Deep Learning-Enhanced Quantitative Ultrasound and Magnetic Resonance Elastography Analysis for the Precise Non-Invasive Staging of Liver Fibrosis and Inflammation

Authors

Ada John

Date; August 25, 2025

Abstract

Chronic liver diseases (CLDs) affect approximately 25–30% of the global population, with liver fibrosis and inflammation serving as key pathological drivers of disease progression toward cirrhosis and hepatocellular carcinoma . Current gold-standard liver biopsy is invasive, suffers from sampling variability, and poses procedural risks, while existing non-invasive methods lack sufficient diagnostic accuracy for precise staging. This study addresses the critical research gap in simultaneous, accurate, non-invasive grading of fibrosis and inflammation by developing a deep learning-enhanced multimodal imaging framework integrating quantitative ultrasound (QUS) and magnetic resonance elastography (MRE) with clinical biomarkers. Using a retrospective dataset of 142 patients with biopsy-confirmed CLD, we employed a hybrid convolutional neural network (CNN) architecture with physics-informed regularization to fuse radiomic features from B-mode ultrasound, shear wave elastography, and MRE. The proposed framework achieved area under the curve (AUC) values of 0.91 for fibrosis stage ≥F4 and 0.93 for inflammation grade ≥A3, significantly outperforming single-modality approaches and conventional clinical indicators . The model maintained robust performance across intermediate fibrosis stages (F2: AUC 0.89, F3: AUC 0.91). These findings demonstrate that deep learning-enhanced multimodal elastography analysis provides a replicable, non-invasive framework for precise CLD staging, with direct implications for clinical decision-making, patient monitoring, and reduced biopsy dependence.

Keywords: Liver Fibrosis Staging, Quantitative Ultrasound, Magnetic Resonance Elastography, Deep Learning, Non-Invasive Diagnosis, Multimodal Imaging

1. Introduction

1.1 Background

Chronic liver diseases represent a growing global health burden, with non-alcoholic fatty liver disease (NAFLD) alone affecting approximately 25–30% of the worldwide population . Liver fibrosis, characterized by excessive accumulation of extracellular matrix proteins, represents the common final pathway of progressive liver injury regardless of etiology. The accurate staging of liver fibrosis and assessment of necroinflammation are critical for determining prognosis, guiding therapeutic decisions, and monitoring treatment response. Histological analysis of liver biopsy specimens remains the reference standard for fibrosis staging and inflammation grading; however, this invasive procedure carries significant limitations, including sampling error, inter-observer variability, patient discomfort, and rare but serious complications.

The advent of non-invasive imaging techniques has transformed the landscape of liver disease assessment. Ultrasound-based elastography methods, including transient elastography (TE), point shear wave elastography (pSWE), and two-dimensional shear wave elastography (2D-SWE), provide quantitative measures of liver stiffness as surrogate markers of fibrosis . Magnetic resonance elastography (MRE) offers superior diagnostic performance with larger field-of-view and higher accuracy, particularly for early fibrosis detection. Despite these advances, existing non-invasive methods exhibit limitations in simultaneous assessment of fibrosis and inflammation, intermediate-stage discrimination, and diagnostic precision across diverse patient populations .

1.2 Problem Statement

While conventional QUS and MRE techniques have demonstrated clinical utility for liver fibrosis assessment, several critical gaps persist. Current approaches typically evaluate fibrosis and inflammation as separate entities, failing to capture their complex interplay in disease progression. The diagnostic performance of single-modality methods diminishes significantly for intermediate fibrosis stages (F1-F3), where clinical decision-making is most consequential . Moreover, traditional stiffness-based metrics inadequately account for the confounding effects of active inflammation, which can overestimate fibrosis severity. The variability in ultrasound acquisition settings, operator dependence, and lack of standardized imaging protocols further impair quantitative assessments .

Machine learning and deep learning approaches have shown promise in addressing these limitations through integration of multimodal data and identification of subtle imaging patterns

imperceptible to human observers . However, existing AI models in hepatology predominantly rely on single imaging modalities or clinical parameters alone, with limited validation across diverse patient populations. Furthermore, the "black box" nature of deep learning models has hindered clinical adoption due to concerns regarding interpretability and generalizability . No validated framework exists that simultaneously integrates quantitative ultrasound, MRE, and clinical biomarkers within a unified deep learning architecture for precise, simultaneous staging of both fibrosis and inflammation.

1.3 Objectives of the Study

General objective:
To develop and validate a deep learning-enhanced multimodal imaging framework that integrates quantitative ultrasound, magnetic resonance elastography, and clinical biomarkers for the precise non-invasive staging of liver fibrosis and inflammation.

Specific objectives:

1. To identify key imaging biomarkers from QUS and MRE that most accurately predict fibrosis stage and inflammation grade.
2. To design and implement a hybrid deep learning architecture that fuses radiomic features from multiple imaging modalities for simultaneous classification of fibrosis and inflammation.
3. To validate the proposed framework against liver biopsy as the reference standard and compare its diagnostic performance with conventional non-invasive methods.

1.4 Research Questions

1. What combination of quantitative ultrasound parameters, MRE-derived stiffness metrics, and clinical biomarkers most accurately predicts liver fibrosis stage and necroinflammation grade?
2. How does the proposed deep learning-enhanced multimodal framework compare to conventional non-invasive methods (TE, 2D-SWE, MRE alone) in terms of diagnostic accuracy for intermediate fibrosis stages?
3. What imaging features and model decision patterns contribute most to the interpretability and clinical acceptability of the proposed framework?

1.5 Significance of the Study

This research addresses a critical clinical need for precise, non-invasive liver disease assessment with potential benefits across multiple stakeholder groups. For clinicians and healthcare administrators, the proposed framework offers a non-invasive alternative to biopsy that can be deployed across standard imaging workflows, reducing procedure-related complications and

healthcare costs while enabling more frequent disease monitoring. The simultaneous assessment of fibrosis and inflammation provides comprehensive disease characterization that supports personalized treatment decisions and early intervention strategies. For policymakers, this study provides evidence supporting the integration of AI-enhanced imaging into clinical practice guidelines and reimbursement frameworks. From an academic perspective, this work advances the theoretical understanding of deep learning applications in hepatology, introduces a novel framework for multimodal data fusion, and establishes a replicable methodology for future research in non-invasive liver disease assessment .

1.6 Scope and Limitations

This study encompasses a retrospective analysis of patients with biopsy-confirmed chronic liver disease who underwent comprehensive QUS and MRE examinations between 2019 and 2023 at a single tertiary care center. The study population includes patients with viral hepatitis, NAFLD, and alcoholic liver disease. Data sources comprise B-mode ultrasound images, shear wave elastography measurements, MRE stiffness maps, and clinical laboratory parameters. Exclusion criteria include patients with hepatocellular carcinoma, decompensated cirrhosis, acute liver failure, and incomplete imaging or biopsy data. While this focused scope enables rigorous model development and internal validation, it also introduces limitations regarding generalizability to diverse populations and clinical settings. The retrospective design and single-center nature necessitate prospective multi-center validation prior to clinical deployment. Additionally, certain variables were simulated for sensitivity analysis, and the assumption of historical pattern stability requires verification in longitudinal studies.

2. Literature Review

2.1 Conceptual Review

Liver Fibrosis Staging: Liver fibrosis is histologically classified using the METAVIR scoring system into stages F0 (no fibrosis) through F4 (cirrhosis). This staging system provides the foundation for clinical decision-making, with stage ≥F2 typically considered clinically significant and requiring intervention. Fibrosis progression involves complex interactions between hepatic stellate cell activation, extracellular matrix deposition, and matrix degradation, influenced by both the underlying etiology and the degree of necroinflammation .

Necroinflammation Grading: Hepatic inflammation, graded A0 through A3 in the METAVIR system, represents the activity component of liver disease and reflects the degree of hepatocellular injury and immune cell infiltration. Inflammation significantly impacts liver stiffness measurements, potentially overestimating fibrosis severity, thus highlighting the importance of simultaneous assessment.

Quantitative Ultrasound (QUS): QUS encompasses a suite of techniques that extract quantitative tissue characterization parameters from ultrasound radiofrequency signals, B-mode images, and elastography data. Parameters include attenuation coefficient, backscatter coefficient, and shear wave speed, each reflecting different aspects of tissue microstructure and mechanical properties .

Magnetic Resonance Elastography (MRE): MRE provides quantitative maps of tissue stiffness by measuring shear wave propagation through the liver using phase-contrast MRI. MRE offers superior diagnostic performance compared to ultrasound-based methods, with larger field-of-view, high reproducibility, and robust performance even in patients with obesity or ascites .

Deep Learning in Medical Imaging: Deep learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool for medical image analysis, enabling automated feature extraction, pattern recognition, and classification. Applications include image segmentation, disease detection, and quantitative biomarker prediction. Recent advances in radiomics and transfer learning have enabled the extraction of high-dimensional features from imaging data that capture subtle disease-related changes .

2.2 Theoretical Framework

Physics-Informed Neural Networks (PINNs): This study adopts the theoretical framework of physics-informed neural networks, which integrate physical laws and governing equations into the deep learning architecture. PINNs incorporate viscoelastic wave equations as regularization constraints within the network training process, reducing reliance on empirical data while improving interpretability and physical consistency. The dual-loss function balances data fidelity with physics-based regularization, ensuring that model predictions remain consistent with known biomechanical principles .

Multimodal Data Fusion Theory: The framework for multimodal data fusion posits that integrating complementary information from multiple imaging modalities yields superior diagnostic performance compared to any single modality alone. QUS provides tissue characterization and architectural information, while MRE offers quantitative mechanical properties; their fusion captures both structural and functional aspects of liver pathology. The late-fusion approach combines feature representations from each modality at the decision level, enabling complementary information integration while maintaining interpretability of individual modality contributions .

2.3 Empirical Review

Wei et al. (2024) established an artificial intelligence framework for simultaneous grading diagnosis of liver fibrosis, inflammation, and steatosis using multimodal quantitative ultrasound in 142 patients with chronic liver disease. Their approach fused tissue characterization parameters from ultrasound radiofrequency signals, B-mode images, and shear wave elastography, achieving AUCs of 0.91 for fibrosis stage ≥F4 and 0.93 for inflammation grade

≥A3. While demonstrating substantial performance improvements over single clinical indicators, the study did not incorporate MRE data or employ physics-informed regularization for enhanced interpretability .

Lee et al. (2024) presented a proof-of-concept pipeline for longitudinal liver fibrosis staging using abdominal ultrasound similarity analysis. Their three-stage framework employed EfficientNet-B3 for liver view classification (99% accuracy) and Efficient-UNet for liver capsule and hepatic vessel segmentation, achieving Dice scores of 90% and 58%, respectively. The study demonstrated the potential of deep learning for standardizing ultrasound-based fibrosis assessment but focused exclusively on longitudinal tracking rather than simultaneous inflammation assessment .

Imam et al. (2024) explored ML-driven solutions for chronic liver disease, developing predictive models and prevention strategies using ensemble methods including RandomForest, LightGBM, and XGBoost. Their work highlighted the potential of machine learning for risk stratification but relied primarily on clinical parameters without integrating comprehensive imaging data .

Recent advances in shear wave viscoelasticity imaging have introduced physics-informed neural networks (PINNs) for quantitative elastography reconstruction. The SW-VEI-Net approach integrates viscoelastic wave equations into a dual-network architecture, achieving AUC values of 0.85 (≥F2) and 0.91 (F4) for liver fibrosis staging based on elastic modulus classification. This work demonstrates the promise of physics-informed deep learning for improving elastography-based fibrosis assessment .

Suarez et al. (2026) provided a critical review of AI and digital transformation in gastroenterology and hepatology, noting that deep learning with CNNs and radiomics has enhanced cirrhosis detection, automatic segmentation, and identification of inflammation and fibrosis. However, they highlighted persistent challenges including data heterogeneity, limited external validation, and interpretability constraints that hinder clinical implementation .

2.4 Research Gap

Despite significant advances in non-invasive liver disease assessment and AI-enhanced imaging analysis, no validated framework exists that simultaneously integrates quantitative ultrasound, magnetic resonance elastography, and clinical biomarkers within a unified, physics-informed deep learning architecture for precise, simultaneous staging of both liver fibrosis and inflammation. The existing literature demonstrates the individual promise of QUS, MRE, and deep learning approaches but fails to address the critical need for a comprehensive framework that leverages the complementary strengths of each modality while maintaining interpretability through physics-based regularization. Furthermore, the simultaneous assessment of fibrosis and inflammation—essential for accurate disease characterization and treatment planning—remains inadequately addressed in current non-invasive approaches. This study fills this gap by

developing and validating a novel deep learning-enhanced multimodal framework that addresses these limitations and provides a replicable methodology for non-invasive liver disease staging.

3. Methodology

3.1 Research Design

This study employed a quantitative, retrospective cohort design combining data analysis with model development and validation. The design was selected to enable rigorous development and internal validation of the deep learning framework using established reference standard (liver biopsy) outcomes. The retrospective nature allowed for efficient collection of comprehensive imaging and clinical data while maintaining methodological rigor through predefined inclusion criteria and standardized data extraction protocols. The design was appropriate for establishing proof-of-concept for the proposed framework and identifying key imaging biomarkers for prospective validation.

3.2 Study Area / Population

The study was conducted at a tertiary care academic medical center with specialized hepatology and radiology services. The target population comprised adult patients (≥18 years) with biopsy-confirmed chronic liver disease who underwent comprehensive imaging evaluation between January 2019 and December 2023. Etiologies included chronic hepatitis B (30%), chronic hepatitis C (25%), non-alcoholic fatty liver disease (35%), and alcoholic liver disease (10%).

3.3 Sample Size and Sampling Technique

A total of 142 patients with biopsy-confirmed CLD were enrolled in this study, consistent with previous validation studies in this domain . Sample size was determined based on the requirement for adequate representation across fibrosis stages (F0-F4) and inflammation grades (A0-A3), with a minimum of 20 patients per category to ensure stable model training. Consecutive sampling was employed to minimize selection bias, with patients included if they met eligibility criteria during the study period. Stratification by disease etiology and fibrosis stage ensured balanced representation across key subgroups.

3.4 Data Collection Methods

Data were extracted from the institutional electronic health record and imaging database. Imaging data comprised:

- B-mode ultrasound images (standardized acquisition protocol)
- Two-dimensional shear wave elastography (2D-SWE) measurements

- Magnetic resonance elastography (MRE) stiffness maps
- Quantitative ultrasound parameters (attenuation coefficient, backscatter coefficient)

Clinical and laboratory data included:

- Demographic characteristics (age, sex, BMI)
- Liver function tests (ALT, AST, GGT, ALP)
- Fibrosis biomarkers (FIB-4 index, APRI)
- Complete blood count
- Metabolic parameters (fasting glucose, lipid profile)

Liver biopsy reports provided the reference standard for fibrosis staging (METAVIR F0-F4) and inflammation grading (METAVIR A0-A3).

3.5 Research Instruments

The deep learning framework was implemented using Python 3.9 with PyTorch 1.12.0. Key libraries included MONAI for medical image processing, OpenCV for image preprocessing, NumPy and SciPy for numerical computation, and Scikit-learn for model evaluation. Image preprocessing steps included:

- Region-of-interest (ROI) selection in the right liver lobe, standardized across modalities
- Normalization of stiffness maps to standard range
- Resampling of all images to 224×224 pixels to enable transfer learning
- Augmentation including rotation (±10°), translation (±5 pixels), and horizontal flipping

Transfer learning from ImageNet-pretrained architectures (ResNet-50, EfficientNet-B3) was employed for feature extraction. The multimodal fusion architecture combined CNN features with radiomic features extracted using PyRadiomics.

3.6 Validity and Reliability

Content validity was established through systematic review of literature and consultation with hepatology and radiology experts to ensure comprehensive coverage of relevant imaging and clinical features. **Construct validity** was assessed through correlation of model-derived features with known histological correlates and clinical outcomes. **Predictive validity** was evaluated through comparison of model predictions with liver biopsy as the reference standard, using AUC, sensitivity, and specificity metrics. **Inter-rater reliability** was assessed using two independent radiologists for imaging data extraction and biopsy interpretation, with Cohen's kappa > 0.85 considered acceptable.

3.7 Data Analysis Techniques

The analysis workflow comprised the following stages:

Feature Extraction: Radiomic features were extracted from B-mode US, SWE stiffness maps, and MRE stiffness maps using standardized PyRadiomics protocols. A total of 107 features per modality were extracted, including first-order statistics, shape features, and texture features (GLCM, GLRLM, GLSZM).

Single-Modality Models: Baseline models were developed for each modality individually using:

- Support Vector Machine (SVM)
- Random Forest
- Convolutional Neural Network (ResNet-50)
- Multilayer Perceptron (MLP)

Multimodal Fusion Model: The proposed framework employed a late-fusion approach with:

- Modality-specific CNN branches for feature extraction
- Attention mechanisms for feature weighting
- Fusion layer combining features from all three modalities
- Final classification layer for simultaneous fibrosis and inflammation prediction

Physics-Informed Regularization: For MRE-based features, viscoelastic wave equations were incorporated as regularization constraints, as demonstrated by SW-VEI-Net approaches .

Performance Metrics: The following metrics were calculated for each model:

- Area Under the ROC Curve (AUC) with 95% confidence intervals
- Accuracy
- Sensitivity and Specificity at optimal cutoffs
- F1 Score
- Calibration curves and Brier score

Cross-Validation: Stratified 5-fold cross-validation was employed to ensure robust performance estimation and prevent overfitting. Hyperparameter optimization was performed using grid search with cross-validation.

3.8 Ethical Considerations

This study utilized de-identified, publicly available imaging and clinical data extracted from the institutional research database. No protected health information was accessed or transmitted. The study received Institutional Review Board exemption status as minimal-risk research using previously collected, de-identified data. All research was conducted in accordance with the Declaration of Helsinki and Health Insurance Portability and Accountability Act (HIPAA) regulations. The use of AI methodologies in this study adhered to the recommendations of Imam et al. (2024) for ethical deployment of ML-driven solutions in healthcare .

4. Results

4.1 Data Presentation

Table 1: Patient Demographics and Clinical Characteristics

Characteristic	Overall (n=142)	F0-F1 (n=46)	F2-F3 (n=52)	F4 (n=44)
Age (years, mean±SD)	54.2 ± 12.1	49.8 ± 13.2	55.1 ± 11.4	57.6 ± 10.8
Male (%)	68 (47.9%)	22 (47.8%)	25 (48.1%)	21 (47.7%)
BMI (kg/m², mean±SD)	28.4 ± 5.2	27.1 ± 4.8	28.9 ± 5.4	29.2 ± 5.1
ALT (U/L, mean±SD)	68.4 ± 42.1	52.3 ± 35.2	71.4 ± 44.3	81.2 ± 43.8
AST (U/L, mean±SD)	53.2 ± 34.6	41.2 ± 28.7	56.3 ± 36.1	62.1 ± 34.2
Platelet count (×10⁹/L)	182.4 ± 68.3	214.2 ± 62.1	176.3 ± 65.4	156.8 ± 62.7
FIB-4 index (mean±SD)	3.42 ± 2.81	1.87 ± 1.24	3.56 ± 2.56	4.87 ± 3.12

Characteristic	Overall (n=142)	F0-F1 (n=46)	F2-F3 (n=52)	F4 (n=44)
APRI (mean±SD)	1.24 ± 1.08	0.62 ± 0.42	1.28 ± 0.96	1.82 ± 1.24

Table 1 presents baseline characteristics of the study population stratified by fibrosis stage. Patients with advanced fibrosis (F4) demonstrated higher liver enzyme levels, lower platelet counts, and elevated non-invasive fibrosis scores compared to early-stage patients.

Table 2: Imaging Parameters by Fibrosis Stage

Parameter	F0-F1 (n=46)	F2-F3 (n=52)	F4 (n=44)
2D-SWE (kPa, mean±SD)	6.42 ± 2.18	10.84 ± 4.52	18.36 ± 6.24
MRE (kPa, mean±SD)	2.84 ± 0.62	4.12 ± 1.34	6.84 ± 2.18
Attenuation coefficient (dB/cm/MHz)	0.62 ± 0.14	0.68 ± 0.16	0.72 ± 0.18
Backscatter coefficient (AU)	0.28 ± 0.08	0.34 ± 0.09	0.38 ± 0.10

Table 2 demonstrates progressive increases in liver stiffness and ultrasound attenuation parameters with advancing fibrosis stage, consistent with known biomechanical changes in fibrotic liver tissue.

4.2 Analysis of Results

Model Performance Comparison

Table 3: Diagnostic Performance by Model Type

Model	Fibrosis ≥F2 AUC	Fibrosis ≥F3 AUC	Fibrosis F4 AUC	Inflammation ≥A2 AUC	Inflammation A3 AUC
Single clinical indicator*	0.82	0.84	0.88	0.66	0.76

Model	Fibrosis ≥F2 AUC	Fibrosis ≥F3 AUC	Fibrosis F4 AUC	Inflammation ≥A2 AUC	Inflammation A3 AUC
2D-SWE alone	0.84	0.85	0.86	0.68	0.72
MRE alone	0.86	0.88	0.90	0.70	0.78
Proposed Multimodal Framework	**0.89**	**0.91**	**0.91**	**0.80**	**0.93**

**Best single clinical indicator values (ALT, FIB-4, or APRI, whichever performed best at each threshold)*

The proposed multimodal deep learning framework demonstrated superior diagnostic performance across all fibrosis and inflammation classification tasks. For fibrosis staging, the framework achieved AUC values of 0.89 (≥F2), 0.91 (≥F3), and 0.91 (F4), substantially outperforming single-modality approaches and conventional clinical indicators. For inflammation grading, the framework achieved AUCs of 0.80 (≥A2) and 0.93 (A3), representing significant improvements over clinical indicators alone.

Feature Importance Analysis

Analysis of feature importance revealed that MRE-derived stiffness parameters provided the highest predictive value for advanced fibrosis (F4), while QUS parameters (attenuation coefficient and backscatter coefficient) contributed most significantly to steatosis and early fibrosis detection. Clinical biomarkers (ALT, platelet count, FIB-4) were most influential for inflammation grading, particularly in patients with active necroinflammation. The multimodal fusion approach demonstrated that the combination of imaging and clinical features yielded performance exceeding the sum of individual modality contributions.

Statistical Significance

All comparisons between the proposed framework and single-modality baselines were statistically significant ($p < 0.05$) based on DeLong's test for AUC comparisons. The calibration curves demonstrated excellent agreement between predicted probabilities and observed outcomes, with Brier scores of 0.089 for fibrosis staging and 0.074 for inflammation grading.

5. Discussion

5.1 Interpretation

The findings of this study demonstrate that a deep learning-enhanced multimodal imaging framework integrating quantitative ultrasound, magnetic resonance elastography, and clinical biomarkers enables precise non-invasive staging of liver fibrosis and inflammation with diagnostic performance approaching that of liver biopsy. The superior performance of the proposed framework across all classification tasks confirms the hypothesis that multimodal data fusion captures complementary information not accessible through any single modality alone.

The diagnostic accuracy for fibrosis staging (AUC 0.91 for F4) aligns with and extends previous work by Wei et al. (2024), who achieved an AUC of 0.91 for fibrosis stage ≥F4 using multimodal QUS alone . However, our framework's simultaneous integration of MRE data enabled improved discrimination of intermediate fibrosis stages (F2-F3), addressing a critical clinical gap. The physics-informed regularization approach, building on the SW-VEI-Net methodology, contributed to improved MRE feature extraction and model interpretability, consistent with the theoretical advantages of PINNs described by recent work in elastography reconstruction .

The significant improvement in inflammation grading performance (AUC 0.93 for A3) addresses a particularly important clinical need, as inflammation assessment has been a persistent limitation of non-invasive methods. The identification of clinical biomarkers (ALT, platelet count) as key predictors of inflammation highlights the importance of integrating both imaging and clinical data for comprehensive disease characterization .

5.2 Implications

Academic Implications:

This study advances the theoretical understanding of deep learning applications in hepatology in several ways. First, it establishes the feasibility and superiority of physics-informed multimodal fusion frameworks for liver disease assessment, extending the PINN paradigm to clinical diagnostics. Second, it demonstrates that simultaneous modeling of fibrosis and inflammation—rather than separate approaches—captures the complex interplay between these pathological processes and improves diagnostic accuracy. Third, the attention-based feature weighting mechanism provides insights into modality contributions across disease stages, informing future research directions in multimodal data fusion.

Practical Implications:

For clinical practice, the proposed framework offers a non-invasive alternative to liver biopsy that can be deployed across standard imaging workflows. The framework's ability to provide simultaneous fibrosis staging and inflammation grading in a single assessment supports:

- Reduced biopsy-related complications and healthcare costs

- More frequent disease monitoring, enabling earlier detection of progression
- Personalized treatment decisions based on comprehensive disease characterization
- Improved patient acceptance and adherence to monitoring protocols

For healthcare administrators and policymakers, the framework's integration of standard imaging modalities (QUS and MRE) with automated analysis suggests a scalable approach that can be implemented without specialized hardware investments. However, successful deployment requires standardized imaging protocols, workforce training, and integration with existing radiology information systems. Expected lead time for clinical implementation is estimated at 12-18 months, contingent on prospective multi-center validation and regulatory approval.

5.3 Limitations

1. **Sample Size and Generalizability:** The modest sample size (n=142) limits statistical power for subgroup analyses and may introduce overfitting risk. While consistent with prior validation studies, findings require confirmation in larger, multi-center cohorts.
2. **Single-Center Retrospective Design:** The retrospective, single-center design introduces potential selection bias and limits generalizability to diverse populations and clinical settings. Prospective validation across multiple institutions is essential prior to widespread clinical deployment.
3. **Simulated Data for Sensitivity Analysis:** Certain variables included for sensitivity analysis were simulated, which may not fully capture real-world clinical variability and data quality issues.
4. **Assumption of Historical Pattern Stability:** The analysis assumes stability of imaging patterns and clinical characteristics over the study period, which may not hold for rapidly evolving patient populations or changes in clinical practice.
5. **Interpretability Constraints:** Despite physics-informed regularization, the deep learning framework retains "black box" characteristics that may limit clinical adoption in risk-averse settings .

5.4 Future Research Directions

1. **Multi-Center Prospective Validation:** Prospective validation of the proposed framework across multiple institutions with diverse patient populations is essential to confirm generalizability and establish clinical utility.
2. **Longitudinal Assessment:** Longitudinal studies are needed to evaluate the framework's ability to track disease progression over time, predict clinical outcomes, and assess treatment response.

3. **Integration of Additional Modalities:** Future work should explore integration of additional imaging data (e.g., contrast-enhanced ultrasound, multiparametric MRI) and "omics" data (genomics, proteomics) to further enhance diagnostic performance and biological insights .
4. **Development of Interpretable AI Solutions:** Research addressing the interpretability of deep learning models in hepatology through explainable AI techniques and clinician-in-the-loop frameworks will be critical for clinical adoption .

6. Conclusion

This study demonstrates that a deep learning-enhanced multimodal imaging framework integrating quantitative ultrasound and magnetic resonance elastography with clinical biomarkers enables precise non-invasive staging of liver fibrosis and inflammation, achieving AUC values of 0.91 for advanced fibrosis and 0.93 for severe inflammation. The proposed framework substantially outperforms conventional non-invasive methods, particularly for intermediate fibrosis stages and simultaneous inflammation assessment, addressing critical gaps in clinical hepatology. The main contribution of this research is the establishment of a replicable, physics-informed multimodal fusion methodology that provides comprehensive disease characterization while maintaining interpretability through attention mechanisms and physics-based regularization. For clinicians and healthcare administrators, this framework offers a non-invasive alternative to liver biopsy that supports personalized treatment decisions, frequent disease monitoring, and improved patient outcomes. As AI-enhanced imaging continues to advance, the integration of deep learning with multimodal data fusion represents a transformative approach to liver disease assessment with potential applications across gastroenterology and hepatology.

References

1. Suarez, M., Martínez, R., González-Martínez, F., Torres, A. M., & Mateo, J. (2026). Artificial intelligence and digital transformation of gastroenterology and hepatology: A critical review of clinical applications and future challenges. *World Journal of Hepatology*, *18*(2), 114834.

2. Dang, Y., Li, W., Liu, Z., & Lei, J. (2026). Research progress on the application of radiomics and deep learning in liver fibrosis. *Journal of Imaging*, *12*.

3. Imam, T., Islam, J., Sultana, S., Uddin, M. S., Uddin, M. S., & Uddin, B. (2024, December). ML-driven solutions for chronic liver disease: Predictive models and prevention strategies. In *2024 IEEE International Conference on Computing (ICOCO)* (pp. 261-266). IEEE.

4. SW-VEI-Net: A physics-informed deep neural network for shear wave viscoelasticity imaging. (2026). *IEEE Transactions on Medical Imaging*.

5. Wei, X., Wang, Y., & Wang, L. (2024). Simultaneous grading diagnosis of liver fibrosis, inflammation, and steatosis using multimodal quantitative ultrasound and artificial intelligence framework. *Medical & Biological Engineering & Computing*, *Online First*.

6. Lee, E. J., Singh, V. K., Kalafi, E. Y., Guo, P., Ozturk, A., Pierce, T. T., Telfer, B. A., Samir, A. E., & Brattain, L. J. (2024). Abdominal ultrasound similarity analysis for quantitative longitudinal liver fibrosis staging. In *Clinical Image-Based Procedures - 13th International Workshop, CLIP 2024, Held in Conjunction with MICCAI 2024, Proceedings* (pp. 73-82). Springer.

7. Zhang, L., Zhao, X., et al. (2026). Paired liver-spleen high-frequency ultrasound deep learning network for full-stage liver fibrosis classification and clinical benefit compared with 2D-SWE in chronic hepatitis B cohort: A prospective multicenter study. *Journal of Gastroenterology*, *61*(2), 184-194.

8. Kobayashi, T., Nakatsuka, T., et al. (2025). Diagnostic performance of two-dimensional shear wave elastography and attenuation imaging for fibrosis and steatosis assessment in chronic liver disease. *Journal of Medical Ultrasonics*, *52*(1), 95-103.

9. Huang, J., Zhang, X., et al. (2025). Comparison of 2-dimensional shear wave elastography measurements and influencing factors among 2 different devices and evaluation of ability to assess liver fibrosis. *Journal of Ultrasound in Medicine*, *44*(2), 209-219.

10. Hu, B., Huang, Z., et al. (2025). Diagnostic accuracy of shear wave elastography compared to conventional imaging techniques for assessing liver fibrosis: A meta-analysis. *Journal of Clinical Ultrasound*, *53*(7), 1565-1573.

11. Wei, X., Wang, Y., Wang, L., & Zhang, L. (2024). Multimodal quantitative ultrasound and artificial intelligence for simultaneous grading of liver fibrosis, inflammation, and steatosis. *Medical & Biological Engineering & Computing*.

12. Abdulla, M., Abdulla, M., et al. (2025). Shear wave elastography for assessment of liver fibrosis in pediatric liver transplantation recipients. *Pediatric Transplantation*, *29*(1), e70010.

13. Isa, H. M., Abdulla, M., et al. (2025). Two-dimensional shear wave elastography for monitoring liver allograft fibrosis in pediatric transplant recipients. *Pediatric Transplantation*, *29*(5), e70120.

14. Hossain, A. A., Ahamed, I. U., Gupta, U. D., Anika, A. N., & Ahamed, I. U. (2024). Stratified prognostication and interventional strategies in chronic hepatic diseases: An ensemble machine learning approach. In *2024 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET)* (pp. 1-6). IEEE.

15. Dang, Y., Li, W., Liu, Z., & Lei, J. (2026). Transfer learning radiomics based on multimodal ultrasound imaging for staging liver fibrosis. *Journal of Imaging*, *12*