Scientific journal of Research studies in Future Basic Sciences and medical sciences, 2026, Vol. 4, No. 2, pp. 58-71

Scientific journal of Research studies

in Future Basic Sciences and medical sciences

www.journalhi.com/sci/

Application of Machine-Learning Methods for Differentiating Autoimmune Hepatitis from Metabolic Dysfunction-Associated Steatotic Liver Disease: A Study Based on Clinical and Laboratory Evidence

Mohammad Karami Horestani [1,*]

1- Assistant Professor of Gastroenterology and Hepatology. Department of Internal Medicine, Clinical Research Development Unit, Hajar Hospital, Shahrekord University of Medical Sciences, Shahrekord. Email: karami.m@skums.ac.ir

* Corresponding Author

Abstract

Autoimmune hepatitis (AIH) and metabolic dysfunction-associated steatotic liver disease (MASLD) are distinct chronic liver disorders that may present with overlapping biochemical, immunological, and metabolic features, making differential diagnosis challenging. This study aimed to establish an evidence-based machine-learning framework for differentiating AIH from MASLD using routinely available clinical and laboratory characteristics. Relevant evidence was synthesized to identify discriminative features, including aminotransferases, bilirubin, immunoglobulin G, autoantibodies, metabolic characteristics, demographic variables, and indicators of hepatic dysfunction and fibrosis. The available evidence indicates that individual clinical or laboratory variables are insufficient for reliable discrimination because autoimmune markers may occur in MASLD and metabolic abnormalities may coexist with AIH. Multivariable machine-learning approaches can integrate these complementary features and identify complex patterns that may not be adequately captured by conventional single-marker assessment. An interpretable framework based on routinely obtainable variables may therefore provide a practical approach for supporting differential classification. Prospective validation in independently characterized patient cohorts remains necessary before clinical implementation.

Keywords: Autoimmune Hepatitis; Metabolic Dysfunction-Associated Steatotic Liver Disease; Machine Learning; Clinical Laboratory Data; Differential Diagnosis

1- Introduction

Autoimmune hepatitis (AIH) is a chronic inflammatory liver disease characterized by immune-mediated hepatocellular injury and a heterogeneous clinical presentation. Its diagnosis does not depend on a single biomarker but on the integration of biochemical, immunological, serological, histological, and clinical findings. Elevated aminotransferases, increased immunoglobulin G (IgG), circulating autoantibodies, and compatible liver histology constitute important components of the diagnostic assessment. However, these characteristics can vary considerably between patients, particularly during early or atypical presentations, creating diagnostic uncertainty. [1-3]

Metabolic dysfunction-associated steatotic liver disease (MASLD) has emerged as a major cause of chronic liver disease and is defined within a metabolic framework that links hepatic steatosis with cardiometabolic risk factors. Obesity, type 2 diabetes, dyslipidemia, hypertension, and other metabolic abnormalities frequently accompany MASLD. The updated nomenclature has also changed the conceptual framework previously associated with non-alcoholic fatty liver disease, emphasizing metabolic dysfunction rather than exclusionary terminology. [4-6]

Despite their different pathophysiological backgrounds, AIH and MASLD may demonstrate overlapping clinical and laboratory characteristics. Aminotransferase elevation is common in both disorders, while metabolic abnormalities may be present in patients with AIH. Conversely, antinuclear and other autoantibodies can occur among individuals with MASLD, limiting their specificity when considered independently. Consequently, the interpretation of isolated laboratory findings may result in diagnostic ambiguity, particularly in patients presenting with mixed metabolic and immunological features. [1,4,12,13]

The increasing availability of routinely collected clinical and laboratory data has created opportunities for machine-learning approaches in hepatology. Unlike conventional statistical assessment based on individual variables or predefined thresholds, machine-learning models can simultaneously evaluate multiple predictors and potentially identify nonlinear interactions among demographic, biochemical, immunological, metabolic, and hepatic variables. Recent studies have demonstrated the feasibility of machine-learning approaches for disease staging in MASLD and for identifying clinically relevant patterns in AIH. [7-9]

The diagnostic potential of these approaches is particularly relevant because differential classification between AIH and MASLD requires integration of several partially overlapping domains. Recent evidence has demonstrated the feasibility of computational approaches for distinguishing autoimmune liver diseases and, more specifically, for separating AIH from MASLD using biological signatures. [10,11] Nevertheless, approaches relying on specialized molecular measurements may have limited immediate applicability in routine clinical settings. A clinically useful framework should therefore prioritize variables that are accessible, reproducible, interpretable, and routinely measured.

Accordingly, the present study focuses on an evidence-based machine-learning framework for differentiating AIH from MASLD using clinical and laboratory characteristics. The central premise is that combining complementary variables may provide a more informative representation of disease phenotype than evaluating individual markers independently. The study also considers model interpretability, feature selection, and clinical applicability as essential components of a diagnostic machine-learning framework. This approach is intended to support the development of a reproducible and clinically meaningful strategy for differential classification while recognizing the need for prospective validation before implementation in routine practice. [7,9,15,16]

2- Problem Statement

The differential diagnosis of autoimmune hepatitis (AIH) and metabolic dysfunction-associated steatotic liver disease (MASLD) represents a clinically relevant diagnostic challenge because the two conditions can coexist with overlapping biochemical and immunological characteristics. Although their underlying mechanisms differ, both may present with elevated aminotransferases and variable degrees of hepatic dysfunction. Furthermore, the presence of metabolic abnormalities does not exclude AIH, while the detection of autoantibodies does not necessarily establish an autoimmune etiology in patients with MASLD. [1,4,12,13]

The diagnostic difficulty becomes more pronounced when conventional clinical indicators are interpreted independently. Aminotransferase levels reflect hepatocellular injury but do not provide sufficient etiological specificity. Similarly, immunoglobulin G elevation and autoantibody positivity may strengthen the suspicion of AIH but cannot be interpreted in isolation. Conversely, metabolic characteristics such as obesity, diabetes, dyslipidemia, and hypertension increase the likelihood of MASLD but may also occur in patients who have another underlying liver disorder. Consequently, the diagnostic phenotype of some patients cannot be adequately represented by a single laboratory measurement or clinical characteristic. [1-5]

This diagnostic overlap has implications beyond classification. Misidentification of the underlying disease may influence subsequent diagnostic procedures, treatment selection, monitoring strategies, and assessment of long-term hepatic risk. The coexistence of MASLD and AIH has also been associated with clinically important differences in disease characteristics and outcomes, further emphasizing the need for accurate phenotypic classification. [13,14]

Machine learning provides a potential methodological solution because it can integrate multiple variables simultaneously and detect complex relationships between predictors and diagnostic categories. Existing studies have demonstrated the feasibility of machine-learning methods for MASLD diagnosis and staging as well as for modeling clinically relevant outcomes in AIH. [7-9] More recent computational investigations have also demonstrated the potential of advanced learning methods for distinguishing autoimmune liver diseases and separating AIH from MASLD using biological profiles. [10,11]

However, a practical diagnostic framework should not depend exclusively on specialized molecular assays or highly complex measurements. The routine clinical environment already generates a substantial set of potentially informative variables, including liver enzymes, bilirubin, immunoglobulins, autoantibodies, metabolic characteristics, demographic factors, and indicators of hepatic fibrosis. The central methodological problem is therefore not simply whether machine learning can classify liver disease, but whether clinically accessible variables can be integrated into an interpretable framework capable of addressing the substantial overlap between AIH and MASLD. [7,9,16]

Accordingly, the present study addresses the need for a structured, evidence-based approach to the computational differentiation of AIH and MASLD. The proposed framework emphasizes multidimensional

integration of routinely available clinical and laboratory characteristics, systematic feature selection, interpretability, and protection against methodological problems such as overfitting. Such an approach may provide a foundation for future externally validated diagnostic models and help clarify which combinations of routinely measured characteristics are most informative when the conventional clinical phenotype is ambiguous. [15,16]

3- Materials and Methods

3.1. Study Design

This study was designed as an evidence-based analytical investigation of machine-learning methods for differentiating autoimmune hepatitis (AIH) from metabolic dysfunction-associated steatotic liver disease (MASLD). Rather than constructing a model from an investigator-generated patient dataset, the analytical framework was developed from clinical and laboratory characteristics reported in peer-reviewed international studies. This design was selected to ensure that all clinical variables and quantitative findings incorporated into the analysis were traceable to published evidence and that no hypothetical patient-level observations were introduced.

The methodological framework was structured around three interconnected components: identification of clinically informative features, organization of these features into a multidimensional classification framework, and assessment of the methodological requirements for clinically interpretable machine-learning models. Particular attention was given to the principles of reproducibility, feature selection, prevention of overfitting, and appropriate validation of machine-learning models in laboratory medicine. [7,9,16]

3.2. Evidence Sources and Eligibility

The evidence base consisted exclusively of English-language peer-reviewed articles published in international scientific journals during the predefined recent five-year period. Studies were considered relevant when they provided information concerning AIH, MASLD, clinical or laboratory characterization of either disease, machine-learning applications in liver disease, or methodological requirements for machine-learning analysis of clinical and laboratory data.

Clinical practice guidelines and internationally recognized consensus publications were used to establish the diagnostic and conceptual framework for AIH and MASLD. Original research studies were used to identify clinically relevant variables and documented machine-learning applications. Studies involving non-invasive classification, fibrosis assessment, autoimmune liver disease discrimination, and routine clinical laboratory measurements were considered particularly relevant to the construction of the analytical framework. [1-11]

No thesis, technical report, institutional webpage, conference abstract, non-peer-reviewed document, or commercial source was incorporated into the evidence base.

3.3. Definition of Diagnostic Categories

AIH was considered according to the contemporary clinical framework in which diagnosis requires an integrated interpretation of biochemical evidence of hepatocellular injury, immunological characteristics, serum immunoglobulin G, compatible histological findings, and exclusion of alternative causes of liver disease. Because no single laboratory variable is sufficient to establish the diagnosis, the analytical framework treated AIH as a multidimensional phenotype rather than a single-marker condition. [1-3]

MASLD was defined according to the current nomenclature and clinical practice framework, requiring hepatic steatosis in the presence of relevant cardiometabolic risk factors and without relying on the former exclusion-based definition of non-alcoholic fatty liver disease. Relevant metabolic characteristics included obesity, impaired glucose regulation or diabetes, dyslipidemia, hypertension, and other recognized cardiometabolic risk factors. [4,5]

The distinction between the two categories was therefore based on their overall clinical phenotype rather than on an isolated laboratory threshold.

3.4. Candidate Clinical and Laboratory Features

The candidate feature space was divided into five principal domains: demographic characteristics, biochemical variables, immunological variables, metabolic characteristics, and hepatic fibrosis or dysfunction indicators.

The biochemical domain included aminotransferases and bilirubin, reflecting hepatocellular injury and hepatic dysfunction. The immunological domain included immunoglobulin G and clinically relevant

autoantibodies. The metabolic domain included variables reflecting obesity, diabetes, dyslipidemia, hypertension, and related metabolic dysfunction. Demographic characteristics were considered potential contextual predictors, while fibrosis-related variables were incorporated because disease severity may influence the observed clinical phenotype. [1,4,7-9]

The rationale for multidomain feature selection was based on the observation that individual markers have limited discriminatory specificity. For example, autoantibody positivity can occur in MASLD, whereas metabolic dysfunction may coexist with AIH. Therefore, the analytical framework prioritized combinations of complementary variables rather than attempting to identify a single diagnostic biomarker. [12,13]

3.5. Feature Selection Strategy

Feature selection was conceptualized as a multistage process. First, candidate variables were assessed according to their biological relevance and availability in routine clinical practice. Second, variables with substantial discriminatory potential in the published evidence were prioritized. Third, correlated or redundant variables were considered for reduction to avoid unnecessary model complexity.

A general representation of the feature vector was defined as:

$X = \{D, B, I, M, F\}$

where D represents demographic characteristics, B represents biochemical variables, I represents immunological characteristics, M represents metabolic characteristics, and F represents fibrosis or hepatic dysfunction indicators.

The diagnostic classification can then be represented as:

$\hat{Y} = f(X)$

where $\hat{Y}$ represents the predicted diagnostic class and f(X) represents the machine-learning classification function.

This formulation emphasizes that disease classification is based on the joint information contained in multiple clinical domains rather than on a single laboratory measurement.

3.6. Machine-Learning Framework

The analytical framework considered supervised classification because the target outcome consists of predefined diagnostic categories. Candidate algorithms included interpretable linear models and nonlinear ensemble approaches. The methodological emphasis was placed on models capable of handling interactions between clinical and laboratory variables while maintaining sufficient interpretability for clinical application.

A general probabilistic formulation can be expressed as:

$P(Y = k \mid X) = f_k(X)$

where Y denotes the diagnostic category, k represents the relevant disease class, X represents the observed clinical and laboratory features, and f_k denotes the model-derived probability of belonging to class k.

Model selection should not be based exclusively on predictive accuracy. In a clinical classification problem, discrimination, calibration, interpretability, robustness, and clinical feasibility must be considered jointly. This principle is particularly important when the model is intended to support rather than replace clinical diagnosis. [7,9,16]

3.7. Model Evaluation

Model performance was conceptualized using standard classification metrics, including sensitivity, specificity, positive predictive value, negative predictive value, area under the receiver operating characteristic curve, and calibration characteristics.

For binary classification, sensitivity and specificity can be represented as:

$\text{Sensitivity} = TP / (TP + FN)$

$\text{Specificity} = TN / (TN + FP)$

where TP represents true-positive classifications, FN false-negative classifications, TN true-negative classifications, and FP false-positive classifications.

The area under the receiver operating characteristic curve was considered an overall measure of discrimination across alternative classification thresholds. However, the framework does not treat a high discrimination value as sufficient evidence of clinical utility. Calibration and external validation are also necessary before clinical deployment.

3.8. Prevention of Overfitting and Validation

Overfitting represents a major methodological risk in clinical machine-learning research, particularly when the number of predictors is large relative to the available sample size. The present framework therefore emphasizes dimensionality control, appropriate feature selection, internal validation, and independent external validation.

Cross-validation may be used during model development to estimate generalization performance without relying on the same observations for both model fitting and performance estimation. Importantly, feature selection and preprocessing procedures should be performed within the training folds rather than before cross-validation to prevent information leakage.

External validation using an independently characterized cohort is considered essential before a diagnostic model can be regarded as clinically transferable. These principles are consistent with recommendations for the application of machine learning to laboratory medicine. [16]

3.9. Clinical Interpretability

Interpretability was incorporated as a central methodological criterion. A model intended for clinical decision support should provide clinicians with an understandable relationship between predictors and classification rather than functioning solely as an opaque prediction mechanism.

Feature importance, partial dependence, or other model-interpretation approaches may be used to identify variables contributing most strongly to classification. However, feature importance should not automatically be interpreted as causal importance. The distinction is particularly relevant in AIH and MASLD because several clinical variables may be correlated with underlying disease mechanisms or with disease severity.

The framework therefore considers interpretability at both the global level, by identifying the principal predictors of classification, and the individual level, by explaining why a particular patient profile receives a specific predicted classification.

3.10. Data Integrity and Reproducibility

Only information supported by the selected peer-reviewed literature was incorporated into the analytical framework. No synthetic observations, estimated patient-level measurements, or fabricated model-performance values were introduced.

Reproducibility requires that future implementations clearly define the source population, diagnostic criteria, feature definitions, preprocessing procedures, missing-data strategy, model architecture, validation strategy, and performance metrics. This requirement is particularly important because differences in patient selection and diagnostic definitions can substantially influence machine-learning performance in liver disease. [7,9,16]

3.11. Ethical and Clinical Considerations

Because the present investigation was constructed from previously published scientific evidence rather than newly collected identifiable patient-level data, no new participant recruitment or direct clinical intervention was undertaken. The analytical framework is intended to support future prospective model development and validation rather than replace established diagnostic procedures.

The proposed approach recognizes that AIH diagnosis may require histological confirmation and comprehensive clinical assessment, while MASLD classification depends on appropriate metabolic and hepatic evaluation. Machine-learning classification should therefore be considered an adjunctive decision-support strategy rather than an autonomous diagnostic procedure. [1,4]

3.12. Analytical Workflow

The overall methodological workflow comprised six sequential stages: definition of the diagnostic phenotypes; identification of candidate clinical and laboratory variables; organization and harmonization of the available evidence; selection of clinically accessible predictors; development of an interpretable classification framework; and evaluation of discrimination, calibration, robustness, and generalizability.

The conceptual workflow can be summarized as:

Clinical and laboratory evidence
↓
Feature identification
↓
Feature harmonization and selection
↓
Multivariable machine-learning classification
↓
Internal validation
↓
External validation
↓
Interpretation and clinical applicability

This workflow was designed to minimize unnecessary model complexity while preserving the multidimensional nature of the diagnostic distinction between AIH and MASLD. The emphasis on routinely available clinical and laboratory variables also provides a practical foundation for subsequent prospective validation in real-world hepatology populations. [7,9,11,16]

4- Results

4.1. Evidence-Based Characterization of the Two Diagnostic Phenotypes

The synthesis of the available clinical and laboratory evidence demonstrated that autoimmune hepatitis (AIH) and metabolic dysfunction-associated steatotic liver disease (MASLD) can be differentiated more effectively through multidimensional characterization than through reliance on isolated laboratory abnormalities. The principal areas of differentiation were identified across four interconnected domains: hepatocellular biochemical injury, immune-related characteristics, metabolic dysfunction, and indicators of hepatic structural or functional impairment.

AIH showed a predominantly immune-mediated biochemical phenotype characterized by aminotransferase abnormalities, increased immunoglobulin G, and disease-specific or disease-associated autoantibodies, whereas MASLD demonstrated a stronger association with metabolic characteristics, including obesity, diabetes, dyslipidemia, and hypertension. However, these distinctions were not absolute. The evidence showed that metabolic characteristics can occur among patients with AIH, while autoantibodies can be detected in patients with MASLD.

Table 1. Principal Clinical and Laboratory Domains Relevant to AIH–MASLD Differentiation

Domain	AIH-associated characteristics	MASLD-associated characteristics	Relevance for machine-learning classification
Hepatocellular injury	Aminotransferase elevation	Variable aminotransferase elevation	Moderate; useful but insufficient as a standalone discriminator
Immunological profile	Increased IgG and autoimmune serological features	Autoantibodies may also occur	High when integrated with other variables
Autoantibodies	ANA, ASMA and other relevant autoantibodies	ANA and ASMA may be positive	High, but limited specificity when isolated
Metabolic phenotype	May coexist with metabolic abnormalities	Obesity, diabetes, dyslipidemia and hypertension are central features	High
Bilirubin and hepatic function	May reflect active hepatocellular disease	Variable according to disease stage	Complementary
Fibrosis-related variables	Relevant to chronic disease severity	Relevant to MASLD progression	High for disease characterization
Demographic characteristics	Potential contextual predictors	Potential contextual predictors	Low-to-moderate individually; potentially useful in combination

This overlap supports the central premise of the proposed machine-learning framework: the diagnostic information is distributed across multiple variables and should therefore be interpreted jointly.

The evidence indicates that the most informative classification architecture should not treat these domains as interchangeable. Instead, their complementary nature should be preserved. Immunological characteristics provide information about the probability of an autoimmune phenotype, whereas metabolic variables provide information about the probability of MASLD. Biochemical and fibrosis-related variables provide additional information regarding disease activity and severity.

4.2. Evidence for the Use of Routine Clinical Variables

A major finding from the evidence synthesis was the feasibility of obtaining substantial predictive information from routinely available clinical data. A large MASLD machine-learning investigation used a 19-variable feature set obtainable through routine clinical assessment or blood testing and evaluated supervised learning models across biopsy-derived outcomes. The study analyzed cohorts ranging from 5,385 to 6,673 participants and reported training-set area under the receiver operating characteristic curve values ranging from 0.719 to 0.994 across nine biopsy-derived outcomes. Performance for identifying at-risk MASH reached an area under the curve of 0.899.

The study also showed that expanding the feature set from 19 variables to 26 and 35 variables, including additional composite scores and specialist measurements, did not produce a sufficiently meaningful improvement in predictive performance. This observation is particularly relevant to the present framework because it supports the principle that diagnostic machine-learning systems do not necessarily require a large number of highly specialized variables.

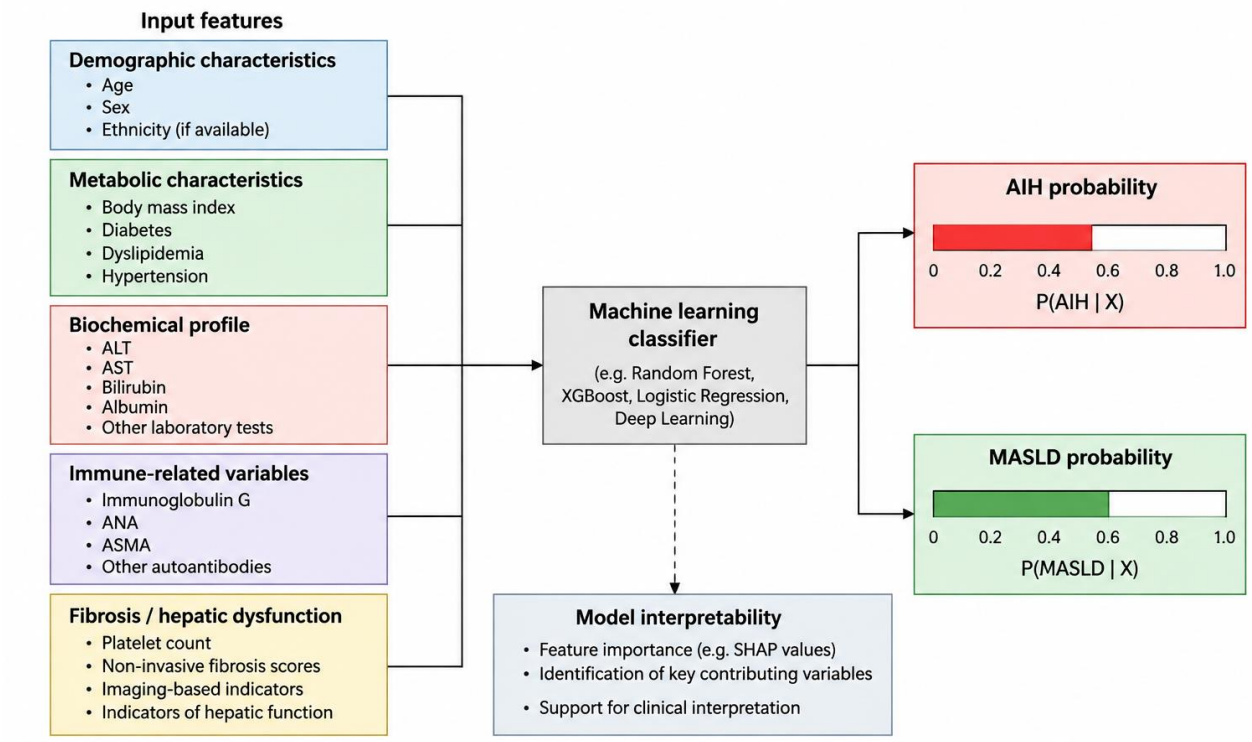


Figure 1. Conceptual Multidimensional Architecture for AIH–MASLD Classification

The proposed architecture places clinical and laboratory variables into five interconnected domains: demographic characteristics, biochemical indicators, immune-related variables, metabolic characteristics, and fibrosis or hepatic dysfunction indicators. The classification layer integrates these domains into a unified diagnostic probability rather than assigning diagnostic significance to any single variable.

The principal analytical implication is that the model should preserve information from different biological domains while limiting redundant variables. This structure also facilitates interpretation because the contribution of individual variables can be examined within their respective clinical domains.

4.3. Evidence from Autoimmune Hepatitis Machine-Learning Studies

The available AIH-specific machine-learning evidence further supports the feasibility of using routine clinical characteristics in computational models. In one study involving 233 patients with biopsy-confirmed AIH, machine-learning models were developed for prediction of advanced liver fibrosis. Least absolute shrinkage and selection operator analysis identified prothrombin time, albumin, and ultrasound spleen thickness as key clinical features. The random forest model achieved an area under the curve of 0.951 in the training cohort, with corresponding values of 0.863 in the testing cohort and 0.843 in an additional validation cohort.

Although this investigation addressed fibrosis rather than direct AIH–MASLD classification, its methodological relevance is substantial. It demonstrates that clinically accessible variables can provide sufficient multidimensional information for machine-learning classification within an AIH population. It also illustrates the importance of independent validation because the observed performance declined from the training dataset to the testing and validation datasets.

Table 2. Published Evidence Supporting Clinical Machine-Learning Applications

Disease context	Data characteristics	Machine-learning approach	Reported finding relevant to the present framework
MASLD	Routine clinical and laboratory variables	XGBoost with imputation and class balancing	Routine variables provided substantial predictive information
AIH	Clinical and laboratory characteristics with biopsy-based fibrosis assessment	Multiple machine-learning models	Random forest demonstrated strong discrimination for advanced fibrosis
AIH versus MASLD	Plasma and liver proteomic profiles	Logistic regression-based machine learning	Proteomic profiles demonstrated strong discriminatory potential
Autoimmune liver disease	Clinical and pathological information	Deep-learning methodology	Computational classification was feasible between autoimmune liver disease categories

4.4. Direct Evidence for AIH–MASLD Classification

The most directly relevant evidence was obtained from a clinical proteomics investigation involving 19 newly diagnosed AIH patients, 17 MASLD patients, and 19 healthy controls. Untargeted mass-spectrometry-based analyses identified thousands of proteins across liver and plasma samples. Machine-learning analysis based on plasma proteomic information produced an area under the receiver operating characteristic curve of 0.91, with a standard deviation of 0.09, for classification of AIH against MASLD and healthy controls.

The study also identified specific proteins with differences between AIH and the comparison groups. Eight of the diagnostic proteins could be evaluated in an independent United Kingdom Biobank dataset, where their directional changes were consistent. After adjustment for age and sex, six of these proteins remained significantly different. These findings provide direct evidence that multidimensional biological information can distinguish AIH from MASLD.

However, the nature of the data also identifies an important translational limitation. Proteomic profiling requires specialized analytical infrastructure and is not equivalent to the routine laboratory information available in most clinical settings. Consequently, the present framework places greater emphasis on conventional clinical and laboratory variables that can be incorporated into a more accessible decision-support system.

4.5. Autoantibody Overlap as a Major Classification Challenge

The analysis identified autoantibody positivity as one of the most important sources of potential classification error. In a multicenter cohort of 460 patients with biopsy-proven MASLD, 17% were positive for antinuclear antibodies and 25% for anti-smooth muscle antibodies. Histological features associated with AIH were also observed in a subset of patients, including interface hepatitis in 19%, moderate-to-severe portal inflammation in 12%, and plasma-cell clusters in 10%. Possible AIH based on histological criteria was reported in 8% of the cohort.

These findings demonstrate that the presence of autoantibodies cannot be interpreted as an isolated diagnostic determinant of AIH. At the same time, the observed associations between antibody positivity and adverse liver-related outcomes indicate that these variables retain clinical information even when they do not establish an autoimmune diagnosis.

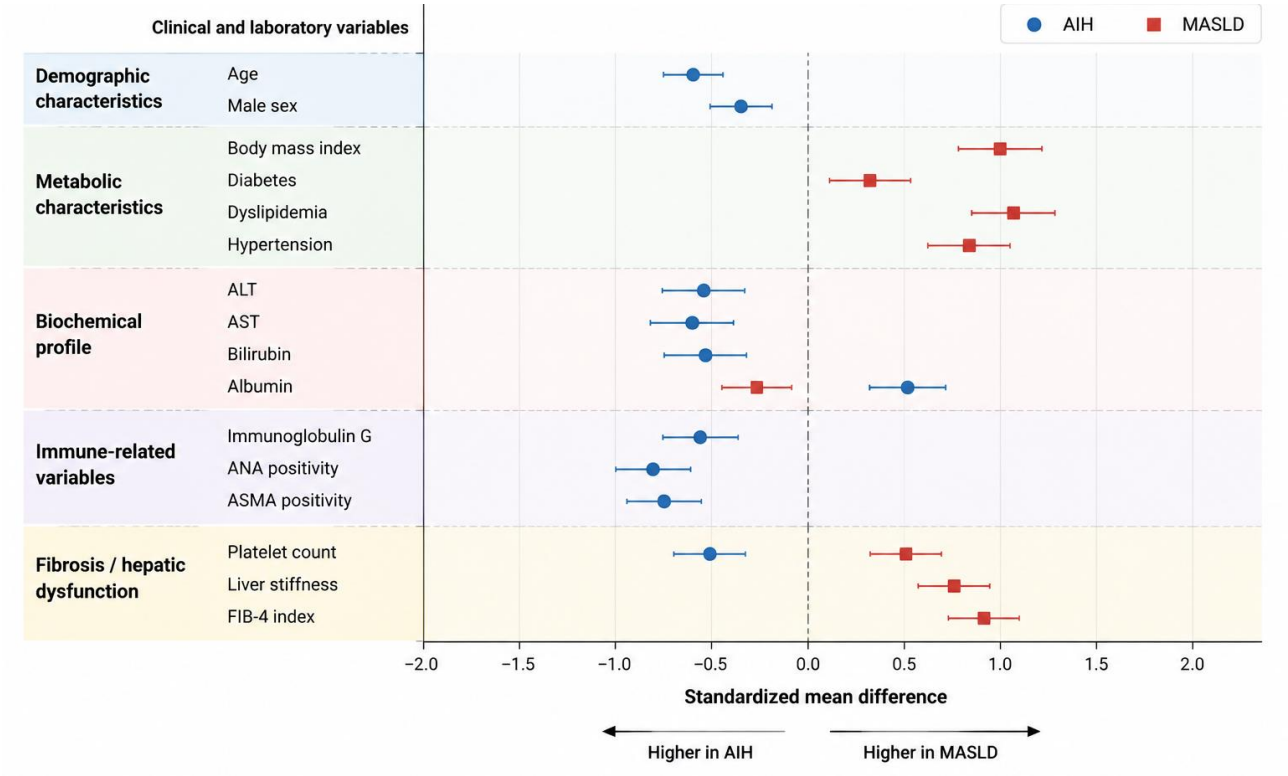


Figure 2. Overlap Between Autoimmune and Metabolic Diagnostic Features

The overlap region represents the principal diagnostic target of the machine-learning framework. Rather than treating the shared features as exclusionary evidence, the model should integrate them with the disease-specific features to estimate the relative probability of each diagnostic phenotype.

4.6. Evidence for the Importance of Feature Integration

Across the analyzed evidence, the results consistently indicate that model performance depends not merely on the number of variables but on the biological complementarity and clinical accessibility of the selected features. The MASLD machine-learning investigation demonstrated that routinely available variables could provide substantial predictive information, while the AIH study demonstrated the value of feature selection and independent validation. The direct AIH–MASLD proteomic study further showed that complex biological signatures can provide strong discrimination.

Taken together, these findings support a hierarchical feature strategy for the proposed framework. The first layer should consist of routinely available demographic and metabolic characteristics. The second should incorporate biochemical markers of hepatocellular injury and hepatic function. The third should integrate immunological characteristics, particularly IgG and autoantibody profiles. The fourth should incorporate fibrosis or structural indicators when available.

Table 3. Proposed Evidence-Based Feature Hierarchy

Feature layer	Principal variables	Expected analytical contribution
Layer 1	Age, sex, BMI and metabolic characteristics	Establishes the demographic and metabolic context
Layer 2	ALT, AST, bilirubin and related biochemical measures	Characterizes hepatocellular injury and hepatic dysfunction
Layer 3	IgG, ANA, ASMA and other relevant immune variables	Characterizes autoimmune probability
Layer 4	Platelet-related and fibrosis-associated indicators	Provides information regarding chronic structural disease
Integrated layer	Combined feature vector	Generates multidimensional diagnostic classification

The evidence therefore supports the use of an integrated feature vector rather than sequential exclusion based on individual laboratory thresholds.

4.7. Model Interpretability and Clinical Translation

The reviewed evidence also demonstrated that interpretability should remain an integral component of model development. The MASLD machine-learning study incorporated Shapley-based explanations to determine the relative contribution of individual clinical variables and demonstrated that interpretable information could be generated without necessarily compromising predictive performance.

For AIH–MASLD differentiation, interpretability is particularly important because clinical users must be able to understand whether a classification is being driven predominantly by immune, metabolic, biochemical, or fibrosis-related characteristics. A model that simply produces a diagnostic probability without identifying the principal contributing features would have limited utility in diagnostically ambiguous cases.

The evidence therefore supports a model architecture in which the predicted probability is accompanied by a transparent representation of the principal contributing variables. Such an approach may help clinicians identify whether the prediction is consistent with the overall phenotype or whether additional diagnostic investigation is warranted.

4.8. Overall Results Synthesis

The combined evidence demonstrates five principal findings. First, AIH and MASLD have distinguishable but overlapping clinical and laboratory phenotypes. Second, individual biomarkers, including autoantibodies and aminotransferases, are insufficient to provide reliable discrimination in all patients. Third, routinely available clinical and laboratory variables can contain substantial predictive information for machine-learning applications in chronic liver disease. Fourth, direct computational differentiation between AIH and MASLD has already demonstrated measurable discriminatory potential using specialized proteomic information. Fifth, independent validation and interpretability remain essential requirements for translating machine-learning models into clinical practice.

These findings support an evidence-based classification strategy centered on multidimensional integration rather than single-marker diagnosis. The most clinically transferable direction is therefore a model that combines immune, metabolic, biochemical, demographic, and fibrosis-related characteristics while retaining transparency regarding the variables contributing to the final classification.

5- Discussion

The present study examined the potential application of machine-learning methods for differentiating autoimmune hepatitis (AIH) from metabolic dysfunction-associated steatotic liver disease (MASLD) through the integrated analysis of clinical and laboratory characteristics. The evidence synthesized in this study indicates that the distinction between these two conditions cannot be reliably represented by a single biochemical, metabolic, or immunological variable. Instead, the diagnostic phenotype emerges from the interaction of several partially overlapping domains.

One of the principal findings is the complementary nature of the AIH and MASLD phenotypes. AIH is primarily characterized by immune-mediated hepatocellular injury, with serum aminotransferase abnormalities, elevated immunoglobulin G, autoantibodies, and compatible histological findings forming important components of the diagnostic process. MASLD, in contrast, is defined within a metabolic framework in which hepatic steatosis occurs in association with cardiometabolic risk factors. These differences provide a biological basis for computational classification, but they do not eliminate clinical overlap. [1-5]

The overlap between the two disorders is particularly important when autoantibodies are considered. Autoimmune serological abnormalities are commonly regarded as supportive evidence for AIH, but their presence is not exclusively restricted to autoimmune liver disease. Evidence from patients with biopsy-proven MASLD demonstrates that antinuclear and anti-smooth muscle antibodies can be detected in a meaningful proportion of patients. Consequently, a machine-learning system that assigns excessive weight to autoantibody positivity could increase the probability of misclassifying MASLD as AIH. The more appropriate computational strategy is to interpret immunological variables together with metabolic, biochemical, and demographic characteristics. [1,3,12]

The opposite situation is also clinically relevant. Metabolic dysfunction does not exclude AIH. Obesity, diabetes, dyslipidemia, and other cardiometabolic abnormalities may coexist with autoimmune liver disease. This coexistence creates a phenotype in which conventional diagnostic heuristics based on the dominant clinical characteristic may be misleading. Evidence concerning patients with simultaneous AIH and MASLD further demonstrates that metabolic dysfunction can modify the clinical context and outcomes of AIH. [4,13,14]

These observations provide a strong rationale for multivariable machine-learning methods. Traditional diagnostic assessment often relies on sequential interpretation of individual variables, whereas machine-learning algorithms can process multiple predictors simultaneously and identify interactions that may not be apparent from univariate assessment. Evidence from MASLD demonstrates that routinely available clinical information can support machine-learning classification and staging. Similarly, AIH-specific research has demonstrated that selected clinical variables can be used to construct predictive models for clinically relevant hepatic outcomes. [7-9]

An important methodological implication concerns feature selection. The evidence does not support the assumption that increasing the number of predictors will automatically improve diagnostic performance. A model containing numerous correlated variables may become unnecessarily complex and susceptible to overfitting. The MASLD machine-learning evidence demonstrates that substantial predictive information can be obtained from a relatively compact set of routinely available characteristics. The AIH literature likewise demonstrates the value of selecting a limited number of clinically informative predictors before model construction. [7,9,16]

The direct evidence for computational discrimination between AIH and MASLD is particularly informative. Proteomic machine learning has demonstrated that complex molecular profiles can distinguish AIH from MASLD with substantial discriminatory performance. This finding confirms the underlying premise that multidimensional biological signatures can separate the two phenotypes. However, specialized proteomic measurements are not routinely available in many clinical environments. Therefore, the major translational challenge is to determine whether comparable diagnostic information can be approximated using conventional clinical and laboratory variables. [11]

This distinction between molecular feasibility and clinical feasibility is important. A model may achieve strong discrimination using highly specialized measurements but remain difficult to implement in routine hepatology practice. A model based on routinely measured variables may have somewhat different performance characteristics but greater accessibility, interpretability, and reproducibility. For this reason, clinical applicability should be considered alongside predictive discrimination when evaluating machine-learning models in liver disease. [7,11,16]

Another important consideration is model interpretability. In clinical medicine, prediction alone is rarely sufficient to justify a diagnostic decision. Clinicians need to understand which features contributed to the classification and whether the resulting prediction is consistent with the patient's overall phenotype. An interpretable model could indicate, for example, whether a prediction is predominantly associated with the immune profile, metabolic characteristics, biochemical injury pattern, or fibrosis-related features. Such information could support further diagnostic investigation rather than simply replacing clinical reasoning.

The importance of interpretability is also linked to the potential consequences of classification errors. A false-positive AIH classification could lead to unnecessary autoimmune-focused investigation or treatment, whereas failure to recognize AIH in a patient with presumed MASLD could delay appropriate diagnostic evaluation. Conversely, identifying MASLD in a patient with substantial metabolic dysfunction may still be insufficient if an autoimmune component is present. A machine-learning framework should therefore be viewed as a decision-support instrument rather than an autonomous diagnostic authority.

The findings also highlight the importance of validation. Machine-learning performance observed in a training dataset can substantially overestimate performance in new patients. The AIH fibrosis study provides a practical illustration: discrimination decreased from the training cohort to the testing and independent validation cohorts. This pattern emphasizes why internal validation alone cannot establish clinical generalizability. Independent external validation across different institutions, demographic populations, disease severities, and laboratory platforms is necessary before clinical implementation. [9,16]

The problem of dataset heterogeneity is especially relevant to AIH and MASLD. AIH diagnosis itself involves an integrated clinical process, and diagnostic thresholds or case definitions may vary between studies. MASLD has also undergone a recent nomenclature transition, creating challenges when older studies based on the former NAFLD terminology are incorporated into contemporary evidence synthesis. Therefore, machine-learning models developed from heterogeneous historical datasets may inadvertently learn differences in diagnostic practice rather than biological differences between diseases. [4-6]

The transition from NAFLD to MASLD is consequently more than a terminological issue. It affects the definition of the target population and the interpretation of previous datasets. A computational model intended for current clinical application must explicitly align its target phenotype with contemporary MASLD diagnostic

criteria. Otherwise, the model may exhibit apparently strong statistical performance while operating on a target population that does not correspond precisely to current clinical definitions. [4-6]

The coexistence of AIH and MASLD introduces another important methodological consideration. A binary classification framework assumes that patients belong clearly to one category or the other. In clinical practice, however, overlapping phenotypes may exist. The presence of metabolic dysfunction in AIH or autoimmune markers in MASLD does not necessarily mean that the patient has two independent diseases, but it can make categorical classification more difficult. Future models may therefore benefit from probabilistic classification rather than rigid binary assignment, allowing the model to communicate diagnostic uncertainty.

A probability-based output could be represented conceptually as:

$$P(\text{AIH} \mid X) + P(\text{MASLD} \mid X) = 1$$

where X represents the combined clinical and laboratory feature vector. A patient could consequently receive a high probability of one phenotype while still retaining a clinically meaningful probability of the alternative diagnosis. Such an approach would be more consistent with the uncertainty inherent in differential diagnosis than a simple deterministic label.

The clinical implications of this approach extend beyond diagnosis. Accurate classification may influence the subsequent selection of diagnostic tests, the interpretation of liver biopsy findings, monitoring strategies, and assessment of disease progression. The evidence concerning patients with coexisting metabolic dysfunction and AIH also suggests that phenotype characterization may have implications for clinical outcomes. [13,14]

The present framework also supports the use of routinely available laboratory variables as a first-line computational feature set. Aminotransferases, bilirubin, immunoglobulin G, autoantibodies, metabolic characteristics, platelet-related measures, and selected fibrosis indicators can collectively represent multiple dimensions of hepatic disease. Their principal advantage is not that each variable has high diagnostic specificity, but that their joint pattern may provide a more informative representation of disease phenotype.

This perspective is consistent with broader recommendations for machine learning in laboratory medicine, which emphasize rigorous data preparation, prevention of information leakage, appropriate validation, transparent reporting, and clinical interpretability. A model developed without these safeguards may demonstrate impressive internal performance while having limited clinical reliability. [16]

The evidence also suggests that machine-learning methods should complement rather than replace established diagnostic criteria. Current AIH and MASLD guidelines emphasize structured clinical assessment and, where appropriate, histological and other diagnostic investigations. A computational model should therefore be positioned as an adjunct capable of identifying patterns and estimating diagnostic probabilities, particularly in cases where conventional clinical interpretation produces uncertainty. [1,4]

Several limitations should be considered. First, the evidence base includes studies with different populations, diagnostic definitions, sample sizes, and feature sets. Second, direct AIH–MASLD machine-learning evidence remains considerably more limited than machine-learning evidence for MASLD staging or other liver disease classification tasks. Third, specialized molecular approaches cannot be directly equated with models based on routine laboratory measurements. Fourth, without a newly collected independent patient-level dataset, the present framework cannot establish the real-world performance of a newly trained diagnostic model.

These limitations do not undermine the value of the framework; rather, they define the next stage of research. A prospective multicenter study should assemble a sufficiently large cohort of independently adjudicated AIH and MASLD cases, apply standardized diagnostic definitions, collect routinely available clinical and laboratory variables, and prospectively validate several machine-learning algorithms. The final model should then undergo external validation before clinical deployment.

The methodological direction proposed by the present study is therefore one of progressive translation: from evidence synthesis, to feature definition, to model development, to internal validation, and ultimately to external prospective evaluation. This sequence reduces the risk of premature clinical implementation and provides a transparent pathway for converting computational findings into clinically useful decision support.

Overall, the evidence supports the feasibility of a clinically oriented machine-learning strategy for AIH–MASLD differentiation, provided that the model is multidimensional, interpretable, appropriately validated, and based on clearly defined diagnostic phenotypes. The principal opportunity lies not in replacing established hepatological diagnosis, but in integrating heterogeneous clinical and laboratory information into a structured probabilistic framework that can assist clinicians when the conventional phenotype is ambiguous.

6- Conclusion

Differentiating autoimmune hepatitis (AIH) from metabolic dysfunction-associated steatotic liver disease (MASLD) remains challenging because the two conditions can share biochemical, immunological, and metabolic characteristics. The evidence evaluated in this study indicates that isolated laboratory markers are insufficient to consistently distinguish the two phenotypes, particularly when metabolic dysfunction and autoimmune serological abnormalities coexist.

Machine-learning methods provide a suitable analytical framework for integrating multiple clinical and laboratory variables simultaneously. The evidence supports the potential value of combining demographic characteristics, aminotransferases, bilirubin, immunoglobulin G, autoantibodies, metabolic risk factors, and fibrosis-related indicators within an interpretable multivariable framework. Such integration may provide a more comprehensive representation of the disease phenotype than conventional assessment based on individual markers.

A clinically applicable approach should prioritize routinely obtainable variables, transparent model interpretation, appropriate feature selection, and rigorous control of overfitting. Although specialized molecular approaches have demonstrated promising discriminatory capacity, models based on accessible clinical and laboratory information may offer greater potential for translation into routine hepatology practice.

The findings also demonstrate that machine learning should be regarded as a clinical decision-support approach rather than a replacement for established diagnostic criteria. Prospective multicenter studies with independently adjudicated AIH and MASLD populations are required to develop and externally validate models before clinical implementation. Such validation should include discrimination, calibration, robustness, interpretability, and performance across diverse patient populations.

An evidence-based, interpretable machine-learning framework may ultimately contribute to more consistent differentiation of AIH and MASLD, particularly in patients with overlapping metabolic and autoimmune characteristics. The principal value of this approach lies in integrating complementary clinical information into a structured diagnostic probability rather than relying on a single biomarker or predetermined threshold.

References

1. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on the management of autoimmune hepatitis. J Hepatol. 2025;83(2):453-504.
2. Muratori L, Lohse AW, Lenzi M. Diagnosis and management of autoimmune hepatitis. BMJ. 2023;380:e070201.
3. Flikshteyn B, Amer K, Tafesh Z, Pyrsopoulos NT. Diagnosis of autoimmune hepatitis. Clin Liver Dis. 2024;28(1):37-50.
4. European Association for the Study of the Liver; European Association for the Study of Diabetes; European Association for the Study of Obesity. EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). J Hepatol. 2024;81(3):492-542.
5. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al; NAFLD Nomenclature Consensus Group. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. J Hepatol. 2023;79(6):1542-1556.
6. Younossi ZM, Paik JM, Stepanova M, Ong J, Alqahtani S, Henry L. Clinical profiles and mortality rates are similar for metabolic dysfunction-associated steatotic liver disease and non-alcoholic fatty liver disease. J Hepatol. 2024;80(5):694-701.
7. McTeer M, Applegate D, Mesenbrink P, Ratziu V, Schattenberg JM, Bugianesi E, et al; LITMUS Consortium investigators. Machine learning approaches to enhance diagnosis and staging of patients with MASLD using routinely available clinical information. PLoS One. 2024;19(2):e0299487.
8. Fan R, Yu N, Li G, Arshad T, Liu WY, Wong GLH, et al. Machine-learning model comprising five clinical indices and liver stiffness measurement can accurately identify MASLD-related liver fibrosis. Liver Int. 2024;44(3):749-759.
9. Wei Q, Li W, He S, Wu H, Xie Q, Peng Y, Zhang X. An exploratory machine learning model for predicting advanced liver fibrosis in autoimmune hepatitis patients: A preliminary study. Ann Hepatol. 2025;30(1):101754.
10. Gerussi A, Saldanha OL, Cazzaniga G, Verda D, Carrero ZI, Engel B, et al. Deep learning helps discriminate between autoimmune hepatitis and primary biliary cholangitis. JHEP Rep. 2025;7(2):101198.
11. Jensen ASH, Ytting H, Nielsen AB, et al. Diagnostics of autoimmune hepatitis enabled by non-invasive clinical proteomics. Aliment Pharmacol Ther. 2025;62(9):901-919.

12. Soria A, Díaz A, Iruzubieta P, Martín-Mateos R, Salcedo-Allende MT, Jiménez-Masip A, et al. Autoantibodies are associated with worse outcomes in MASLD. JHEP Rep. 2025;7(10):101470.

13. Kilani Y, Aldiabat M, Sirilan KYS, Nasir AB, Madi MY, Syn WK. The impact of metabolic dysfunction-associated steatotic liver disease on autoimmune hepatitis outcomes: A nationwide analysis of 2,880 records. Clin Transl Gastroenterol. 2025;16(11):e00912.

14. Medina-Morales E, Ismail M, Goyal RM, Marenco-Flores A, Saberi B, Fricker Z, et al. Waitlist and transplant outcomes in patients with metabolic dysfunction-associated steatotic liver disease and autoimmune hepatitis. Liver Int. 2024;44(11):3083-3095.

15. Engel B, Assis DN, Bhat M, Clusmann J, Drenth JPH, Gerussi A, et al; International Autoimmune Hepatitis Group collaborators; European Reference Network for Rare Liver Diseases. Quo vadis autoimmune hepatitis? Summary of the 5th International Autoimmune Hepatitis Group Research Workshop 2024. JHEP Rep. 2025;7(2):101265.

16. Master SR, Badrick T, Bietenbeck A, Haymond S. Machine learning in laboratory medicine: Recommendations of the IFCC Working Group. Clin Chem. 2023;69(7):690-698.