accumulation, which leads to downstream effects of inflammation and fibrosis[5]. There is growing evidence showing a complex interaction of independent, combined, and modifying effects of alcohol and metabolic factors on the onset and progression of SLD, highlighting the multifactorial background of liver disease. Given the emerging complexity, more holistic approaches could be useful for risk prediction, diagnosis treatment planning, and prognosis[2].

NEED FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE MODELS IN MASLD/ALD

Clinical decision making in hepatology relies on establishing an accurate diagnosis, understanding the dynamics of progression, ascertaining intrinsic prognostic risks, determining the population to benefit most from therapy, and identifying the most appropriate therapy. SLD is a slowly progressive disease, and both for MASLD and ALD, a majority do not go on to have progressive liver disease. Instead, non-liver-related events (cardiovascular, renal, and extrahepatic malignancies) frequently determine ultimate patient outcomes[6]. Large data across fourteen studies encompassing more than 17,000 individuals have shown that liver-related and all-cause mortality is increased, with those at the highest risk defined as having greater than stage 2 fibrosis[7]. Hence, it becomes imperative to diagnose, identify this group at the highest risk, and devise appropriate risk-based therapeutic strategies that translate to tangible benefits. Clinical risk prediction models, hence, find their application in both MASLD and ALD. Such models help clinicians make better decisions in both diagnostic and prognostic domains, including the development of cirrhosis, first decompensation, and incident organ failures[8]. While conventional modeling techniques have long been in vogue, clinical data's emerging complexity and magnitude have paved the scope for machine learning (ML) and artificial intelligence (AI)-based modeling in SLD, specifically for MASLD/ALD phenotypes.

BASICS OF ML AND AI

AI is the fastest-growing domain of computer science, with broad scopes and acceptance across disciplines ranging from education to finance, e-commerce, human resources, and rapid healthcare evolution[9]. In simplistic terms, AI enables computers and machines to simulate human intelligence and problem-solving capabilities. Amidst the broad domain of AI, ML is a specific subdomain that focuses on using data and algorithms to enable AI to imitate how humans learn, gradually improving its accuracy with the ultimate aim of performing complex tasks. ML itself can be "supervised ML" when priorly labeled datasets are used to train algorithms, classify data, or predict outcomes, or it can be "unsupervised ML" where algorithms are aimed to analyze unlabeled datasets (subsets called clusters)[10].

The AI/ML development lifecycle [Figure 2A] begins with the collection of reliable and valid data, followed by crucial steps including data cleaning, exploratory analyses, data partitioning, model training, validation, generalization, and ultimately, impact assessments. From the outset, clear expectations and deliverables

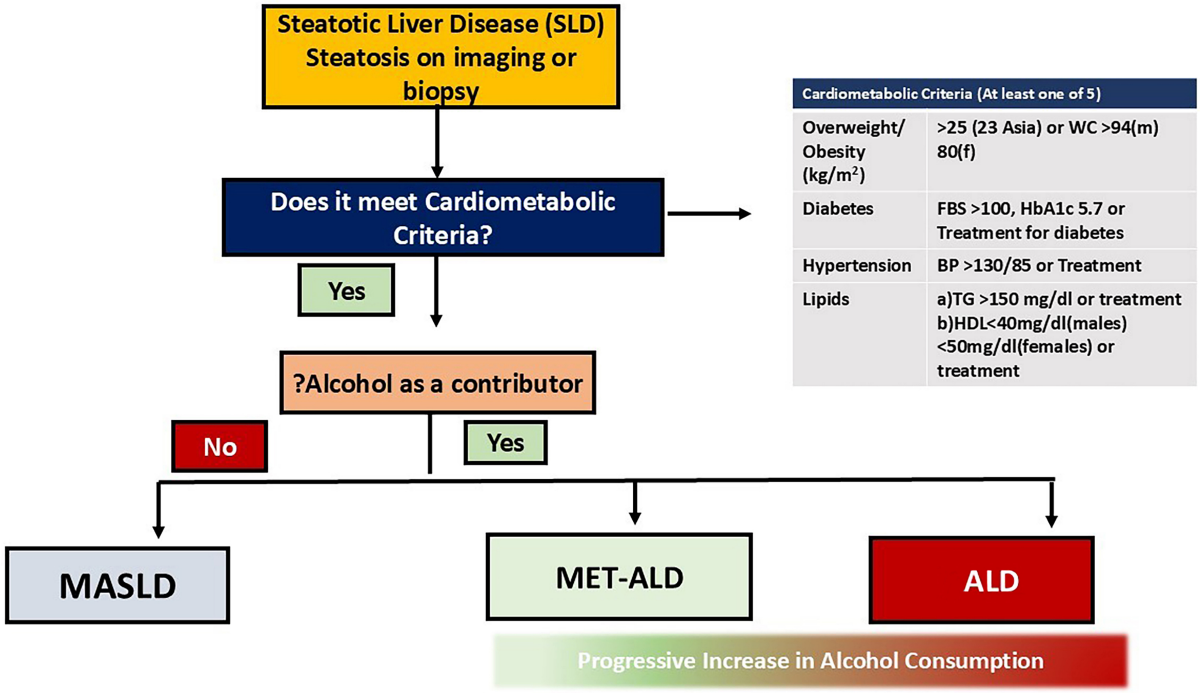

Figure 1. Depicting the current classification of steatotic liver disease into common phenotypes. FBS: Fasting blood sugar; HBA1C: glycosylated hemoglobin; MASLD: metabolic dysfunction-associated steatotic liver disease; ALD: alcohol-associated liver disease; WC: waist circumference; HDL: high-density lipoprotein; TG: triglyceride; BP: blood pressure.

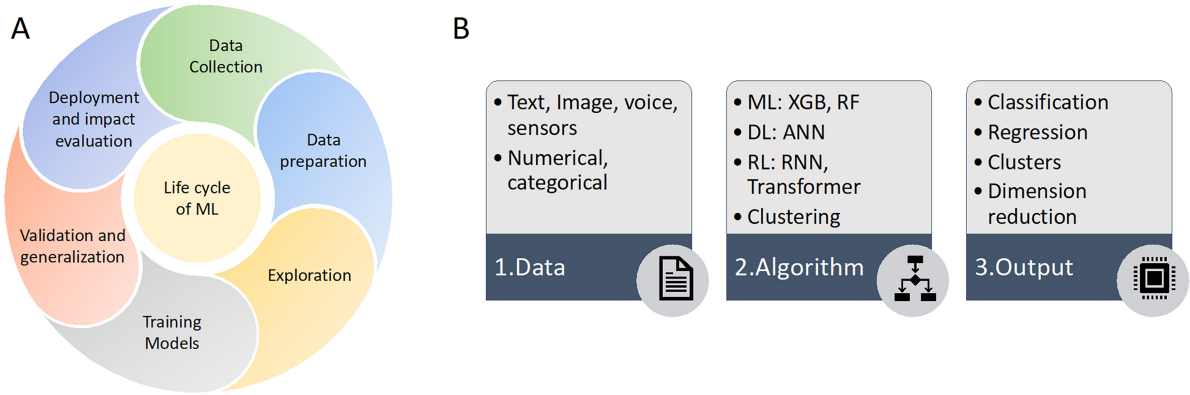

Figure 2. Basics of machine learning. (A) life cycle of ML models; (B) Key ingredients of ML models, i.e., data, algorithm, and output. ML: Machine learning; XGB: extreme gradient boost; RF: random forest; DL: deep learning; ANN: artificial neural network; RNN: recurrent neural networks; RL: reinforcement learning.

must be established. This involves defining the research question that the AI/ML process seeks to answer and highlighting its public health or scientific relevance[11].

The data collection phase is next, which may encompass both structured data (e.g., spreadsheets) and unstructured data such as text, voice, images, and sensor outputs. Data cleaning follows, where duplicate entries, errors, outliers, and missing data are addressed, and validation steps and annotations are applied. Exploratory analysis helps identify the distribution of key features across categories of interest[12]. Subsequently, the dataset is split into training and validation cohorts. The models are trained, validated, and

generalized across diverse populations and geographical regions. Explainability becomes a critical aspect, allowing for the interpretation of the most influential variables and insights into how predictions are made for individual cases[12,13]. Finally, real-world deployment enables the evaluation of models' impacts on clinical and socio-economic outcomes relevant to the disease under study.

In essence, three core components underpin any AI/ML model: data, algorithm, and the desired output [Figure 2B]. The interplay of these components leads to three primary AI/ML paradigms: supervised, unsupervised, and reinforcement learning. Evaluating the effectiveness of ML and deep deep learning (DL) models in medical imaging involves assessing their performance using appropriate metrics and ensuring their clinical utility through rigorous validation. Common evaluation metrics include accuracy, sensitivity, specificity, Dice similarity coefficient, Jaccard index, and area under the receiver operating characteristic curve (AUC-ROC). These metrics provide insights into a model's ability to identify and classify medical images correctly. However, selecting the appropriate reference standard, such as expert annotations or histopathological findings, is crucial for meaningful evaluation. Inadequate or biased reference standards can lead to misleading performance assessments[14].

Determining the clinical utility of ML or DL models extends beyond statistical performance; it requires evaluating their impact on clinical decision making and patient outcomes. Key considerations include the model's integration into clinical workflows, its interpretability by healthcare professionals, and its generalizability across diverse patient populations. Prospective clinical trials and real-world studies are essential to assess these aspects, ensuring that the deployment of such models leads to tangible improvements in healthcare delivery[15].

In supervised learning (e.g., linear or logistic regression, decision trees, Bayesian networks, support vector machines), the model is trained on labeled data. Here, the algorithm learns the relationship between input features - such as patient demographics or medical imaging - and labeled outcomes. In unsupervised learning techniques (e.g., principal component analysis, clustering methods), the data are not labeled, and the algorithm identifies hidden structures or patterns within the dataset. In reinforcement learning, models iteratively learn from the consequences of their actions to determine optimal behavior in a given context[16]. Supervised learning is one of the most widely used approaches in healthcare. For instance, consider a structured dataset containing patient histories, clinical findings, laboratory results, and a diagnosis label. The first step involves training an ML model using this dataset. Once trained, the model is tested on a separate dataset to evaluate its ability to correctly identify diagnoses. When presented with a new case, the trained model uses key features to classify and predict the diagnosis[17].

In contrast, unsupervised learning might involve an algorithm tasked with analyzing liver tumor images. In this case, the algorithm is not pre-trained on the dataset and has no prior knowledge of the features it will encounter. The algorithm's task is to identify image features on its own by clustering the dataset based on image similarities[18].

DL has garnered significant attention in recent years. These models are inspired by neural networks in the human brain, where dendrites receive, soma processes, and axons transmit information. Similarly, DL models consist of multiple layers of neurons connected by mathematical functions. The input layer gathers the data, the hidden layers process it, and the output layer provides predictions, classifications, or insights, depending on the task at hand. Unlike traditional ML models, neural networks can automatically extract features from data and learn independently[19]. DL models are especially effective in handling large, complex, and high-dimensional datasets. The rise in the volume and complexity of healthcare data, along with

advancements in computational power, has led to increased DL adoption. Modern AI/ML advancements often leverage DL techniques such as convolutional neural networks (CNNs) for image recognition and recurrent neural networks (RNNs) for predictive modeling[20]. Figure 3 showcases various AI-based models and their applications in hepatology.

ROLE OF ML MODELS AND AI IN PREDICTING DIAGNOSIS AND PROGNOSIS IN ALD

Knowledge gaps

The field of AI and ML, in particular, has rapidly expanded. However, the use of such models must be directed toward specific knowledge gaps. In ALD, specific knowledge gaps exist in predicting and identifying excessive alcohol use and making decisions in linkage to care, deciphering the ever-expanding field of metabolomics to identify clinically important pathways, distinguishing ALD from MASLD, and identifying the most at-risk group for short-term mortality among alcoholic hepatitis, thus mandating urgent and novel treatments. In the following sections, we review the literature on these specific gaps in knowledge and the application of ML-based models in such domains.

Role in alcohol use disorder

Identifying excessive alcohol usage is one key to addressing the issue at its root cause and developing preventive strategies. Preliminary studies from a two-center cohort study comparing multiple ML algorithms showed the highest predictive performance of the elastic-net machine-learning algorithm, with AUCs of 0.86 and 0.85 in Canadian and Australian cohorts, respectively, using the validated "Detection of alcohol and drug problems in adolescents (DEP-ADO)" questionnaire as a reference[21]. Those with alcohol use disorder need to be identified for appropriate treatment. Lee *et al.* demonstrated a 10-measure alternating decision tree (ADT) derived from a more extensive set of 178 clinical measures that best classified individuals as treatment-seeking or non-treatment-seeking for AUD. All alternative methods were inferior to the ADT method for accuracy and kappa statistic as a classifier, except for the simple logistic model. In addition, the comparison decision tree approaches produced either a single large, complex decision tree with 576 nodes or many small decision trees. Accordingly, neither random tree, random forest (RF), nor logistic regression was clinically useful. In contrast, the ADT approach did aim to reduce a large number of measures from the dataset into a streamlined battery composed of a subset of measures[22]. To further evidence in the field, Roberts *et al.* have demonstrated the use of ML models for predicting heavy alcohol use in those undergoing treatment for alcohol use disorder. The authors used RF algorithms and "leave sites out" partitioning to validate the models externally. ML-based models had the best classification accuracy compared to logistic regression in both the internal cross-validation and external data set for the tendency to engage in heavy drinking in between therapy sessions[23]. Lastly, ML-based models have also been used to predict alcohol use disorder remission, using multiple input sources such as electroencephalogram (EEG) source-level functional brain connectivity, polygenic risk scores (PRS), medications, and demographic information. In the male model with PRS, EEG functional connectivity and marital and employment status features demonstrated the highest accuracy of 86.04% compared to single domain-based models. Additionally, the authors identified several discriminatory factors related to neuroticism, depression, aggression, years of education, and alcohol consumption phenotypes[24].

ROLE IN ALD

Pathophysiological and diagnostic approaches

Gut microbiota alterations are pivotal in the development and downstream complications of ALD. Park *et al.* demonstrated the use of gut microbiota-based (Fecal 16S rRNA sequencing data) ML algorithms. Authors used both supervised (support vector machine, RF, multilevel perceptron, and convolutional neural network) and unsupervised (independent component analysis, principal component analysis, linear discriminant analysis, and random projection) models, with the neural network combined with principal

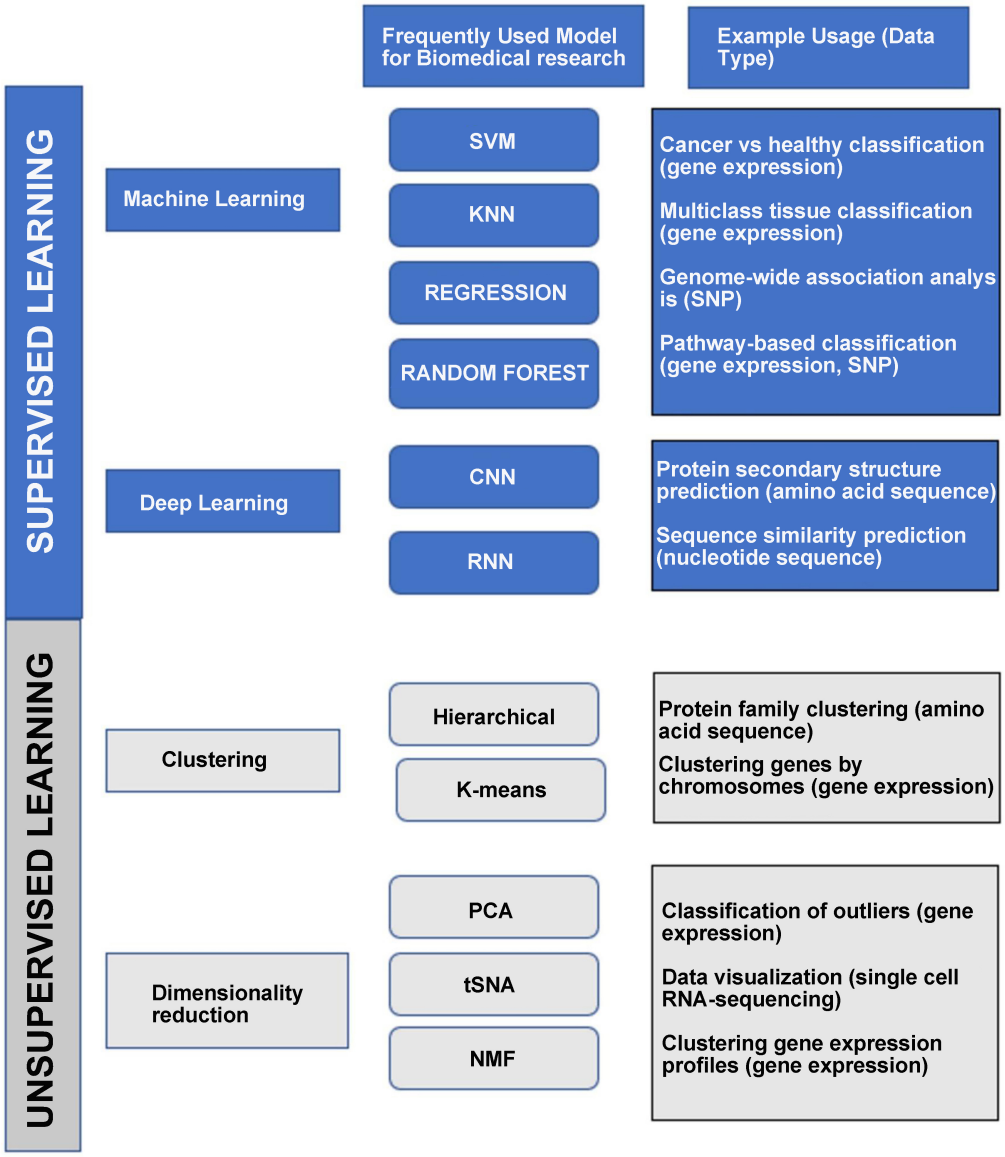

Figure 3. Types of machine learning, their objectives, name of algorithms, and key examples in medicine. CNN: Convolutional neural networking; PCA: principal component analysis; KNN: K nearest neighbor; SVM: support vector machine algorithm; RNN: recurrent neural networks; NMF: non-negative matrix factorization.

component analysis achieving the best AUCs of > 0.90 as compared to independent component analysis, random projection, and support vector machine algorithm (SVM)[25]. Similarly, magnetic resonance imaging-based DL models have been shown in a proof-of-principle study to aid in discriminating liver cirrhosis etiology based on alcohol as an etiology[26]. Multi-omic approaches for developing diagnostic and prognostic liquid biopsy strategies are promising, specifically in ALD. Listopad *et al.* used transcriptomics and proteomics in liver tissue and peripheral blood mononuclear cells to develop highly accurate ML-based multi-omic models. Such computational approaches to identify blood-based diagnostic biomarkers can, in the future, contribute to developing highly precise blood tests, mitigating the need for liver biopsy[27].

Distinguishing ALD from other aetiologies

MASLD is the fastest-growing liver disease globally. Paradoxically, alcohol consumption also shows global upward trends. Based on this, distinguishing alcohol and non-alcohol aetiologies becomes essential. Markers like aspartate transaminase/alanine transaminase ratio, gamma-glutamyl transpeptidase, mean corpuscular volume, and composite scores like alcohol-non alcohol index are modestly helpful[28]. A recent study explored such markers in the setting of MASLD and ALD and found that although good at extremes, such scores perform only modestly in the intermediate cohort of Met-ALD[29]. Hence, a knowledge gap remains in differentiating such grey zones where ML models may have a role. In this context, Sowa *et al.* used ML techniques relying on the ALT/AST ratio, adipokines, and cytokines to distinguish MASLD from ALD (AUC = 0.91). In addition, using cytokine-based ML, the severity of ALD was also predicted (AUC = 0.98)[30]. However, further application to the cohort of Met-ALD is required. A similar challenge often exists in differentiating acute alcoholic hepatitis from acute cholangitis in those consuming alcohol, as both present as jaundice with elevated aminotransferase. Ahn *et al.* used ten readily available laboratory parameters and compared eight supervised ML techniques (decision tree, naive Bayes, logistic regression, k-nearest neighbor, support vector machine, artificial neural networks, RF, and gradient boosting) to differentiate the entities. Using a feature selection strategy to choose the best five variables, ML models showed excellent performances with AUCs of 0.98 and 0.97 in derivation and validation sets, respectively[31].

Role in mortality prediction

Accurate prediction of patients at greatest risk of adverse outcomes and the need for definitive therapy remains a significant challenge in patients with alcoholic hepatitis. Non-invasive traditional scores like Maddrey's discriminant function (mDF), model for end-stage live disease (MELD), and Glasgow alcoholic hepatitis score (GAHS) have been used but have their limitations. Gao *et al.* used ML approaches using the gradient boosting analysis of bacterial and metabolic pathways datasets to achieve excellent 30-day mortality prediction (AUC = 0.87). Notably, multiple pathways were related to purine nucleoside biosynthesis, which plays an essential role in *in vivo* bacterial growth in humans[32]. In a large multicentric study across 23 centers and 12 countries, Dunn *et al.* proposed the Alcoholic Hepatitis Artificial Intelligence Ensemble score, integrating age, gender, cirrhosis, and nine laboratory values with center-specific mortality rates. The score outperformed conventional scores like mDF, MELD, and GAHS, with an AUC of 0.81 and 0.79 in the derivation and validation cohorts[33].

Role in predicting post-transplant alcohol relapse

Relapse to alcohol use after liver transplantation is one of the key deterrents to the applicability of liver transplant in alcoholic hepatitis and alcohol-related cirrhosis. Multiple scores have been used to predict alcohol relapse after transplantation[34]. Lee *et al.* used an XGBoost-based ML model to predict post-LT harmful alcohol use. With a median follow-up of 4.4 (IQR 3.0-6.0) years post-LT, the model showed an AUC of 0.930 (95%CI: 0.862-0.998) in the training cohort and 0.692 (95%CI: 0.666-0.718) in the validation cohort[35]. However, it must be pointed out that although having a robust AUC, the overall poor positive predictive value even in the highest susceptible group (maximum score of 11) is poor.

Applications and translation with AI-based devices and approaches

Wearable devices using AI-based algorithms are important translational outcomes in ALD. In a pilot study, Jalal *et al.* demonstrated a wearable blood alcohol concentration monitoring device to reflect the volatility and variation of alcohol concentration on the skin according to blood alcohol concentration changes[36]. In a small pilot study, the use of digital phenotyping for alcohol use disorder (AWARE application), with features like accelerometer magnitude, number of calls, and location entropy, showed a significant association with alcohol craving[37]. Such data provide new insights into the use of smartphone sensors as markers for alcohol craving and mood in ALD and alcohol use disorder. Table 1 summarizes some key

Table 1. Selected studies showing the application of machine learning and artificial intelligence in alcohol-related liver disease

Authors	Study domain and goals	Number of subjects	Performance
Roberts *et al.*[23]	Predicting heavy alcohol use during outpatient treatment	1,383	Training cohort: AUC = 0.89 Validation cohorts: AUC range = 0.80 to 0.87
Kinreich *et al.*[24]	Predicting alcohol use disorder remission using EEG, polygenic risk scores and demographic variables	1,376	AUC of 0.86 in the training cohort
Park *et al.*[25]	Gut microbial analysis to distinguish ALD from MASLD	Derivation: 263 ALD 201 MASLD (For validation: 126 ALD and 84 MASLD)	Derivation cohort: AUC of CNN combined with PCA (AUCs) > 0.93 Validation AUC 0.90
Dunn *et al.*[33]	Predicting 90-day mortality in alcoholic hepatitis	Derivation: 860 patients Validation: 859 patients	Validation cohort 30- and 90-day AUCs were 0.811 (0.779-0.844) and 0.799 (0.769-0.830), respectively
Ahn *et al.*[31]	Differentiating acute alcoholic hepatitis from acute cholangitis	Derivation: 459 patients Validation: 205 143 physicians used as comparators	Derivation AUC 0.98 Validation AUC 0.97 Physician accuracy 0.79
Lee *et al.*[35]	Predicting post liver transplant alcohol relapse	91 in the training cohort and 25 in the validation cohort	Training cohort: AUC 0.930, PPV: 0.891 Validation cohort: AUC 0.692, PPV 0.82

ALD: Alcohol-associated liver disease; MASLD: metabolic dysfunction-associated steatotic liver disease; CNN: convolutional neural network; PCA: principal component analysis; PPV: positive predictive value; AUC: area under curve; EEG: electroencephalogram.

studies in ALD[23,24,25,33,31,35].

ROLE OF ML MODELS AND AI IN PREDICTING DIAGNOSIS AND PROGNOSIS IN MASLD

Knowledge gaps

Similar to ALD, AI and ML have seen a rapid increase in MASLD. However, some of the key areas that remain as definitive knowledge gaps revolve around variability in histopathology reporting, predicting the presence of steatosis, identifying at-risk metabolic dysfunction-associated steatohepatitis (MASH), identifying significant and advanced fibrosis, and ultimately predicting clinically significant portal hypertension (CSPH) and major adverse liver events.

Role in MASLD histology

Manual histological assessment is the gold standard for diagnosing and monitoring disease progression in MASH, but it has high interobserver variability. ML-based approaches using histological samples have been shown to accurately characterize disease severity and heterogeneity[38]. AI-based analysis has also been shown to be extremely sensitive in demonstrating treatment-induced reversal of fibrosis compared to conventional scoring[39]. Lastly, using AI-assisted tools (fibrosis) has been shown to improve inter-pathologist agreement among pathologists with varying experience[40]. Other studies have looked at the development and validation of AI models, leveraging microscopy along with insights from an expert hepatopathologist. A total of 25 liver biopsies from the trial on drug Belapectin were randomly selected for training, and an additional 10 for validation. The outcomes emphasized the crucial role of disease-specific customization of AI models, based on expert pathologist training, in improving accuracy and applicability in clinical trials[41]. Ballooned hepatocytes represent a specific histopathological problem, where even among expert hepatopathologists, there is poor agreement regarding the number of ballooned hepatocytes seen on the same digitized histology images. AI has been shown to provide a more reliable way to assess the range of injury recorded as "hepatocyte ballooning"[42]. Figure 4 shows a schematic for understanding concepts of applying AI-based models in digital pathology in MASH.

Predicting the onset of MASLD

While MASLD is one of the most common causes of liver disease globally, it is also of interest to predict

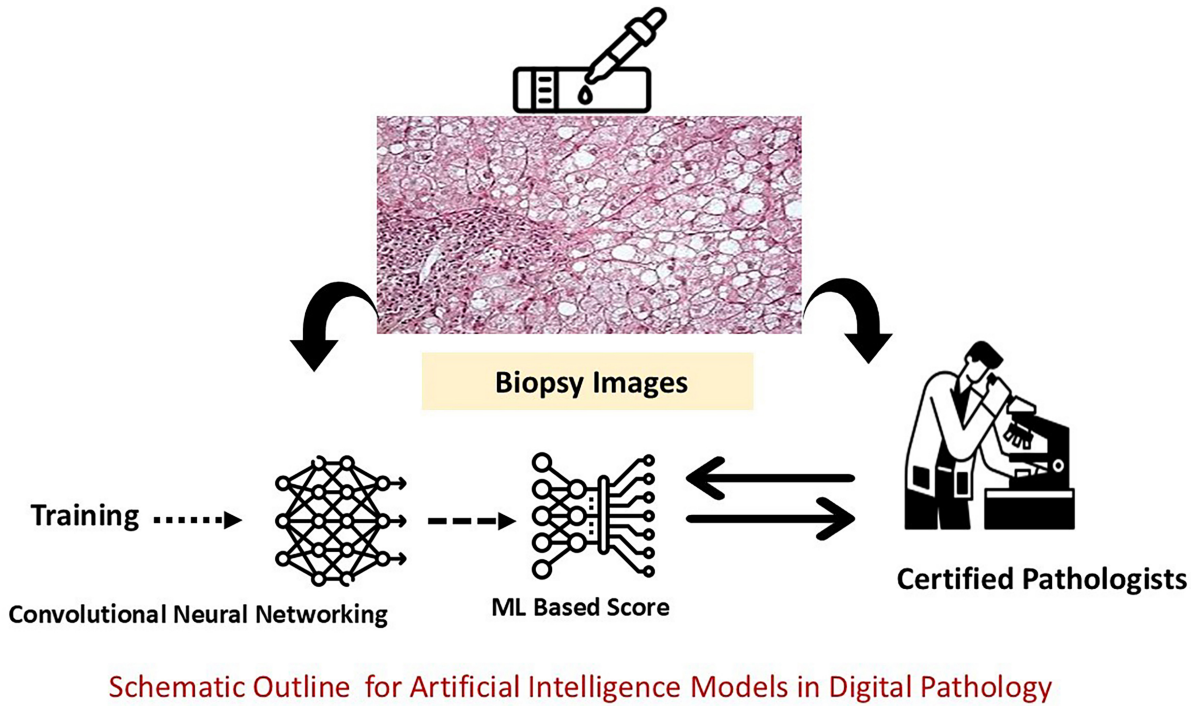

Figure 4. Schematic representation of AI applications in digital pathology. ML: Machine learning.

who will develop incident MASLD. Such premises make an ideal case scenario for the application of ML algorithms driven by large datasets. Lim *et al.* worked to answer such a question using a development dataset comprised of 25,599 individuals from a South Korean non-alcoholic fatty liver disease (NAFLD) registry using Random Survival Forest and Extra Survival Trees. In derivation and validation sets of 331,107 and 543,874 person-months of follow-up, MASLD incidence was 25.7% and 14.4%, with a median time to MASLD onset of being 60 (IQR 38-75) and 24 (IQR 13-37) months, respectively. Both the ML models achieved a C-index of > 0.7 in the validation cohort[43]. While Lim *et al.*[43] used single-point variables, Deng *et al.* used a dynamic model using five constantly updated checkup data to predict the risk of NAFLD at and after their sixth health checkup. Intriguingly, the authors showed that the DL model's predictive performance improved over time, with AUC increasing from 0.72 at baseline to 0.818 when five consecutive checkups were included[44].

Predicting MASLD progression and identifying MASH

Using electronic health records of patients who received an index diagnosis of MASLD, Ghandian *et al.* used a gradient-boosted ML algorithm (XGBoost) and multi-layer perceptron models to predict the development of non-alcoholic steatohepatitis (NASH) or fibrosis within four years using demographic features, vital signs, and laboratory measurements. The XGBoost algorithm achieved an AUC of 0.79 for predicting progression to NASH and 0.87 for fibrosis based on International Classification of Diseases, Tenth Revision (ICD-10) codes[45]. Identifying at-risk MASH using conventional regression models is an important knowledge gap. Njei *et al.* used five simple indices: ALT, GGT, platelet count, waist circumference, and age to predict high-risk MASH, defined as a FAST score of ≥ 0.35 and ≥ 0.67. An XGBoost model achieved high diagnostic accuracy with an AUC of 0.95 compared to 0.5 for FIB-4, APRI, and NAFLD Fibrosis score[46]. Similarly, Naderi Yaghouti *et al.* used RF with sequential forward selection to obtain an accuracy of 81.2% ± 6.4% in predicting NASH as per the NAFLD activity score[47]. Along similar lines, Lee *et al.* showed better performance with independent ML models using only clinical predictors for

MASH and at-risk MASH[48].

Role in predicting significant and advanced fibrosis

Predicting significant and advanced fibrosis is one of the cornerstones of MASLD risk stratification. Most conventional risk prediction models are based on regression techniques. Charu *et al.*, in an elegant study, used an ensemble machine-learning algorithm, the "super learner", to benchmark the performance of clinical risk prediction algorithms, especially those based on simple regression techniques. The "superlearner" exhibited excellent discriminative abilities for fibrotic MASH as compared to existing models [Fibrosis-4 (FIB-4), NAFLD fibrosis score, Forns, AST to Platelet Ratio Index (APRI), BARD, and Steatosis-associated fibrosis estimator (SAFE)], with AUCs of 0.79 (95%CI: 0.73-0.84) and 0.74 (95%CI: 0.68-0.79) in the derivation and validation sets, respectively, thus suggesting a role to benchmark the performance of conventional clinical risk prediction models[49]. In an extensive study in patients with biopsy-proven MASLD, Verma *et al.* tested twenty-one ML models (training cohort, N = 1,153, testing cohort, N = 283, validation cohort, N = 220) using clinical and biochemical parameters. The ML models showed 7%-12% better discrimination than FIB-4 for significant fibrosis. Optimized RF yielded the best negative predictive value. Compared to FIB-4, RF could pick ten times more patients with significant fibrosis (SF), reduce unnecessary referrals by 28%, and prevent missed referrals by 78%[50].

Predicting CSPH and clinical complications

The development of CSPH marks an epoch in the natural history of compensated advanced chronic liver disease. The current BAVENO recommendations define CSPH on varying probabilities based on non-invasive tests like LSM as CSPH present at LSM ≥ 25 kPa and absent at CSPH: LSM < 15 kPa and PLT ≥ 150 × 10^9/L with a grey zone for those with LSM between 20-25 kPa and PLT < 150 × 109/L, or LSM between 15-20 kPa and PLT < 110 × 10^9/L[51]. Looking at histological features to predict CSPH is a novel approach that was demonstrated by Noureddin *et al.* to construct a score called SNOF (septa, nodules, and fibrosis) using 448 histological variables. A SNOF score of ≥ 11.78 reliably distinguished CSPH (AUC = 0.85), opening insights into using histology to estimate CSPH[52]. Hepatic decompensation and all mortality form the hard endpoints in MASLD research. The use of digital pathology to predict such endpoints is feasible. Using an artificial intelligence-based image processing algorithm of > 150 image outputs, Kendall *et al.* developed models for hepatic decompensation [Clinical Outcome Decompensation Index-Fibrosis ("CODI-F")] and mortality [Clinical Outcome Mortality Index-Fibrosis ("COMI-F")]. Both tools could directly predict hard endpoints and demonstrated predictive value at least equivalent to traditional or computational ordinal fibrosis scores[53].

Applications and translation with AI-based devices and approaches

An electrocardiogram (ECG) is one of the simplest and most readily available tests in medical practice. ECG has been shown to predict the presence of cirrhosis. Data from 5,212 patients with cirrhosis who underwent liver transplantation were used to construct an AI-Cirrhosis-ECG (ACE) score, which showed excellent accuracy (AUC = 0.908, 84.9% sensitivity, 83.2% specificity)[54]. The score has also been used to detect MASLD, using 3,468 MASLD cases and 25,407 controls with an AUC of 0.69 (original cohort) and 0.62 (validation cohort). The AI model performance was similar or superior to age- and sex-adjusted models using body mass index (BMI) (AUC = 0.71), presence of diabetes, hypertension or hyperlipidemia (AUC = 0.68), or diabetes alone (AUC = 0.66)[55]. Digital therapeutics has seen a rapid evolution and application to provide evidence-based interventions to prevent, manage, and treat diseases. Sato *et al.* used a computer-generated smartphone to demonstrate histological improvement in patients with NASH, supported by weight loss with a > 50% reduction in fibrosis stage in those with significant fibrosis[56]. Animal model studies have shown that an on-skin impedance sensor and an attention-based deep-learning technique can detect MASLD early, although such techniques await clinical translation[57]. Table 2 summarizes some of the

Table 2. Selected studies showing the application of machine learning and artificial intelligence in MASLD

Authors	Study domain and goals	Number of subjects	Performance and comments
Taylor-Weiner *et al.*[38]	Characterizing histological disease severity and heterogeneity	166 slides	Cohen's kappa 0.81 (model agreement) with pathologist consensus Superior prognostic utility compared with manual pathological features
Lim *et al.*[43]	Predicting the onset of MASLD	Derivation: 1,331, 107 Validation: 543, 874	C-index of > 0.7 in the validation cohort Random survival forest 0.751 (95%CI: 0.742-0.759), extra survival trees 0.752 (95%CI: 0.744-0.762)
Ghandian *et al.*[45]	Predicting progression to MASH or Fibrosis from MASLD	141,293 patients, 4,384 and 4,472 of whom were eventually diagnosed with NASH or fibrosis, respectively	AUROC of 0.79 and 0.87 for prediction of progression to NASH and fibrosis, respectively
Verma *et al.*[50]	Prediction of significant fibrosis in MASLD	Training cohort: 1,153 Testing cohort: 283 Validation cohort: 220	NPV in overall set 0.94/0.75 Testing cohort 0.79/0.58) Validation cohort 0.85/0.55 RF could pick > 10 times more patients with significant fibrosis and reduce unnecessary referrals by 28%
Kendall *et al.*[53]	Histological features to predict decompensation in MASLD	452 liver biopsy sections	Decompensation index (> .31), HR 5.96, $P < .001$
Udompap *et al.*[55]	Prediction of MASLD using ECG	3,468 MASLD cases and 25,407 controls	Derivation 0.69 Validation 0.62
Wang *et al.*[57]	Predicting MASLD using skin sensors	Animal model	Derivation 1.0

MASLD: Metabolic dysfunction-associated steatotic liver disease; AUORC: area under receiver operative curve; NPV: negative predictive value; RF: random forest; ECG: electrocardiogram; MASH: metabolic dysfunction associated steatohepatitis; NASH: non-alcoholic steatohepatitis.

key studies on MASLD[38,43,45,50,53,55,57].

SUMMARY OF STUDIES ON APPLICATIONS OF IMAGING DATA IN ALD AND MASLD

DL and ML have significantly advanced the use of medical imaging in diagnosing, treating, and predicting outcomes for ALD and MASLD. These technologies enhance the analysis of imaging modalities, enabling more accurate detection and characterization of liver abnormalities. For instance, DL algorithms have been developed to automatically detect hepatic steatosis on CT images, achieving high diagnostic accuracy and facilitating early intervention strategies. In prognostic realms, AI models have been employed to predict disease progression and patient outcomes in SLD. By analyzing imaging data alongside clinical parameters, these models can stratify patients based on risk, aiding in personalized treatment planning. We summarize few key studies in Table 3[58,59,60,61].

LARGE LANGUAGE MODELS

Large language models (LLMs) have emerged as transformative tools in healthcare, offering the potential to analyze vast and complex datasets, generate clinical predictions, and enhance personalized medicine[62]. These models, based on DL architectures, are capable of processing large amounts of unstructured data such as electronic health records, medical literature, and multi-omic datasets, providing clinicians with valuable insights for patient management. These models have shown remarkable potential in handling clinical tasks, including data extraction, literature summarization, content generation, predictive modeling, clinical decision support, and patient-provider chatbots[63]. However, general-purpose LLMs like OpenAI's ChatGPT are trained on publicly available datasets and are not specifically optimized for clinical applications. As a result, when tasked with clinical queries, these models may produce outputs that contain inaccuracies, incomplete data, or "hallucinations" - fabricated information that lacks a factual basis[64]. Despite these shortcomings, LLMs are believed to hold immense promise for biomedical and clinical purposes due to the complexity and growing volume of medical knowledge. For instance, it was estimated

Table 3. AI applications in using imaging data

Authors23456	Domain	Principle	ML models	Number of subjects	Performance metrics
Ghosh *et al.*[58]	Prognosis	Reviewed AI methods applied to omics data and imaging in liver diseases to predict disease progression	Various AI methodologies	Not specified	AI models showed promise in prognosticating liver disease outcomes, including those related to ALD
Fujii *et al.*[59]	Diagnosis	Image analysis and quantitative assessment of the contour of the sagittal section of the left lobe of the liver	Deep learning	486	Liver surface roughness was correctly judged by AI. Image analysis showed the thickness of the left lobe was inversely correlated with the liver fibrosis stage
Santoro *et al.*[60]	Diagnosis	AI to non-invasively assist US-Mediated diagnosis of early-stage steatotic liver	Deep learning	134	Discrimination capacity by AUC between patient with steatosis and patient without steatosis was better for AI-based hepatorenal index than manual than HRIM (AUC: 0.87 *vs.* 0.82, respectively). ROC analysis showed an AUC = 0.98 for HRIA with a 1.64 cut-off in distinguishing between mild and moderate/severe groups
Li *et al.*[61]	Staging	AI-powered models utilizing non-contrast MRI, including T1WI and T2FS, can accurately stage LF	Two CM	1,726	Fusion models yielded the highest AUC among the CMs, achieving AUCs of 0.8-0.9 for significant to advanced fibrosis and cirrhosis and significantly surpassed transient elastography (only for staging ≥ F2 and ≥ F3 grades), serum biomarkers, and three junior radiologists for staging LF. Radiologists, with the aid of the OMs, could achieve a more accurate LF assessment

AI: Artificial intelligence; ML: machine learning; US: ultrasound; LF: liver fibrosis; AUC: area under curve; ROC: receiver operative curve; CMs: classification models; MRI: magnetic resolution imaging; HRIM: hepatorenal index manual; HRIA: hepatorenal index automatic; CM: classification models; OM: optimal model.

that two research papers were added to PubMed every minute, a number that has likely increased in recent years[65]. Thus, efforts to integrate clinical practice guidelines and medical literature into LLMs are gaining traction as a way to adapt general-purpose models for specialized clinical applications.

The strategies to make LLMs more clinically focused involve several steps. The first involves fine-tuning the original model, although this approach is computationally intensive. The second strategy is prompt engineering, where users provide specific prompts to guide the model's responses, although this method can only accommodate small datasets and may require multiple iterations. The third, more promising approach is retrieval-augmented generation (RAG). In this framework, an LLM is enhanced by coupling it with an information retrieval system that provides relevant external data, such as clinical guidelines. The external dataset is vectorized and encoded using embedding models, and this structured information is integrated into the LLM to guide its outputs. RAG offers two significant benefits: it allows the LLM to handle large volumes of documents, providing a solid knowledge base, and it reduces the risk of generating incorrect or fabricated information by narrowing the solution space for the model's outputs. This method has been explored in the treatment of patients with hepatitis C, and the LLM-generated responses showed 99% accuracy compared to expert-provided answers, as measured by text similarity scores[66]. Other LLMs like "LiVersa" trained on 30 AASLD guidance documents could enhance the accuracy of responses[63].

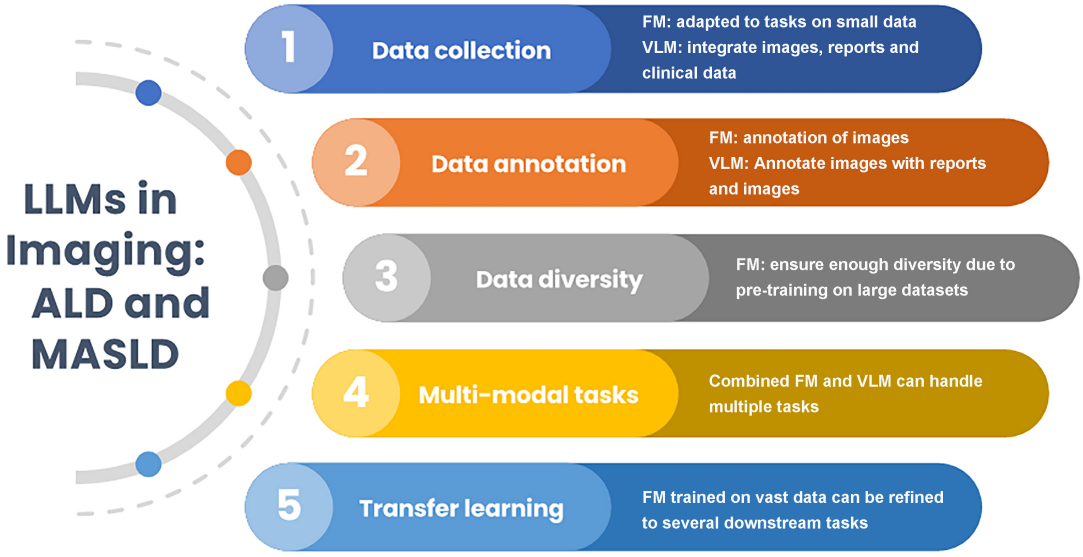

Figure 5. Various applications of FM and VLM in imaging (pathology or radiology) for ALD and MASLD. FM: Foundation models; VLM: vision language models; ALD: alcohol-associated liver disease; MASLD: metabolic dysfunction-associated liver disease; LLMs: large language models.

Computational pathology, utilizing advanced AI models like foundation models (FMs) and vision-language models (VLMs), offers transformative potential in the management of ALD and MASLD[67]. FMs, trained on vast amounts of pathology data through self-supervised learning, can adapt to various downstream tasks without requiring large-scale annotated datasets, which are often challenging to collect. FMs can be deployed to analyze liver tissue samples, identify disease patterns, grade severity, and even predict disease progression by integrating histological images with clinical data[68]. This can help clinicians make timely decisions regarding treatment strategies and prognosis. VLMs add another layer of capability by combining image analysis with natural language processing, enabling the integration of pathology reports and other clinical narratives into predictive models[69]. For instance, in MASLD, VLMs can analyze liver biopsy images alongside pathology reports to provide a comprehensive assessment of disease severity or predict which patients may progress to advanced fibrosis. In ALD, VLMs can similarly assist in diagnosing complications like cirrhosis by interpreting both imaging data and associated clinical descriptors. These models, which can generate natural language predictions, offer explainability and interpretability, critical for assessing varied presentations in ALD and MASLD. Overall, the combination of FMs and VLMs represents a new era in computational pathology, offering more personalized and accurate management strategies for liver diseases like ALD and MASLD [Figure 5].

LIMITATIONS

While the universe of AI and ML witnesses rapid expanses, understanding intrinsic limitations is essential. A simple glance at ML-based models shows them to have excellent "diagnostic accuracy numbers". However, they often suffer from clinical explainability and applicability[70]. The settings where such models are designed to be applied are of paramount importance, as complex models with large laboratory variables will not be applicable when devising screening and risk-stratifying strategies. Secondly, such models are at risk of replicating and amplifying intrinsic biases in data, potentially leading to misdiagnosis and misclassification, and this mandates the use of diverse training sets and regular audits. Lastly, there remains

a lack of standardization of ML-based modeling, posing challenges for comparison and reproducibility across studies[71].

CONCLUSION

In conclusion, AI and ML-based modeling have emerged as exciting frontiers in both ALD and MASLD. Substantial knowledge gaps remain, which conventional modeling systems have failed to address. The use of ML-based approaches should address such specific lacunas, translating to clinically meaningful objectives and aiding in decision making.

DECLARATIONS

Authors' contributions

Writing of the original draft, review and editing: Roy A
Conceptualization, supervision, writing of the original draft, review and editing: Verma N

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Younossi ZM, Wong G, Anstee QM, Henry L. The global burden of liver disease. *Clin Gastroenterol Hepatol*. 2023;21:1978-91. DOI PubMed
2. Åberg F, Byrne CD, Pirola CJ, Männistö V, Sookoian S. Alcohol consumption and metabolic syndrome: clinical and epidemiological impact on liver disease. *J Hepatol*. 2023;78:191-206. DOI PubMed
3. Hsu CL, Loomba R. From NAFLD to MASLD: implications of the new nomenclature for preclinical and clinical research. *Nat Metab*. 2024;6:600-2. DOI PubMed PMC
4. Rinella ME, Lazarus JV, Ratziu V, et al; NAFLD Nomenclature consensus group. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *J Hepatol*. 2023;79:1542-56. DOI PubMed
5. Malnick SDH, Alin P, Somin M, Neuman MG. Fatty liver disease-alcoholic and non-alcoholic: similar but different. *Int J Mol Sci*. 2022;23:16226. DOI PubMed PMC
6. Ha J, Yim SY, Karagozian R. Mortality and liver-related events in lean versus non-lean nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2023;21:2496-507.e5. DOI PubMed
7. Ng CH, Lim WH, Hui Lim GE, et al. Mortality outcomes by fibrosis stage in nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2023;21:931-9.e5. DOI PubMed PMC
8. Strandberg R, Jepsen P, Hagström H. Developing and validating clinical prediction models in hepatology - an overview for clinicians. *J Hepatol*. 2024;81:149-62. DOI PubMed
9. Bhat M, Rabindranath M, Chara BS, Simonetto DA. Artificial intelligence, machine learning, and deep learning in liver transplantation. *J Hepatol*. 2023;78:1216-33. DOI PubMed

10. Kalapala R, Rughwani H, Reddy DN. Artificial intelligence in hepatology- ready for the primetime. *J Clin Exp Hepatol*. 2023;13:149-61. DOI PubMed PMC
11. Ashmore R, Calinescu R, Paterson C. Assuring the machine learning lifecycle: desiderata, methods, and challenges. *ACM Comput Surv*. 2022;54:1-39. DOI
12. Xing EP, Ho Q, Xie P, Wei D. Strategies and principles of distributed machine learning on big data. *Engineering*. 2016;2:179-95. DOI
13. Blanco-Justicia A, Domingo-Ferrer J. Machine learning explainability through comprehensible decision trees. Springer International Publishing; 2019. pp. 15-26. DOI
14. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Res Notes*. 2022;15:210. DOI PubMed PMC
15. Farah L, Murris JM, Borget I, Guilloux A, Martelli NM, Katsahian SI. Assessment of performance, interpretability, and explainability in artificial intelligence-based health technologies: what healthcare stakeholders need to know. *Mayo Clin Proc Digital Health*. 2023;1:120-38. DOI
16. Lo Vercio L, Amador K, Bannister JJ, et al. Supervised machine learning tools: a tutorial for clinicians. *J Neural Eng*. 2020;17:062001. DOI
17. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. *Behav Ther*. 2020;51:675-87. DOI PubMed PMC
18. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. *Am J Epidemiol*. 2019;188:2222-39. DOI PubMed
19. Georgevici AI, Terblanche M. Neural networks and deep learning: a brief introduction. *Intensive Care Med*. 2019;45:712-4. DOI PubMed
20. Al-Askar H, Radi N, Macdermott Á. Recurrent neural networks in medical data analysis and classifications. Elsevier; 2016. pp. 147-65. DOI
21. Afzali MH, Sunderland M, Stewart S, et al. Machine-learning prediction of adolescent alcohol use: a cross-study, cross-cultural validation. *Addiction*. 2019;114:662-71. DOI PubMed
22. Lee MR, Sankar V, Hammer A, et al. Using machine learning to classify individuals with alcohol use disorder based on treatment seeking status. *EClinMed*. 2019;12:70-8. DOI PubMed PMC
23. Roberts W, Zhao Y, Verplaetse T, et al. Using machine learning to predict heavy drinking during outpatient alcohol treatment. *Alcohol Clin Exp Res*. 2022;46:657-66. DOI PubMed PMC
24. Kinreich S, McCutcheon VV, Aliev F, et al. Predicting alcohol use disorder remission: a longitudinal multimodal multi-featured machine learning approach. *Transl Psychiatry*. 2021;11:166. DOI PubMed PMC
25. Park IG, Yoon SJ, Won SM, et al. Gut microbiota-based machine-learning signature for the diagnosis of alcohol-associated and metabolic dysfunction-associated steatotic liver disease. *Sci Rep*. 2024;14:16122. DOI PubMed PMC
26. Luetkens JA, Nowak S, Mesropyan N, et al. Deep learning supports the differentiation of alcoholic and other-than-alcoholic cirrhosis based on MRI. *Sci Rep*. 2022;12:8297. DOI PubMed PMC
27. Listopad S, Magnan C, Day LZ, et al. Identification of integrated proteomics and transcriptomics signature of alcohol-associated liver disease using machine learning. *PLOS Digit Health*. 2024;3:e0000447. DOI PubMed PMC
28. Cerović I, Mladenović D, Ješić R, et al. Alcoholic liver disease/nonalcoholic fatty liver disease index: distinguishing alcoholic from nonalcoholic fatty liver disease. *Eur J Gastroenterol Hepatol*. 2013;25:899-904. DOI PubMed
29. Roy A, De A, Kulkarni AV, et al. Alcoholic liver disease/nonalcoholic fatty liver disease index for classification of patients with steatotic liver disease. *J Obes Metab Syndr*. 2024;33:222-8. DOI PubMed PMC
30. Sowa JP, Atmaca Ö, Kahraman A, et al. Non-invasive separation of alcoholic and non-alcoholic liver disease with predictive modeling. *PLoS One*. 2014;9:e101444. DOI PubMed PMC
31. Ahn JC, Noh YK, Rattan P, et al. Machine learning techniques differentiate alcohol-associated hepatitis from acute cholangitis in patients with systemic inflammation and elevated liver enzymes. *Mayo Clin Proc*. 2022;97:1326-36. DOI PubMed
32. Gao B, Wu TC, Lang S, et al. Machine learning applied to omics datasets predicts mortality in patients with alcoholic hepatitis. *Metabolites*. 2022;12:41. DOI PubMed PMC
33. Dunn W, Li Y, Singal AK, et al. An artificial intelligence-generated model predicts 90-day survival in alcohol-associated hepatitis: a global cohort study. *Hepatology*. 2024;80:1196-211. DOI PubMed PMC
34. Lee BP, Vittinghoff E, Hsu C, et al. Predicting low risk for sustained alcohol use after early liver transplant for acute alcoholic hepatitis: the sustained alcohol use post-liver transplant score. *Hepatology*. 2019;69:1477-87. DOI PubMed PMC
35. Lee BP, Roth N, Rao P, et al. Artificial intelligence to identify harmful alcohol use after early liver transplant for alcohol-associated hepatitis. *Am J Transplant*. 2022;22:1834-41. DOI PubMed PMC
36. Jalal AH, Arbabi S, Ahad MA, Alam F, Ahmed MA. Wearable alcohol monitoring device for the data-driven transcutaneous alcohol diffusion model. *Sensors*. 2024;24:4233. DOI PubMed PMC
37. Wu T, Sherman G, Giorgi S, et al. Smartphone sensor data estimate alcohol craving in a cohort of patients with alcohol-associated liver disease and alcohol use disorder. *Hepatol Commun*. 2023;7:e0329. DOI PubMed PMC
38. Taylor-Weiner A, Pokkalla H, Han L, et al. A machine learning approach enables quantitative measurement of liver histology and disease monitoring in NASH. *Hepatology*. 2021;74:133-47. DOI PubMed PMC
39. Naoumov NV, Brees D, Loeffler J, et al. Digital pathology with artificial intelligence analyses provides greater insights into treatment-induced fibrosis regression in NASH. *J Hepatol*. 2022;77:1399-409. DOI PubMed

40. Soon GST, Liu F, Leow WQ, Wee A, Wei L, Sanyal AJ. Artificial intelligence improves pathologist agreement for fibrosis scores in nonalcoholic steatohepatitis patients. *Clin Gastroenterol Hepatol*. 2023;21:1940-2.e3. DOI PubMed
41. Goodman Z, Akbary K, Noureddin M, et al. Enhancing histology detection in MASH cirrhosis for artificial intelligence pathology platform by expert pathologist training. *Liver Int Comm*. 2024;5:e70007. DOI
42. Brunt EM, Clouston AD, Goodman Z, et al. Complexity of ballooned hepatocyte feature recognition: defining a training atlas for artificial intelligence-based imaging in NAFLD. *J Hepatol*. 2022;76:1030-41. DOI PubMed PMC
43. Lim DYZ, Chung GE, Cher PH, Chockalingam R Jr, Kim W, Tan CK. Use of machine learning to predict onset of NAFLD in an all-comers cohort-development and validation in 2 large asian cohorts. *Gastro Hep Adv*. 2024;3:1005-11. DOI PubMed PMC
44. Deng Y, Ma Y, Fu J, et al. A dynamic machine learning model for prediction of NAFLD in a health checkup population: a longitudinal study. *Heliyon*. 2023;9:e18758. DOI PubMed PMC
45. Ghandian S, Thapa R, Garikipati A, et al. Machine learning to predict progression of non-alcoholic fatty liver to non-alcoholic steatohepatitis or fibrosis. *JGH Open*. 2022;6:196-204. DOI PubMed PMC
46. Njei B, Osta E, Njei N, Al-Ajlouni YA, Lim JK. An explainable machine learning model for prediction of high-risk nonalcoholic steatohepatitis. *Sci Rep*. 2024;14:8589. DOI PubMed PMC
47. Yaghouti AR, Zamanian H, Shalbaf A. Machine learning approaches for early detection of non-alcoholic steatohepatitis based on clinical and blood parameters. *Sci Rep*. 2024;14:2442. DOI PubMed PMC
48. Lee J, Westphal M, Vali Y, et al; LITMUS investigators. Machine learning algorithm improves the detection of NASH (NAS-based) and at-risk NASH: a development and validation study. *Hepatology*. 2023;78:258-71. DOI PubMed
49. Charu V, Liang JW, Mannalithara A, Kwong A, Tian L, Kim WR. Benchmarking clinical risk prediction algorithms with ensemble machine learning for the noninvasive diagnosis of liver fibrosis in NAFLD. *Hepatology*. 2024;80:1184-95. DOI PubMed
50. Verma N, Duseja A, Mehta M, et al. Machine learning improves the prediction of significant fibrosis in Asian patients with metabolic dysfunction-associated steatotic liver disease - the Gut and Obesity in Asia (GO-ASIA) study. *Aliment Pharmacol Ther*. 2024;59:774-88. DOI PubMed
51. Mendizabal M, Cançado GGL, Albillos A. Evolving portal hypertension through Baveno VII recommendations. *Ann Hepatol*. 2024;29:101180. DOI PubMed
52. Noureddin M, Goodman Z, Tai D, et al. Machine learning liver histology scores correlate with portal hypertension assessments in nonalcoholic steatohepatitis cirrhosis. *Aliment Pharmacol Ther*. 2023;57:409-17. DOI PubMed PMC
53. Kendall TJ, Chng E, Ren Y, Tai D, Ho G, Fallowfield JA. Outcome prediction in metabolic dysfunction-associated steatotic liver disease using stain-free digital pathological assessment. *Liver Int*. 2024;44:2511-6. DOI PubMed
54. Ahn JC, Attia ZI, Rattan P, et al. Development of the AI-cirrhosis-ECG score: an electrocardiogram-based deep learning model in cirrhosis. *Am J Gastroenterol*. 2022;117:424-32. DOI PubMed PMC
55. Udompap P, Liu K, Attia IZ, et al. Performance of AI-enabled electrocardiogram in the prediction of metabolic dysfunction-associated steatotic liver disease. *Clin Gastroenterol Hepatol*. 2024;Epub ahead of print. DOI
56. Sato M, Akamatsu M, Shima T, et al. Impact of a novel digital therapeutics system on nonalcoholic steatohepatitis: the NASH app clinical trial. *Am J Gastroenterol*. 2023;118:1365-72. DOI PubMed PMC
57. Wang K, Margolis S, Cho JM, et al. Non-invasive detection of early-stage fatty liver disease via an on-skin impedance sensor and attention-based deep learning. *Adv Sci*. 2024;11:e2400596. DOI PubMed PMC
58. Ghosh S, Zhao X, Alim M, Brudno M, Bhat M. Artificial intelligence applied to 'omics data in liver disease: towards a personalised approach for diagnosis, prognosis and treatment. *Gut*. 2025;74:295-311. DOI PubMed
59. Fujii I, Matsumoto N, Ogawa M, et al. Artificial intelligence and image analysis-assisted diagnosis for fibrosis stage of metabolic dysfunction-associated steatotic liver disease using ultrasonography: a pilot study. *Diagnostics*. 2024;14:2585. DOI PubMed PMC
60. Santoro S, Khalil M, Abdallah H, et al. Early and accurate diagnosis of steatotic liver by artificial intelligence (AI)-supported ultrasonography. *Eur J Intern Med*. 2024;125:57-66. DOI PubMed
61. Li C, Wang Y, Bai R, et al. Development of fully automated models for staging liver fibrosis using non-contrast MRI and artificial intelligence: a retrospective multicenter study. *EClinMed*. 2024;77:102881. DOI PubMed PMC
62. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. 2023;29:1930-40. DOI PubMed
63. Ge J, Sun S, Owens J, et al. Development of a liver disease-specific large language model chat interface using retrieval-augmented generation. *Hepatology*. 2024;80:1158-68. DOI PubMed PMC
64. Hatem R, Simmons B, Thornton JE. A call to address AI "hallucinations" and how healthcare professionals can mitigate their risks. *Cureus*. 2023;15:e44720. DOI PubMed PMC
65. Novoa J, Chagoyen M, Benito C, Moreno FJ, Pazos F. PMIDigest: interactive review of large collections of pubmed entries to distill relevant information. *Genes*. 2023;14:942. DOI PubMed PMC
66. Kresevic S, Giuffrè M, Ajcevic M, Accardo A, Crocè LS, Shung DL. Optimization of hepatological clinical guidelines interpretation by large language models: a retrieval augmented generation-based framework. *NPJ Digit Med*. 2024;7:102. DOI PubMed PMC
67. der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med*. 2021;27:775-84. DOI PubMed

68. Waqas A, Bui MM, Glassy EF, et al. Revolutionizing digital pathology with the power of generative artificial intelligence and foundation models. *Lab Invest.* 2023;103:100255. DOI PubMed
69. Lu MY, Chen B, Williamson DFK, et al. A visual-language foundation model for computational pathology. *Nat Med.* 2024;30:863-74. DOI PubMed PMC
70. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. *Nat Med.* 2022;28:1773-84. DOI PubMed
71. Narayanan P, Wu T, Shah VH, Curtis BL. Insights into ALD and AUD diagnosis and prognosis: exploring AI and multimodal data streams. *Hepatology.* 2024;80:1480-94. DOI PubMed