



Published in final edited form as:

NEJM AI. 2025 March ; 2(3): . doi:10.1056/aioa2400948.

Opportunistic Screening of Chronic Liver Disease with Deep-Learning–Enhanced Echocardiography

Yuki Sahashi, M.D., M.Sc.¹, Milos Vukadinovic, B.S.^{1,2}, Fatemeh Amrollahi, Ph.D.³, Hirsh Trivedi, M.D.⁴, Justin Rhee, B.S.⁵, Jonathan Chen, M.D., Ph.D.³, Susan Cheng, M.D., M.M.Sc., M.P.H.¹, David Ouyang, M.D.^{1,6}, Alan C. Kwan, M.D.¹

¹Department of Cardiology, Smidt Heart Institute, Cedars–Sinai Medical Center, Los Angeles

²Department of Bioengineering, University of California Los Angeles, Los Angeles

³Center for Biomedical Informatics Research, Department of Medicine, Stanford University, Palo Alto, CA

⁴Karsh Division of Gastroenterology and Hepatology, Department of Medicine, Cedars–Sinai Medical Center, Los Angeles

⁵School of Medicine, Brown University, Providence, RI

⁶Division of Artificial Intelligence in Medicine, Cedars–Sinai Medical Center, Los Angeles

Abstract

BACKGROUND: Chronic liver disease (CLD) affects more than 1.5 billion adults, most of whom are asymptomatic and undiagnosed. Echocardiography is broadly performed and visualizes the liver, but this information is not clinically leveraged for CLD diagnosis.

METHODS: We developed and evaluated EchoNet-Liver, a deep-learning, computer-vision pipeline that can identify high-quality subcostal images from full echocardiogram studies and detect the presence of cirrhosis and steatotic liver disease (SLD). This retrospective observational study included adult patients from two large urban academic medical centers who received both echocardiography and abdominal imaging — either ultrasound or magnetic resonance imaging (MRI) — within 30 days. The model predictions were compared with diagnoses from clinical evaluations of paired abdominal ultrasound or MRI studies.

RESULTS: A total of 1,596,640 echocardiogram videos from 66,922 studies and 24,276 patients at Cedars–Sinai Medical Center (CSMC) were used to develop EchoNet-Liver. In a held-out CSMC test cohort, EchoNet-Liver detected cirrhosis with an area under the receiver operating characteristic curve (AUROC) of 0.837 (95% confidence interval [CI], 0.828 to 0.848) and SLD with an AUROC of 0.799 (95% CI, 0.788 to 0.811). In a separate test cohort with paired abdominal MRI studies, EchoNet-Liver detected cirrhosis with an AUROC of 0.704 (95% CI, 0.699 to 0.708) and SLD with an AUROC of 0.725 (95% CI, 0.707 to 0.762). In an external test cohort of 106 patients (5280 videos), the model detected cirrhosis with an AUROC of 0.830 (95% CI, 0.799 to 0.859) and SLD with an AUROC of 0.769 (95% CI, 0.733 to 0.813).

CONCLUSIONS: Deep-learning assessment of clinical echocardiography enables opportunistic screening for SLD and cirrhosis. The application of this algorithm can identify patients who may benefit from further diagnostic testing and treatment for CLD. (Funded by KAKENHI [Japan Society for the Promotion of Science, 24K10526] and others.)

Introduction

Chronic liver disease (CLD) affects an estimated 1.5 billion people worldwide and 100 million in the United States and can result in malignancy, end-stage liver disease, or mortality.¹ The prevalence of CLD is sharply increasing,^{2,3} particularly with respect to steatotic liver disease (SLD), resulting from an increased burden of obesity and metabolic disease. The vast majority of CLD cases are undiagnosed.^{4,5} With the joint risk factors of diabetes, obesity, and alcohol abuse, subclinical liver disease affects many individuals with known cardiovascular disease, as well as those with nonspecific findings of fatigue, shortness of breath, and abdominal distension.⁴

Multiple approaches are available for screening and diagnosis of CLD, including serologic risk scores, qualitative and quantitative ultrasound and magnetic resonance imaging (MRI), and invasive biopsy.^{6,7} However, accuracy, availability, and cost limit the capacity of these pathways to address the problem of underdiagnosis.

Echocardiography, or ultrasound of the heart and associated structures, is a first-line cardiovascular diagnostic test, and is frequently performed across the spectrum of patients with metabolic and cardiovascular diseases.⁸ Included within a standard echocardiographic examination are subcostal views, which visualize the inferior vena cava and provide clear visualization of hepatic tissue and liver contour. The full clinical value of these images for the identification of hepatic disease is underexplored, as cardiologists are not trained in the assessment of liver pathologies by ultrasound.

Artificial intelligence (AI) can identify diseases and characteristics that may not be readily observable by the human eye,^{9–13} and it can predict disease progression¹⁴ and mortality,¹⁵ and improve measurement accuracy of cardiac parameters.^{16–19} Our study aims to validate an AI computer-vision approach that leverages echocardiographic images and videos to detect CLD. We hypothesize that a deep-learning pipeline can identify high-quality subcostal-view videos and detect SLD and cirrhosis in a high-throughput fashion.

To that end, we developed EchoNet-Liver, a deep-learning computer-vision pipeline, and trained the model and evaluated its performance with internal and external validations across multiple cohorts (Fig. 1). AI-based analysis of hepatic tissue visible within standard subcostal echocardiography videos may permit opportunistic screening of CLD from standard echocardiography without additional costs.

Methods

COHORT SELECTION

We included patients 18 years of age or older without liver transplant who received both an echocardiogram and an abdominal ultrasound within 30 days at Cedars–Sinai Medical

Center (CSMC). Clinical diagnoses from the abdominal ultrasound report — including normal liver, steatotic liver, and cirrhotic liver — were paired with the echocardiogram images as labels for training and validation. We also created an independent test cohort of patients who received an echocardiogram and an abdominal MRI, to evaluate the performance of EchoNet-Liver based on different disease definitions and populations. Studies from patients within any test cohort (ultrasound or MRI) were excluded from the associated training cohort (Tables S1–S4 in the Supplementary Appendix).

Electronic health records were used to determine cohort demographics, comorbidities, and characteristics as well as pertinent *International Classification of Diseases*, Ninth Revision or *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision codes within 1 year of the echocardiogram study. A cohort of patients from Stanford Health Care (SHC) who had received both an echocardiogram and an abdominal ultrasound within 30 days were included as external validation. An overall flow chart of model development and evaluation is provided in Figure 1. Approval for this study was obtained from the CSMC and SHC institutional review boards, and the requirement for informed consent was waived for retrospective data analysis without patient contact.

DISEASE DEFINITIONS

Echocardiography studies were matched to the closest abdominal ultrasound study within 30 days. For echocardiography studies associated with more than one abdominal ultrasound study within the 30-day timeframe, CLD labels were derived from the abdominal ultrasonic study that was performed nearest in time to the echocardiography study. The CLD diagnoses from the abdominal ultrasound reports were categorized as normal, cirrhosis, or SLD, based on text describing the liver parenchyma. Patients who had reports that specified normal liver parenchyma were identified as controls. In the abdominal MRI group, diagnosis labels for cirrhosis were taken from the clinical MRI report. For SLD, we focused on the subset of cases for whom the MRI-derived proton density fat fraction (MRI-PDFF) or magnetic resonance spectroscopy (MRS)–derived fat signal was measured, with larger than 5.0% being diagnostic of SLD.²⁰ Deidentified patient sample reports of abdominal ultrasound and MRI for each normal, cirrhosis, and SLD cases are described in Supplementary Appendix, Methods 1.

VIEW SELECTION AND IMAGE PROCESSING

A standard echocardiogram study often contains 50 to 100 videos, of which typically only one or two capture the liver with sufficient image quality for assessment of echotexture and contour. Echocardiography videos were initially obtained as Digital Imaging and Communications in Medicine (DICOM) files and underwent deidentification and processing into audio–video–interleave format. We developed a pipeline of two deep-learning models for view classification (i.e., identification of the subcostal-view videos) and quality control (i.e., exclusion of videos with severe motion artifacts or low image quality). Representative ground-truth images for the view classification and quality-control model are described in Figure S1. Detailed methods for both models are summarized in Supplementary Appendix, Methods 2.

CHRONIC LIVER DISEASE DETECTION

For the training and evaluation of the liver disease detection models, we utilized an image-based model (DenseNet²¹) to focus on hepatic tissue texture for the prediction of cirrhosis and SLD. Given the focus on texture, the input data were still images from high-quality subcostal echocardiogram videos at native resolution (480×640 pixels). We trained the model to minimize binary cross-entropy loss using an AdamW optimizer with an initial learning rate of 1×10^{-5} , at a batch size of 40 for 100 epochs. During each epoch, a randomly sampled frame from each video was used as an independent training example. A variety of hyperparameters and architectures were compared prior to the development of the final EchoNet-Liver model (Table S5). Model training and evaluation were conducted using Python 3.8.0, PyTorch 2.2, and torchvision 0.17.

STATISTICAL ANALYSIS

Model performance was determined by measuring the area under the receiver operating characteristic (AUROC) curve on held-out test cohorts. The area under the precision–recall curve (AUPRC), sensitivity, specificity, positive predictive value, and negative predictive value were similarly reported at the Youden Index. All 95% confidence intervals were calculated with 10,000 bootstrapping samples by echocardiography study level. Data analysis was performed using both Python (version 3.8.0) and R (version 4.2.2) programming languages. This study was carried out following the TRIPOD+AI guideline.²² Saliency maps were generated to identify the areas of interest for the classifier across all test datasets. Each saliency map was produced using gradient-weighted class activation mapping,²³ which captures the gradient information directed into the final convolutional layer of the trained deep-learning model. We input the final layer of the fourth Dense Block for this approach.

CODE AND DATA AVAILABILITY

Code is available at <https://github.com/echonet/liver/>, and training data are available with the submission of a research protocol and are subject to approval by the institutional review boards of CSMC and SHC.

Results

PATIENT CHARACTERISTICS

A total of 1,596,640 videos, from 66,922 CSMC echocardiography studies of 24,276 patients (conducted between July 2012 and June 2022), were identified and split 8:1:1 by patient into training, validation, and held-out test cohorts. These patients exhibited a variety of comorbidities typical among those who receive both echocardiography and abdominal ultrasound examinations (Table 1), including prevalent hypertension (28.8%), hyperlipidemia (20.1%), diabetes (19.3%), hepatitis B (1.2%), and hepatitis C (3.7%). The average body mass index was 26.5 ± 6.2 , and 5.3% of the patients regularly consumed alcohol. The median duration between the echocardiography and abdominal ultrasound examinations was 0 days (interquartile range, –4 to +3 days). From the total CSMC cohort, a test cohort for the AI pipeline was selected, consisting of 1486 patients with similar

demographics to the total cohort (Table 1). After view classification and quality control, the CLD detection algorithms were applied to 1072 subcostal videos from 431 patients for cirrhosis and 1142 subcostal videos from 459 patients for SLD. There were 69 (6.4%) cirrhosis videos from 34 (7.9%) patients, and 143 (12.5%) SLD videos from 69 (15.0%) patients based on abdominal ultrasound reports (Fig. 1). Patient characteristics in each model (i.e., view-classifier model, quality-control model, cirrhosis-detection model, and SLD-detection model) are shown in Tables S1–S4.

We identified an additional test dataset of 164,579 echocardiogram videos from 6751 studies among 2335 CSMC patients who underwent abdominal MRI for cirrhosis, MRI-PDFF or MRS for SLD, plus echocardiography, within 365 days (Table 1). The patients in this additional test cohort were not included in the training and validation cohorts of the original disease detection models. From this cohort, after view classification and quality control, the final MRI cirrhosis-detection cohort consisted of 6243 subcostal videos from 2137 patients, containing 1661 (26.6%) cirrhosis videos from 576 (27.0%) patients; and the final SLD-detection cohort consisted of 250 subcostal videos from 85 patients, containing 95 (38.0%) SLD videos from 38 (44.7%) patients (Fig. 1).

In the SHC cohort for EchoNet-Liver evaluation, our test cohort included 106 studies from 106 patients (Table 1). After view-classifier and quality control, we identified 130 subcostal videos from 66 individual patients for CLD detection. Among all videos, 20 (15.4%) videos from 10 (15.2%) patients were positive for cirrhosis, and 21 (16.2%) videos from 11 (16.7%) were positive for SLD, based on abdominal ultrasound.

VIEW-CLASSIFIER AND QUALITY CONTROL PERFORMANCE

A total of 11,419 subcostal-view videos were used for training a subcostal-view classifier model; all other videos were labeled as nonsubcostal controls. On an independent test set of 100 CSMC studies, including 2315 videos, the view-classifier model identified 186 out of 196 subcostal-view videos with an AUROC of 0.991 (95% confidence interval [CI], 0.986 to 0.997); the sensitivity was 0.949 (95% CI, 0.920 to 0.970), and the specificity was 0.999 (95% CI, 0.999 to 1.000). In the SHC population, the subcostal-view classifier identified 149 out of 228 subcostal videos from 5280 total videos with an AUROC of 0.965 (95% CI, 0.955 to 0.976); the sensitivity was 0.654 (95% CI, 0.590 to 0.720), and the specificity was 0.993 (95% CI, 0.991 to 0.995). In a comparison of image quality by two cardiologists, the quality-control model demonstrated an AUROC of 0.855 (95% CI, 0.812 to 0.901) in the CSMC videos and an AUROC of 0.785 (95% CI, 0.727 to 0.840) in the SHC videos.

DISEASE DETECTION PERFORMANCE

In the held-out CSMC test ultrasound dataset, EchoNet-Liver detected cirrhosis with an AUROC of 0.837 (95% CI, 0.828 to 0.848) and SLD with an AUROC of 0.799 (95% CI, 0.788 to 0.811), as shown in Figure 2. The algorithm showed a sensitivity of 0.696 (95% CI, 0.672 to 0.726) and a specificity of 0.846 (95% CI, 0.841 to 0.853) for detecting cirrhosis; and a sensitivity of 0.741 (95% CI, 0.722 to 0.763) and a specificity of 0.720 (95% CI, 0.712 to 0.728) for detecting SLD.

In the SHC external-validation test cohort, EchoNet-Liver detected cirrhosis with an AUROC of 0.830 (95% CI, 0.799 to 0.858) and SLD with an AUROC of 0.769 (95% CI, 0.733 to 0.813). In the SHC test cohort, EchoNet-Liver demonstrated a sensitivity of 0.800 (95% CI, 0.733 to 0.842) and a specificity of 0.709 (95% CI, 0.686 to 0.740) for detecting cirrhosis, and a sensitivity of 0.667 (95% CI, 0.611 to 0.765) and specificity of 0.780 (95% CI, 0.760 to 0.830) for detecting SLD. All evaluation metrics for the CSMC held-out test cohort and the SHC external-validation test cohort, including AUROC, AUPRC, positive predictive value, negative predictive value, sensitivity, and specificity at Youden Index are summarized in Table 2. Table S6 and Figure S2 are presented to demonstrate how these metrics change as the model output thresholds are adjusted.

To check the robustness of EchoNet-Liver disease detection, the discriminative ability was confirmed in the held-out CSMC test ultrasound dataset by disease severity using the Fibrosis-4 Index score for liver fibrosis (Table S7). In this sensitivity analysis, the performance was similar for both types of disease, regardless of severity. Furthermore, we performed sensitivity analyses in which we chose the most recent echocardiography from each patient if the patient had undergone several echocardiogram examinations. In this analysis, EchoNet-Liver detected the presence of cirrhosis with an AUROC of 0.871 (95% CI, 0.862 to 0.883) and an AUPRC of 0.477 (95% CI, 0.402 to 0.503), and it detected the presence of SLD with an AUROC of 0.796 (95% CI, 0.782 to 0.810) and an AUPRC of 0.464 (95% CI, 0.424 to 0.483).

COMPARISON WITH DIAGNOSIS BY MRI

To evaluate the performance of EchoNet-Liver across multiple diagnostic pathways, the algorithm was evaluated in a cohort of patients who received both echocardiography and abdominal MRI at CSMC. In the MRI-paired cohort, EchoNet-Liver detected the presence of cirrhosis with an AUROC of 0.704 (95% CI, 0.699 to 0.708), as shown in Figure 2A, and an AUPRC of 0.493 (95% CI, 0.484 to 0.501), as shown in Figure S3A. When evaluated in the MRI-PDFF/MRS steatosis cohort, the model detected the presence of SLD with an AUROC of 0.725 (95% CI, 0.707 to 0.762) and an AUPRC of 0.619 (95% CI, 0.588 to 0.670), as shown in Figure 2B and Figure S3B.

COMPARISONS WITH CLINICAL DIAGNOSIS OF LIVER DISEASE

Variations in model performance occur depending on the imaging modality with which EchoNet-Liver is compared, as even abdominal ultrasound and MRI have significantly discordant test characteristics. In the CSMC cohort, when using abdominal MRI as the reference standard, abdominal ultrasound had a sensitivity of 0.91 and a specificity of 0.90 for cirrhosis, and a sensitivity of 0.79 and specificity of 0.71 for SLD, thus demonstrating heterogeneity based on diagnostic approach and cohort definitions. Notably, in our clinical database of 280,615 echocardiography studies, only 25,532 (9.1%) were conducted in patients who had abdominal imaging within 30 days, and 81,607 (29.1%) were conducted in patients who had abdominal ultrasound at any time. This finding suggests an unmet need, as most patients receiving cardiac imaging do not also get abdominal imaging. Moreover, when we examine all 125,293 patients with echocardiography, regardless of any history of

liver diagnostic imaging, 2722 (2.17%) had a prior clinical diagnosis of cirrhosis, and 4763 (3.80%) had a prior diagnosis of SLD.

To estimate the potential proportions of unrecognized diagnoses, we applied EchoNet-Liver to a temporally split test sample of 1241 clinical echocardiography studies independent of contemporaneous abdominal ultrasound and without previous clinical diagnosis. Based on the selected cutoffs shown in Table S6, 14 patients (1.13%) were predicted as suspicious for cirrhosis, and 99 (7.98%) were predicted as suspicious for SLD. Although the true prevalences of undetected CLD will require prospective validation, these percentages suggest a definite potential for identifying undiagnosed at-risk patients who may benefit from an opportunistic screening approach.

MODEL EXPLAINABILITY

We generated saliency maps for representative echocardiography images from the CSMC held-out dataset and the SHC external dataset (Fig. 3). For both cirrhosis and SLD, EchoNet-Liver highlighted the liver and showed diffuse activation throughout the hepatic parenchyma in the subcostal echocardiographic images as regions of interest.

Discussion

In this study, we demonstrated the strong performance of a deep-learning pipeline (EchoNet-Liver) for detecting cirrhosis and SLD from clinical echocardiography images. Using routinely acquired transthoracic echocardiogram studies, EchoNet-Liver identified previously collected subcostal-view images containing the liver and classified them as having a high suspicion of CLD. The discriminative ability of the model was confirmed among patients in a geographically distinct external health care system, those with paired abdominal MRI imaging, and those with a clinical diagnosis of CLD without abdominal imaging. Across diverse populations and disease definitions, deep-learning-enhanced echocardiography enabled high-throughput, automated detection of CLD, which could enable opportunistic screening for asymptomatic liver disease.

CLD often remains undiagnosed due to the asymptomatic nature of early disease. Despite the significant prevalence and morbidity of CLD, routine screening is not recommended due to a lack of evidence of its cost-effectiveness.²⁴ Opportunistic screening from echocardiogram images can identify high-risk patients with concurrent cardiovascular risk and liver disease in a cost-effective manner. By harnessing preexisting imaging indicated for other diagnostic purposes, our AI-enhanced workflow can increase the utility of existing imaging examinations and augment clinical suspicion for subclinical disease.²⁵ A similar benefit is possible in other modalities, including screening of coronary calcium in nongated chest computer tomography and cardiomyopathy in chest x-rays.

By incorporating view classification, quality control, and disease detection in one pipeline, automation with AI can enable high-throughput evaluation of this noninvasive and common cardiovascular diagnostic test. Furthermore, distinguishing between hepatic congestion from chronic heart failure and CLD, or their coexistence, can often be challenging. Our deep-learning model can aid the clinician in thinking about the differential diagnosis of these

conditions. However, further evaluation and prospective testing are necessary before clinical adoption.

There are several limitations in the present study. First, it is a retrospective study conducted at two tertiary care centers for patients who have undergone both abdominal ultrasound and echocardiography, which may result in selection bias and limited sample sizes for testing. Second, the spectrum of age, gender, race, and comorbidities in the study dataset may not represent the general population and may be biased toward patients with excess comorbidity, necessitating further external validation and prospective trials. Third, the model was developed with a cohort of patients who underwent both abdominal ultrasound and echocardiography within a 30-day period, and thus for whom the prevalence of liver disease is likely higher than among the general population receiving echocardiography.

The true clinical utility of this model will depend on whether or not its application to a general echocardiography population can efficiently detect undiagnosed CLD. Furthermore, the utility can vary depending on several factors, including disease prevalence in the general population, the costs of additional testing, and the severity of the disease. To assess the clinical utility and optimal thresholds of the model based on the clinical context, a prospective trial focusing on patients who have not previously been diagnosed with liver disease will be required. Comparison of EchoNet-Liver performance in cohorts with different disease definitions naturally results in different metrics, as disparate imaging modalities such as abdominal ultrasound and MRI have different sensitivities and specificities for diagnosis of CLD.

In developing EchoNet-Liver, we tried to balance the scale of training data with the clinical acceptability of performance, and we chose to use radiologists' diagnoses based on abdominal ultrasound as the training label. As such, EchoNet-Liver's performance might recapitulate some of the biases of CLD diagnosis by abdominal ultrasound. However, we demonstrated consistent discriminative ability across datasets from different modalities and geographically distinct populations.

In totality, this study suggests that the clinical utility of high-throughput disease screening using AI is promising, particularly for early disease, and can enhance the utility of preexisting imaging data.^{11,12,26} Further studies are warranted to establish the optimal clinical workflow for opportunistic liver-disease screening among cardiovascular-disease patients and downstream treatment. By improving the diagnosis of subclinical CLD, we may be able to limit or reverse disease progression^{27–29} and improve care by triaging patients toward appropriate clinical and diagnostic management.^{24,30}

In conclusion, we found that EchoNet-Liver, a deep-learning pipeline using echocardiography to detect the presence of SLD and cirrhosis, had strong performance in multiple populations and disease definitions. The findings were consistent across two institutions and in comparison with abdominal MRI. Deep learning applied to echocardiography may offer opportunistic and cost-effective screening for CLD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Disclosures

Author disclosures and other supplementary materials are available at ai.nejm.org.

Dr. Sahashi reports support from KAKENHI (Japan Society for the Promotion of Science, 24K10526), overseas research grant from SUNRISElab, Japan Heart Foundation and Ogawa Foundation and honoraria for lectures from m3.com inc. Dr. Ouyang reports support from the National Institutes of Health (NIH, NHLBI R00HL157421 and R01HL173526) and Alexion, and consulting or honoraria for lectures from EchoIQ, Ultramix, Pfizer, InVision, the Korean Society of Echocardiography, and the Japanese Society of Echocardiography. Dr. Kwan reports consulting fees from InVision and support from the American Heart Association (AHA, 23CDA1053659) and National Institutes of Health (NIH, KL2TR001882).

References

1. Moon AM, Singal AG, Tapper EB. Contemporary epidemiology of chronic liver disease and cirrhosis. *Clin Gastroenterol Hepatol* 2020;18:2650–2666. DOI: 10.1016/j.cgh.2019.07.060. [PubMed: 31401364]
2. Tian H, Zhang K, Hui Z, et al. Global burden of non-alcoholic fatty liver disease in 204 countries and territories from 1990 to 2019. *Clin Res Hepatol Gastroenterol* 2023;47:102068. DOI: 10.1016/j.clinre.2022.102068. [PubMed: 36513252]
3. Sepanlou SG, Safiri S, Bisignano C, et al. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2020;5:245–266. DOI: 10.1016/S2468-1253(19)30349-8. [PubMed: 31981519]
4. Bajaj JS, Silvey SG, Rogal S, et al. Undiagnosed cirrhosis and hepatic encephalopathy in a national cohort of veterans with dementia. *JAMA Netw Open* 2024;7:e2353965. DOI: 10.1001/jamanetworkopen.2023.53965. [PubMed: 38294815]
5. Castera L, Laouenan C, Vallet-Pichard A, et al. High prevalence of NASH and advanced fibrosis in type 2 diabetes: a prospective study of 330 outpatients undergoing liver biopsies for elevated ALT, using a low threshold. *Diabetes Care* 2023;46:1354–1362. DOI: 10.2337/dc22-2048. [PubMed: 37043830]
6. Sterling RK, Patel K, Duarte-Rojo A, et al. AASLD practice guideline on blood-based non-invasive liver disease assessments of hepatic fibrosis and steatosis. *Hepatology* 2024;81:321–357. DOI: 10.1097/HEP.0000000000000845. [PubMed: 38489523]
7. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, et al. AASLD practice guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology* 2023;77:1797–1835. DOI: 10.1097/HEP.0000000000000323. [PubMed: 36727674]
8. Virnig BA, Shippee ND, O'Donnell B, Zeglin J, Parashuram S. Trends in the use of echocardiography, 2007 to 2011: Data Points #20. In: Data points publication series. Rockville, MD: Agency for Healthcare Research and Quality (US), May 13, 2014 (<https://pubmed.ncbi.nlm.nih.gov/24967475/>).
9. Yuan N, Duffy G, Dhruva SS, et al. Deep learning of electrocardiograms in sinus rhythm from US veterans to predict atrial fibrillation. *JAMA Cardiol* 2023;8:1131. DOI: 10.1001/jamacardio.2023.3701. [PubMed: 37851434]
10. Hughes JW, Yuan N, He B, et al. Deep learning evaluation of biomarkers from echocardiogram videos. *EBioMedicine* 2021;73:103613. DOI: 10.1016/j.ebiom.2021.103613. [PubMed: 34656880]
11. Duffy G, Cheng PP, Yuan N, et al. High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *JAMA Cardiol* 2022;7:386–395. DOI: 10.1001/jamacardio.2021.6059. [PubMed: 35195663]

12. Vrudhula A, Duffy G, Vukadinovic M, Liang D, Cheng S, Ouyang D. High-throughput deep learning detection of mitral regurgitation. *Circulation* 2024;150:923–933. DOI: 10.1161/CIRCULATIONAHA.124.069047. [PubMed: 39129623]
13. Goto S, Mahara K, Beussink-Nelson L, et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nat Commun* 2021;12:2726. DOI: 10.1038/s41467-021-22877-8. [PubMed: 33976142]
14. Oikonomou EK, Holste G, Yuan N, et al. A multimodal video-based AI biomarker for aortic stenosis development and progression. *JAMA Cardiol* 2024;9:534. DOI: 10.1001/jamacardio.2024.0595. [PubMed: 38581644]
15. Lau ES, Di Achille P, Kopparapu K, et al. Deep learning-enabled assessment of left heart structure and function predicts cardiovascular outcomes. *J Am Coll Cardiol* 2023;82:1936–1948. DOI: 10.1016/j.jacc.2023.09.800. [PubMed: 37940231]
16. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* 2020;580:252–256. DOI: 10.1038/s41586-020-2145-8. [PubMed: 32269341]
17. Kwan AC, Chang EW, Jain I, et al. Deep learning-derived myocardial strain. *JACC Cardiovasc Imaging* 2024;17:715–725. DOI: 10.1016/j.jcmg.2024.01.011. [PubMed: 38551533]
18. Tromp J, Seekings PJ, Hung CL, et al. Automated interpretation of systolic and diastolic function on the echocardiogram: a multicohort study. *Lancet Digit Health* 2022;4:e46–e54. DOI: 10.1016/S2589-7500(21)00235-1. [PubMed: 34863649]
19. He B, Kwan AC, Cho JH, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. *Nature* 2023;616:520–524. DOI: 10.1038/s41586-023-05947-3. [PubMed: 37020027]
20. Caussy C, Alquraish MH, Nguyen P, et al. Optimal threshold of controlled attenuation parameter with MRI-PDFF as the gold standard for the detection of hepatic steatosis. *Hepatology* 2018;67:1348–1359. DOI: 10.1002/hep.29639. [PubMed: 29108123]
21. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI: IEEE, 2017:2261–2269.
22. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024;385:e078378. DOI: 10.1136/bmj-2023-078378. [PubMed: 38626948]
23. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*. Piscataway, NJ: IEEE, 2017:618–626. DOI: 10.1109/ICCV.2017.74.
24. Leung PB, Davis AM, Kumar S. Diagnosis and management of nonalcoholic fatty liver disease. *JAMA* 2023;330:1687–1688. DOI: 10.1001/jama.2023.17935. [PubMed: 37843854]
25. Pickhardt PJ. Value-added opportunistic CT screening: state of the art. *Radiology* 2022;303:E41. DOI: 10.1148/radiol.229010. [PubMed: 35604845]
26. Sandhu AT, Rodriguez F, Ngo S, et al. Incidental coronary artery calcium: opportunistic screening of previous nongated chest computed tomography scans to improve statin rates (NOTIFY-1 Project). *Circulation* 2023;147:703–714. DOI: 10.1161/CIRCULATIONAHA.122.062746. [PubMed: 36342823]
27. Harrison SA, Bedossa P, Guy CD, et al. A phase 3, randomized, controlled trial of resmetirom in NASH with liver fibrosis. *N Engl J Med* 2024;390:497–509. DOI: 10.1056/NEJMoa2309000. [PubMed: 38324483]
28. Newsome PN, Buchholtz K, Cusi K, et al. A placebo-controlled trial of subcutaneous semaglutide in nonalcoholic steatohepatitis. *N Engl J Med* 2021;384:1113–1124. DOI: 10.1056/NEJMoa2028395. [PubMed: 33185364]
29. Sanyal AJ, Bedossa P, Fraessdorf M, et al. A phase 2 randomized trial of survodutide in MASH and fibrosis. *N Engl J Med* 2024;391:311–319. DOI: 10.1056/NEJMoa2401755. [PubMed: 38847460]
30. Tapper EB, Parikh ND. Diagnosis and management of cirrhosis and its complications: a review. *JAMA* 2023;329:1589–1602. DOI: 10.1001/jama.2023.5997. [PubMed: 37159031]

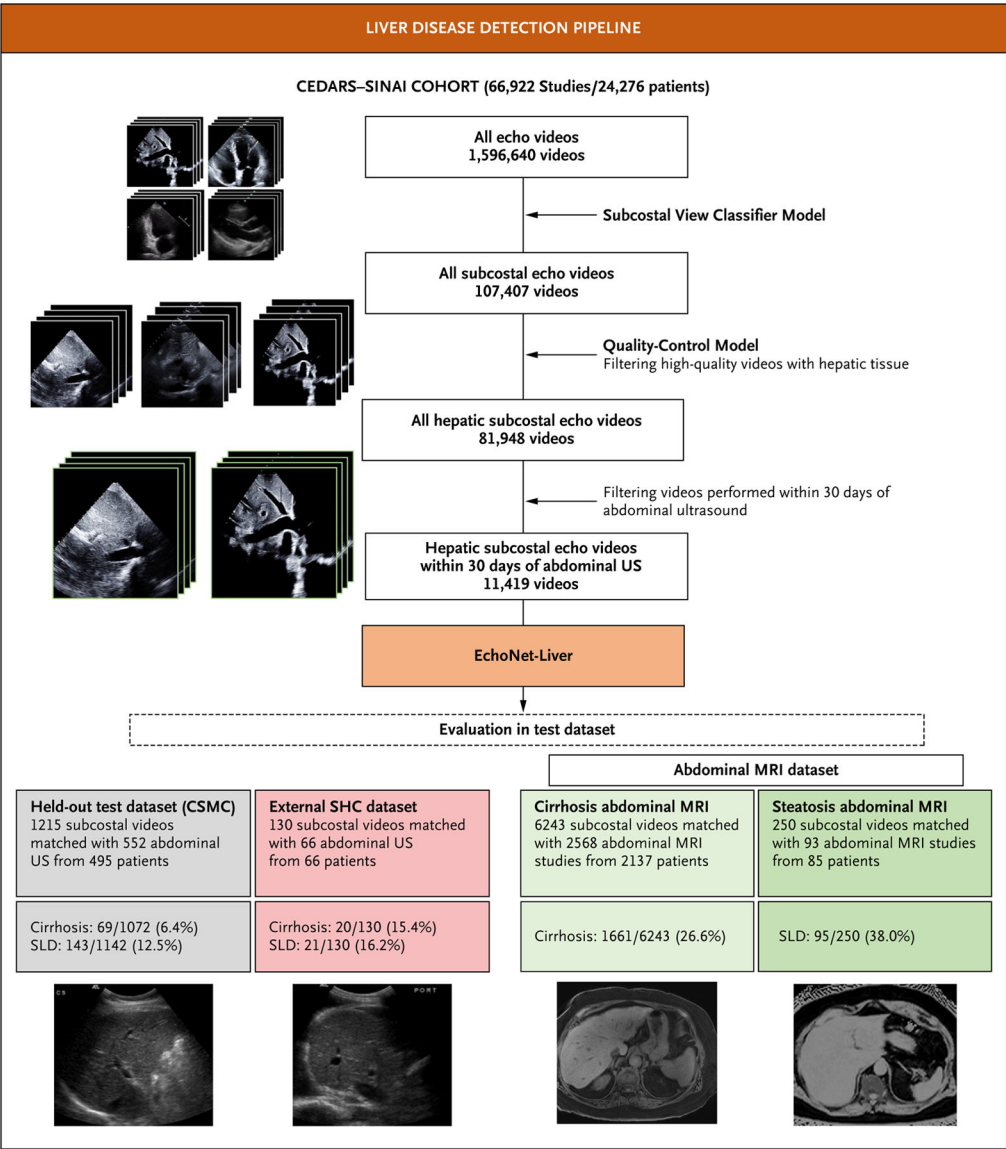


Figure 1. Overview of the Study Pipeline. About 1.6 million echocardiogram videos from 66,992 patients were used to train EchoNet-Liver, an automated pipeline that includes deep-learning–based view classification, image quality assessment, and detection of chronic liver disease. Evaluation of EchoNet-Liver was performed using held-out test cohorts. These cohorts included patients with paired echocardiograms and abdominal ultrasounds (CSMC, gray), an ultrasound external test dataset (SHC, red), and abdominal MRI-based diagnosis (CSMC, light green for cirrhosis, dark green for SLD). Case counts are provided on a per-video basis for each test subcohort. CSMC denotes Cedars–Sinai Medical Center; MRI, magnetic resonance imaging; SHC, Stanford Health Care; and SLD, steatotic liver disease.

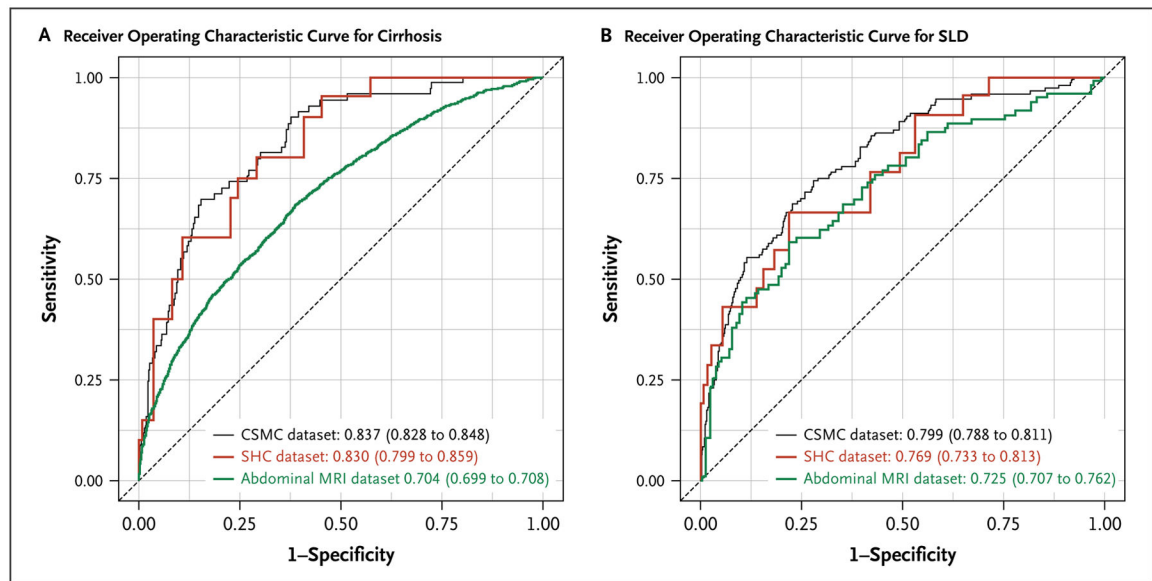


Figure 2. Model Performance of EchoNet-Liver

Performance of a deep-learning model using high-quality subcostal echocardiography videos for cirrhosis (Panel A) and SLD (Panel B). The model was evaluated in an internal CSMC held-out test dataset (black lines), an external SHC abdominal ultrasound dataset (red lines), and an abdominal MRI test dataset (green lines). Performance for cirrhosis was similar between the two ultrasound-labeled test cohorts (CSMC and SHC), but appeared worse for MRI, whereas the performance was similar across all three test cohorts for SLD. Figures for each dataset show the area under the receiver operating characteristic curve, with 95% confidence intervals in parentheses. CSMC denotes Cedars–Sinai Medical Center; MRI, magnetic resonance imaging; SHC, Stanford Health Care; SLD, steatotic liver disease; and US, ultrasound.

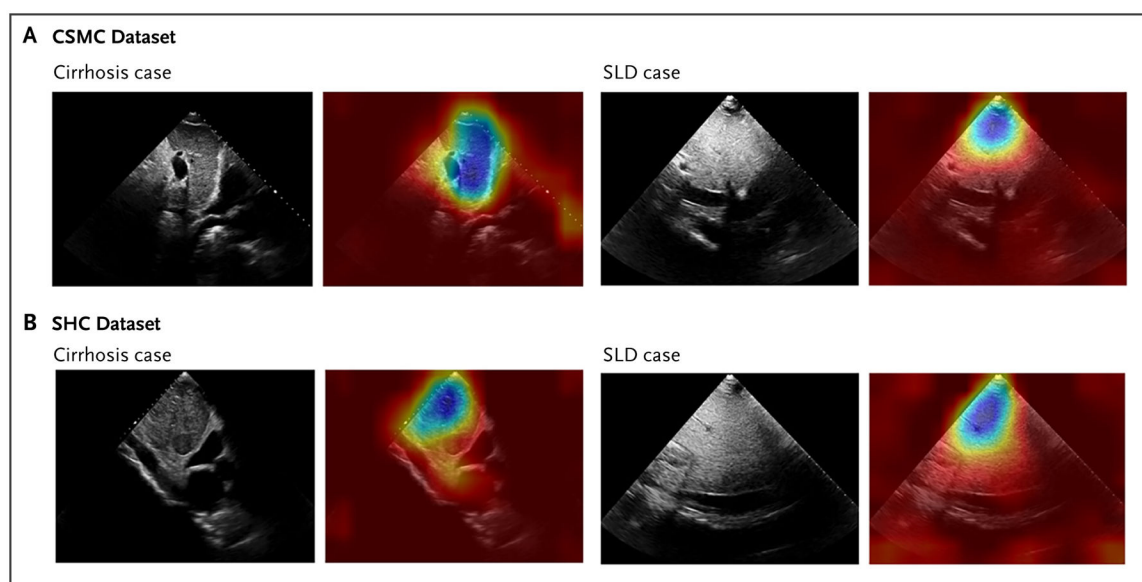


Figure 3. Representative Saliency Maps for the Two Test Datasets

Corresponding pairs of input echocardiogram frames and gradient-weighted class activation mapping visualizations of both cirrhosis and SLD from the CSMC dataset (Panel A) and the SHC dataset (Panel B). Blue shading represents regions or pixels that were more heavily weighted by the detection algorithm, whereas red-shaded pixels were less heavily weighted. The results demonstrate higher weighting or attention toward regions of liver parenchyma, and lower weighting in regions outside the liver or field of view. CSMC denotes Cedars–Sinai Medical Center; and SHC, Stanford Health Care.

Table 1.

Patient Characteristics of the Overall Study Cohorts*

Characteristic	CSMC Cohort (N=24,276)	CSMC Ultrasound Test Cohort (N=1486)	CSMC MRI Test Cohort (N=2335)	SHC External Validation Cohort (N=106)
Studies, n	66,922	6996	6751	106
Video/image files, n	1,596,640	163,736	164,579	5280
Age (years), mean (SD)	65.1 (17.1)	63.0 (17.3)	62.7 (14.6)	59.3 (16.8)
BMI, mean (SD)	27.1 (6.6)	26.9 (6.6)	26.9 (6.2)	27.0 (7.9)
Female sex, n (%)	11,892 (49.1)	666 (44.9)	1200 (51.4)	48 (45.3)
Race, n (%) [†]				
White	16,347 (67.3)	972 (65.4)	1572 (67.3)	49 (46.2)
Black	3962 (16.3)	279 (18.8)	335 (14.3)	2 (1.9)
Asian	1699 (7.0)	95 (6.4)	184 (7.9)	21 (19.8)
Other	2268 (9.3)	140 (9.4)	244 (10.5)	34 (32.0)
Comorbidity, n (%)				
Hypertension	6989 (28.8)	454 (30.6)	549 (23.5)	72 (67.9)
Dyslipidemia	4879 (20.1)	328 (22.1)	367 (15.7)	45 (42.5)
Diabetes	4675 (19.3)	310 (20.9)	449 (19.2)	43 (40.6)
Stroke	1626 (6.7)	113 (7.6)	107 (4.6)	7 (6.6)
Atrial fibrillation	802 (3.3)	71 (4.8)	42 (1.8)	36 (34.0)
Heart failure	2964 (12.2)	276 (18.6)	105 (4.5)	47 (44.3)
Coronary artery disease	4517 (18.6)	342 (23.0)	259 (11.1)	29 (27.4)
LVEF, mean (SD)	58.1 (14.9)	56.5 (16.5)	63.0 (10.1)	55.4 (13.8)
Active smoking	1208 (5.0)	100 (6.7)	101 (4.3)	4 (3.8)
HBV	284 (1.2)	15 (1.0)	72 (3.1)	9 (8.5)
HCV	908 (3.7)	53 (3.6)	201 (8.6)	9 (8.5)
Regular alcohol consumption	1243 (5.1)	92 (6.2)	211 (9.0)	26 (24.5)

* BMI denotes body mass index (the weight in kilograms divided by the square of the height in meters); CSMC, Cedars-Sinai Medical Center; HBV, hepatitis B virus; HCV, hepatitis C virus; LVEF, left ventricular ejection fraction; MRI, magnetic resonance imaging; SD, standard deviation; and SHC, Stanford Health Care.

[†] Race was self-reported by patients.

Table 2. Prediction of Liver Disease by Deep-Learning Analysis of Echocardiography Using Labels from Abdominal Ultrasound Held-Out Test Population and External Dataset *

Cohort	Prevalence	AUROC	AUPRC	Sensitivity	Specificity	PPV	NPV
CSMC held-out test cohort							
Cirrhosis	6.4%	0.837 (0.828 to 0.847)	0.305 (0.259 to 0.329)	0.696 (0.672 to 0.726)	0.846 (0.841 to 0.853)	0.238 (0.215 to 0.253)	0.976 (0.974 to 0.979)
SLD	12.5%	0.799 (0.788 to 0.811)	0.404 (0.372 to 0.429)	0.741 (0.722 to 0.763)	0.720 (0.712 to 0.728)	0.275 (0.259 to 0.288)	0.951 (0.948 to 0.956)
SHC external validation test cohort							
Cirrhosis	15.4%	0.830 (0.799 to 0.858)	0.496 (0.416 to 0.578)	0.800 (0.733 to 0.842)	0.709 (0.686 to 0.740)	0.333 (0.256 to 0.381)	0.951 (0.944 to 0.963)
SLD	16.2%	0.769 (0.733 to 0.813)	0.516 (0.423 to 0.579)	0.667 (0.611 to 0.765)	0.780 (0.760 to 0.830)	0.368 (0.294 to 0.452)	0.924 (0.915 to 0.953)

* Figures in parentheses are 95% confidence intervals. AUROC denotes area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; CSMC, Cedars-Sinai Medical Center; NPV, negative predictive value; PPV, positive predictive value; SHC, Stanford Health Care; and SLD, steatotic liver disease.