1	Opportunistic Screening of Chronic Liver Disease with Deep Learning Enhanced		
2	Echocardiography		
3			
4	Short title: Deep learning liver disease detection from echocardiography		
5			
6	Authors and affiliations:		
7	Yuki Sahashi MD, MSc ¹ ; Milos Vukadinovic BS ^{1,2} , Fatemeh Amrollahi PhD ³ , Hirsh		
8	Trivedi MD ⁴ , Justin Rhee BS ⁵ , Jonathan Chen MD, PhD ³ , Susan Cheng MD, MMSc,		
9	MPH^1 , David Ouyang $MD^{1,6}$, Alan C. Kwan MD^1 .		
10			
11	1. Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center,		
12	Los Angeles, CA		
13	2. Department of Bioengineering, University of California Los Angeles, Los		
14	Angeles, CA		
15	3. Bioinformatics Research, Department of Medicine, Stanford University, Palo		
16	Alto, CA		
17	4. Karsh Division of Gastroenterology and Hepatology, Department of Medicine,		
18	Cedars-Sinai Medical Center, Los Angeles, CA		
19	5. School of Medicine, Brown University, Providence, RI		
20	6. Division of Artificial Intelligence in Medicine, Cedars-Sinai Medical Center,		
21	Los Angeles, CA		
22			
23			
24	Keywords: Opportunistic screening, Echocardiography, Deep Learning, Convolutional		
25	neural network, Cirrhosis, Steatotic Liver Disease, Non-Alcoholic Fatty Liver Disease		
26			
27	Correspondence: <u>David.ouyang@cshs.org</u> , <u>alan.kwan@cshs.org</u>		
28	127 S. San Vicente Blvd, AHSP A3600		
29	Los Angeles, CA 90048		
30			
31	Word Count: 2,288 words		

32 **KEY POINTS:**

Question: Can a deep learning algorithm applied to echocardiography videos
 effectively identify chronic liver diseases including cirrhosis and steatotic liver disease
 (SLD)?

36

Findings: This retrospective observational cohort study utilized 1,596,640 echocardiography videos from 66,922 studies of 24,276 patients. The deep learning model with a computer vision pipeline (EchoNet-Liver) demonstrated strong performance to detect cirrhosis and SLD. External validation at a geographically distinct site demonstrated similar discriminative ability.

42

43 Meaning: The application of EchoNet-Liver to echocardiography could aid
44 opportunistic screening of chronic liver diseases, providing a unique cost-effective
45 angle to improve patient management.

- 46
- 47

48

49 Abbreviations:

- 50 AI : Artificial intelligence
- 51 AUROC: Area under receiver operating characteristic
- 52 AUPRC: Area Under precision recall curve

53 CLD: chronic liver disease

54 CVD: cardiovascular disease

55 LVEF: Left ventricular ejection fraction

56 PR: Precision-recall

57 SLD: steatotic liver disease

- 58 PPV: positive predictive value
- 59 NPV: negative predictive value
- 60 MRI: magnetic resonance imaging
- 61 MRS: magnetic resonance spectroscopy
- 62 PDFF: proton density fat fraction

64 ABSTRACT

- 65 Importance: Chronic liver disease affects more than 1.5 billion adults worldwide,
- 66 however the majority of cases are asymptomatic and undiagnosed. Echocardiography is
- ⁶⁷ broadly performed and visualizes the liver; but this information is not leveraged.
- 68 **Objective:** To develop and evaluate a deep learning algorithm on echocardiography
- 69 videos to enable opportunistic screening for chronic liver disease.
- 70 **Design:** Retrospective observational cohorts
- 71 **Setting:** Two large urban academic medical centers
- 72 **Participants:** Adult patients who received echocardiography and abdominal imaging

(either abdominal ultrasound or abdominal magnetic resonance imaging) with ≤ 30 days

- ⁷⁴ between tests, between July 4, 2012, to June 4, 2022.
- 75 Exposure: Deep learning model predictions from a deep-learning computer vision
- 76 pipeline that identifies subcostal view echocardiogram videos and detects the presence
- 77 of cirrhosis or steatotic liver disease (SLD).
- Main Outcome and Measures: Clinical diagnosis by paired abdominal ultrasound or
 magnetic resonance imaging (MRI).
- Results: A total of 1,596,640 echocardiogram videos (66,922 studies from 24,276 80 81 patients) from Cedars-Sinai Medical Center (CSMC) were used to develop EchoNet-Liver, an automated pipeline that identifies high quality subcostal images from 82 83 echocardiogram studies and detects the presence of cirrhosis or SLD. In the held-out 84 CSMC test cohort, EchoNet-Liver was able to detect the presence of cirrhosis with an AUC of 0.837 (0.789 - 0.880) and SLD with an AUC of 0.799 (0.758 - 0.837). In a 85 86 separate test cohort with paired abdominal MRIs, cirrhosis was detected with an AUC of 0.704 (0.689-0.718) and SLD was detected with an AUC of 0.726 (0.659-0.790). In 87 88 an external test cohort of 106 patients (n = 5,280 videos), the model detected cirrhosis with an AUC of 0.830 (0.738 - 0.909) and SLD with an AUC of 0.768 (0.652 – 0.875). 89
- 90 Conclusions and Relevance: Deep learning assessment of clinical echocardiography 91 enables opportunistic screening of SLD and cirrhosis. Application of this algorithm may 92 identify patients who may benefit from further diagnostic testing and treatment for 93 chronic liver disease.

94 Introduction

Chronic liver disease (CLD) affects an estimated 1.5 billion people worldwide 95 96 and 100 million in the United States and can result in malignancy, end-stage liver disease, or mortality¹. The prevalence of chronic liver disease is sharply increasing^{2,3}, 97 particularly related to steatotic liver disease (SLD) as a result of increased burden of 98 obesity and metabolic disease, however the vast majority of patients are undiagnosed^{4,5}. 99 This issue affects many individuals with known cardiovascular disease as well as 100 individuals with even severe end-stage liver disease, such as cirrhosis, which is 101 frequently missed in the earlier stages of fibrosis⁴. 102

103 Multiple approaches are available for screening and diagnosis of chronic liver disease, including serological risk scores, qualitative and quantitative ultrasound and 104 magnetic resonance imaging (MRI), and invasive biopsy^{6,7}. However, accuracy, 105 availability, and cost all limit the sufficiency of these pathways to address 106 107 underdiagnosis. Echocardiography, or ultrasound of the heart and associated structures, is a first-line diagnostic ultrasound test, and is frequently performed across the spectrum 108 of patients with metabolic and cardiovascular diseases⁸. Included within a standard 109 echocardiographic examination are subcostal views which visualize the inferior vena 110 cava and provide clear visualization of hepatic tissue quality as well as liver contour. 111 The full clinical utility of these images for identification of hepatic disease is 112 underutilized as cardiologists are not trained in the assessment of liver pathologies by 113 ultrasound. 114

Artificial intelligence (AI) can identify diseases and characteristics that may 115 not be readily observable by the human eye^{9-13} , predict disease progression¹⁴, 116 mortality¹⁵, and improve measurement accuracy of cardiac parameters^{16–19}. Our study 117 aims to develop and validate an AI computer vision approach to leverage 118 echocardiographic images and videos to detect CLD. We hypothesize that a 119 deep-learning pipeline can identify high-quality subcostal view videos and detect SLD 120 and cirrhosis in a high-throughput fashion. We trained and evaluated the model 121 122 performance in internal and external validations across multiple cohorts. AI-based analysis of hepatic tissue visible within standard subcostal echocardiography videos 123

124 may provide opportunistic screening of chronic liver disease from standard

125 echocardiography without additional costs.

127 Methods

128 Cohort Selection

129 We included adult patients over 18 years who received an echocardiogram at Cedars-Sinai Medical Center (CSMC) within 30 days of abdominal ultrasound and no 130 history of liver transplant between July 2012 and June 2022. Clinical diagnoses from 131 the abdominal ultrasound report including normal liver, steatotic liver, and cirrhotic 132 liver, were paired with the echocardiogram images as labels for training and validation. 133 In addition, we created an independent test cohort of patients with echocardiography 134 and a matched abdominal MRI for cross-modal validation. Studies from patients 135 included within any test cohort (ultrasound or MRI) were excluded from training cohort 136 (Supplementary Tables 1 - 4). A cohort of patients from Stanford Healthcare (SHC) 137 138 who had echocardiography and abdominal ultrasound within 30 days were included as external validation. An overall flowchart for model development and evaluation is 139 provided in Figure 1. Approval for this study was obtained from the Cedars-Sinai 140 Medical Center and Stanford Healthcare Institutional Review Boards, and the 141 requirement for informed consent was waived for retrospective data analysis without 142 143 patient contact.

144

145 **Disease Definitions**

146

Echocardiography studies were matched to the closest abdominal ultrasound 147 study within 30 days. CLD labels were derived from the abdominal ultrasound reports 148 149 and categorized studies into normal, cirrhosis, or SLD based on text describing the liver 150 parenchyma. If an echocardiography study had greater than one abdominal ultrasound 151 within 30 days, labels were taken from the radiologist confirmed clinical report closest in time. Controls were identified when the report specified the liver parenchyma was 152 153 normal In the abdominal MRI group, diagnosis labels for cirrhosis was taken from the clinical MRI report; for SLD, we focused on the subset of cases where magnetic 154 155 resonance imaging-derived proton density fat fraction (MRI-PDFF) or magnetic resonance spectroscopy (MRS) fat signal was measured, with larger than 5.0% being 156 diagnostic of SLD²⁵. 157

159 View Selection and Image Processing

A standard echocardiogram study often contains 50-100 videos, of which 160 typically only 1-2 videos per study capture the liver in sufficient quality for assessment 161 of the liver echotexture and contour. Echocardiography videos were initially obtained as 162 163 Digital Imaging and Communications in Medicine (DICOM) files and underwent 164 de-identification and processing into AVI videos. We developed a pipeline of two deep learning models for view classification (identification of the subcostal view videos) and 165 quality control (excluding videos with severe motion artifact and low image quality). 166 Two video-based convolutional neural networks $(R2+1D)^{23}$ were used with standardized 167 112×112 pixel videos for input. A dataset of 11,778 manually curated videos of 4,991 168 patients was used model training, classified by both view type and image quality 169 (representative ground truth images in **Supplemental Figure 1**). To evaluate the 170 performance of view selection and quality checking, we randomly selected 100 CSMC 171 172 echocardiogram studies (n = 2,315 videos) and two cardiologists manually identified high quality subcostal and low-quality subcostal view videos for comparison with 173 model output. 174

175

176 Chronic Liver Disease Detection

177 For the training and evaluation of the liver disease detection models, we utilized an image-based model (DenseNet²⁶), to focus on hepatic tissue texture for the 178 prediction of cirrhosis and SLD. Given the focus on texture, the input data were still 179 frame images from high quality subcostal echocardiogram videos at native resolution 180 $(480 \times 640 \text{ pixels})$. We trained the model to minimize binary cross-entropy loss using an 181 AdamW optimizer with an initial learning rate of 1×10^{5} at a batch size of 40 for 100 182 183 epochs. A variety of hyperparameters and architectures were compared prior to development of the final EchoNet-Liver model (Supplemental Table 5). All model 184 training and evaluation were conducted using Python 3.8, PyTorch 2.2, and torchvision 185 0.17. 186

187

188 **Statistical analysis**

189 Model performance was determined by measuring the area under the receiver 190 operating curve (AUC) on held out test cohorts. Area under the precision-recall curves

191 (AUPRC), sensitivity, specificity, positive predictive value, and negative predictive value were similarly reported at the Youden index. All 95% confidence intervals were 192 193 calculated with 10,000 bootstrapping samples. Data analysis was performed using both Python (version 3.8.0) and R (version 4.2.2) programming languages. This study was 194 carried out following the TRIPOD-AI guideline (Supplemental Material 1)²⁷. Saliency 195 maps were generated to identify the areas of interest for the classifier across all test 196 datasets. Each saliency map was produced using Grad-CAM²⁸, which captures the 197 gradient information directed into the final convolutional layer of the trained deep 198 learning model. We input the final layer of the fourth DenseBlock for this approach. 199 200

201 Code and Data Availability

202 Code is available at <u>https://github.com/echonet/liver/</u> and training data is 203 available with submission of a research protocol and approval by CSMC and SHC 204 IRBs.

206 **RESULTS**

207 **Patient characteristics**

208 Using a total of 1,596,640 videos from 66,922 CSMC echocardiography studies of 24.276 patients were identified and split 8:1:1 by patient into training, 209 validation, and held-out test cohorts. The patient cohorts exhibited a variety of 210 comorbidities consistent with patients that receive both echocardiography and 211 abdominal ultrasound studies with prevalent hypertension (28.8%), hyperlipidemia 212 (20.1%), diabetes (19.3%), hepatitis B (1.2%) and hepatitis C (3.7%) (Table 1). The 213 average BMI was 26.5 ± 6.2 , and 5.3% of the patients regularly consumed alcohol. The 214 215 median duration between the echocardiography and abdominal ultrasound examinations was 0 days (interquartile range, -4 to +3 days). In this cohort, there were 371 (8.8%) 216 217 cirrhosis and 645 (14.0%) SLD cases based on abdominal ultrasound reports.

We identified an additional test population of 6,959 echocardiogram videos 218 from 4,145 studies across 2,335 CSMC patients who underwent abdominal MRI for 219 cirrhosis and MRI with PDFF/MRS for SLD, and echocardiography within 365 days. 220 The patients in this additional test cohort were not included in training and validation 221 cohort of disease detection models. In the SHC cohort for EchoNet-Liver evaluation, we 222 223 identified 66 studies of individual patients who have a total of 130 high-quality subcostal videos. Of all patients, 10 and 11 patients were diagnosed with cirrhosis and 224 225 SLD respectively based on abdominal ultrasound. Patient characteristics are presented in Table 1. Patient characteristics in each model development (i.e., view-classifier 226 227 model, quality-control model, cirrhosis detection model and SLD detection model) were 228 demonstrated in Supplemental Tables 1-4.

- 229
- 230

231 <u>View Classifier and Quality Control Performance Across Two Institutions</u>

A total 11,419 subcostal view videos were used for training a subcostal view classifier model, with all other videos labeled as non-subcostal controls. On an independent test set of 100 CSMC studies (n = 2,315 videos), the view classifier model identified 186 out of 196 subcostal view videos with an AUC of 0.991 (0.984 – 0.997). The sensitivity was 0.949 (0.910-0.974) and specificity was 0.999 (0.998-1.00). In the SHC population, the subcostal view classifier identified 149 out of 228 subcostal videos

from a total of 5,280 total videos with an AUC of 0.965 (0.956 - 0.974). The sensitivity was 0.654 (0.591 - 0.716) and specificity was 0.993 (0.991 - 0.995). In comparison of image quality by two cardiologists, the quality-control model demonstrated an AUC of 0.855 (0.800-0.905) in CSMC and an AUC of 0.785 (0.722-0.843) in SHC.

242

243 Disease Detection Performance Across Two Institutions

244 In the held-out CSMC test ultrasound dataset, the EchoNet-Liver detected chronic liver disease with an AUC of 0.837 (95% CI 0.789 - 0.880) for cirrhosis and 245 0.799 (0.758 - 0.837) for SLD (Figure 2). For cirrhosis, the AUPRC was 0.309 (0.206 -246 0.417), PPV was 0.238 (0.180 - 0.299), and NPV was 0.976 (0.965 - 0.986). For SLD, 247 the AUPRC was 0.408 (0.325 - 0.491), the PPV was 0.275 (0.232 - 0.319) and the NPV 248 249 was 0.951 (0.935 - 0.966). In the CSMC test cohort, the algorithm showed a sensitivity of 0.696 (0.582 - 0.803) and a specificity of 0.847 (0.824 - 0.868) for detecting of 250 cirrhosis and sensitivity of 0.741 (0.669 - 0.812) and specificity of 0.720 (0.692 - 0.747) 251 for detecting SLD. In the SHC cohort, EchoNet-Liver detected cirrhosis and SLD with 252 an AUC of 0.830 (0.738 - 0.909) and 0.768 (0.652 - 0.875) respectively. For cirrhosis, 253 the AUPRC was 0.500 (0.278 - 0.708), PPV was 0.334 (0.203 - 0.471), and NPV was 254 255 0.952 (0.901 - 0.989). For SLD, the AUPRC was 0.518 (0.299 - 0.705), the PPV was 0.370 (0.219 - 0.528) and the NPV was 0.924 (0.864 - 0.976). In the SHC test cohort, 256 EchoNet-Liver demonstrated a sensitivity of 0.802 (0.611 - 0.957) and a specificity of 257 0.709 (0.623 - 0.789) for detecting of cirrhosis and sensitivity of 0.667 (0.450 - 0.867) 258 259 and specificity of 0.781 (0.701 - 0.855) for detecting SLD. Evaluation metrics are 260 summarized in Table 2.

261

262 Comparison with Diagnosis by Magnetic Resonance Imaging

In order to evaluate the performance of EchoNet-Liver across multiple diagnostic pathways, the algorithm was evaluated in a cohort of patients that both received echocardiography and abdominal MRI at CSMC. In the MRI paired cohort, EchoNet-Liver detected the presence of cirrhosis with an AUC of 0.704 (95% CI 0.689–0.718) (**Figure 2-A**). and a AUPRC of 0.493 (95% CI 0.468–0.518) (**Supplementary Figure 2-A**). When evaluated in the MRS/PDFF steatosis cohort, the model detected the presence of SLD with an AUC of 0.726 (95% CI 0.659–0.790)

270 (Figure 2-B) and an AUPRC of 0.626 (95%CI 0.519–0.731) (Supplementary Figure

271 **2-B**)

272

273 Model Explainability

274 We generated saliency maps for representative echocardiography images from

the two test datasets (CSMC held-out dataset and SHC external dataset shown in Figure

276 3). For both cirrhosis and SLD, EchoNet-Liver highlighted the liver of the subcostal

277 echocardiographic images as regions of interest, with a diffuse activation throughout the

278 hepatic parenchyma.

280 Discussion

281 In this study, we demonstrated strong performance of a deep learning pipeline 282 (EchoNet-Liver) for detecting cirrhosis and SLD from clinical echocardiography images. The discriminative ability of the model was confirmed in a geographically distinct 283 284 external health-care system cohort as well as in a cohort of patients with paired abdominal MRI imaging. Across diverse populations and disease definitions, deep 285 learning-enhanced echocardiography enabled high-throughput automated detection of 286 chronic liver disease, which could enable opportunistic screening for asymptomatic 287 liver disease. 288

Chronic liver disease often remains undiagnosed due to the asymptomatic 289 nature of early disease. Despite the significant prevalence and morbidity, routine 290 screening is not recommended given the high cost of imaging and lack of evidence of its 291 cost-effectiveness²³. Opportunistic screening from echocardiogram images can identify 292 a high-risk population of patients with concurrent cardiovascular risk and liver disease 293 in a cost-effective manner. By harnessing pre-existing imaging indicated for other 294 diagnostic reason, our AI-enhanced workflow can increase the potential utility of 295 imaging examinations²⁹. By incorporating view classification, quality control, and 296 disease detection in one pipeline, automation with AI can enable high-throughput 297 evaluation of this non-invasive and common cardiovascular diagnostic test. 298

299 There are several limitations in the present study to consider. This study is a 300 retrospective study conducted at two tertiary care centers for patients who have 301 undergone both abdominal ultrasound and echocardiography, which may result in 302 selection bias. The spectrum of age, gender, race, and comorbidities in the study dataset may not represent the general population and may bias towards patients with more 303 304 comorbidity, necessitating further external validation. Comparison between disparate imaging modalities such as ultrasound and MRI will have inherent limitations due to 305 306 different modality-specific accuracy. However, in totality this study suggests that the clinical utility of high throughput disease screening using AI is promising, particularly 307 for early disease, and enhances the utility of pre-existing imaging data^{11,12,30}. Further 308 309 studies are warranted to establish the optimal clinical workflow for opportunistic liver disease screening among CVD patients and downstream treatment. By improving 310

diagnosis of subclinical CLD, we may be able to limit or reverse disease progression $^{20-22}$ and improve patient care by triaging patients toward appropriate clinical and diagnostic management^{23,24}.

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 with comparison to abdominal magnetic resonance imaging. Deep learning applied to echocardiography may offer an opportunity for opportunistic and cost-effective screening for chronic liver disease.

- 320
- 321

322 Disclosures

YS reports support from the KAKENHI (Japan Society for the Promotion of Science: 323 24K10526), oversea research grant from SUNRISElab, Japan Heart Foundation and 324 Ogawa Foundation and honoraria for lectures from m3.com inc. DO reports support 325 from the National Institute of Health (NIH; NHLBI R00HL157421 and R01HL173526) 326 and Alexion, and consulting or honoraria for lectures from EchoIQ, Ultromics, Pfizer, 327 InVision, the Korean Society of Echocardiography, and the Japanese Society of 328 Echocardiography. ACK reports consulting fees from InVision and support from the 329 330 American Heart Association (AHA; 23CDA1053659) and National Institutes of Health (NIH; UL1TR001881). 331

- 332
- 333

334 Author contributions

- 335 Concept and design: YS, DO, AK
- 336 Code and programming: YS, MV, JR
- 337 Acquisition, analysis, or interpretation of data: YS, DO, AK, FA
- 338 Drafting and Critical revision of the manuscript: YS, DO, AK, HT, SC JC
- 339 Statistical analysis: YS, DO, AK
- 340 Obtained funding: AK, DO, SC
- 341 Supervision: AK, DO, SC

<u>Tables</u>

Table 1: Patient characteristics of the study cohort

Characteristic	CSMC cohort,	CSMC Ultrasound	CSMC MRI Test	SHC External
	N = 24,276	Test cohort	cohort,	Validation cohort,
		N = 1,486	N = 2,335	N = 106
Number of studies	66,922	6,996	6,751	106
Number of video files	1,596,640	163,736	164,579	5,388
Age, 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)
Gender, female	11,892 (49.1%)	666 (44.9%)	1,200 (51.4%)	48 (45.3%)
Race/Ethnicity				
White	16,347 (67.3%)	972 (65.4%)	1,572 (67.3%)	49 (46.2%)
Black	3,962 (16.3%)	279 (18.8%)	335 (14.3%)	2 (1.9%)
Asian	1,699 (7.0%)	95 (6.4%)	184 (7.9%)	21 (19.8%)
Other	2,268 (9.3%)	140 (9.4%)	244 (10.5%)	34 (32.0%)
Hypertension	6,989 (28.8%)	454 (30.6%)	549 (23.5%)	72 (67.9%)
Dyslipidemia	4,879 (20.1%)	328 (22.1%)	367 (15.7%)	45 (42.5%)
Diabetes	4,675 (19.3%)	310 (20.9%)	449 (19.2%)	43 (40.6%)
Stroke	1,626 (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	2,964 (12.2%)	276 (18.6%)	105 (4.5%)	47 (44.3%)
Coronary Artery	4,517 (18.6%)	342 (23.0%)	259 (11.1%)	29 (27.4%)
Disease				

LVEF, mean (SD)	58.1 (14.9)	56.5 (16.5)	63.0 (10.1)	55.4 (13.8)
Active Smoking	1,208 (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%)
Alcohol	1,243 (5.1%)	92 (6.2%)	211 (9.0%)	26 (24.5%)

346

347 LVEF: Left Ventricular Ejection Fraction; BMI: Body Mass Index; SD: Standard Deviation; SLD: Steatotic liver disease,

348 HBV: hepatitis B, HCV: hepatitis C

349

	AUROC	AUPRC	Sensitivity	Specificity	PPV	NPV
(a) CSMC h	neld-out test cohor	<u>t_</u>				
Cirrhosis	0.837	0.309	0.696	0.847	0.238	0.976
	(0.789 - 0.880)	(0.206 – 0.417)	(0.582 - 0.803)	(0.824 - 0.868)	(0.180 - 0.299)	(0.965 - 0.986)
SLD	0.799	0.408	0.741	0.720	0.275	0.951
	(0.758 - 0.837)	(0.325 – 0.491)	(0.669 - 0.812)	(0.692 - 0.747)	(0.232 - 0.319)	(0.935 - 0.966)
(b) SHC external validation test cohort						
Cirrhosis	0.830	0.500	0.802	0.709	0.334	0.952
	(0.738 - 0.909)	(0.278 – 0.708)	(0.611 - 0.957)	(0.623 - 0.789)	(0.203 - 0.471)	(0.901 - 0.989)

Table 2: Prediction of liver disease by deep learning analysis of echocardiography using labels from abdominal ultrasound held-out test

352 population.

353 PPV: positive predictive value. NPV: positive predictive value. AUROC: Area under receiver operating characteristic. SLD: Steatotic

(0.450 - 0.867)

0.781

(0.701 - 0.855)

0.370

(0.219 - 0.528)

0.924

(0.864 - 0.976)

0.667

354 liver disease.

0.768

(0.652 - 0.875)

0.518

(0.299 - 0.705)

SLD

355

356 Figures

357 Figure 1: Overview of the study pipeline



358

More than 2M echocardiogram videos were used to train EchoNet-Liver, an automated pipeline for deep learning view classification, image quality assessment, and detection of chronic liver disease. Evaluation of EchoNet-Liver was performed 3 held-out test cohorts, including patients with paired echocardiograms and abdominal ultrasounds (CSMC), abdominal MRI (CSMC), and an ultrasound external test dataset (SHC). SLD: steatotic liver disease, MRI: magnetic resonance imaging.

365 Figure 2. Model Performance of Echo-Net-Liver



Performance of a deep learning model using high-quality subcostal echocardiography
videos for cirrhosis (A) and SLD (B). Model was evaluated in an internal CSMC
held-out test dataset (black), external SHC abdominal ultrasound dataset (red) and
abdominal MRI test dataset (green).

371 US: ultrasound, MRI: Magnetic resonance imaging, SLD: Steatotic liver disease

Figure 3: Representative saliency maps for the two test datasets

388



389

Corresponding pairs of input echocardiogram frames and Grad-CAM visualization of both cirrhosis and steatotic liver disease (SLD) (CSMC dataset (A) and Stanford Healthcare dataset (B)) Pixels brighter in color and closer to blue were more salient to model predictions, while those darker in color were less important to the model's final prediction.

395 **Reference**

- Moon AM, Singal AG, Tapper EB. Contemporary Epidemiology of Chronic Liver
 Disease and Cirrhosis. *Clin Gastroenterol Hepatol*. 2020;18(12):2650-2666.
- 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(1):102068.
- GBD 2017 Cirrhosis Collaborators. 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(3):245-266.
- 406 4. Bajaj JS, Silvey SG, Rogal S, et al. Undiagnosed Cirrhosis and Hepatic
 407 Encephalopathy in a National Cohort of Veterans With Dementia. *JAMA Netw Open*.
 408 2024;7(1):e2353965-e2353965.
- 409 5. Castera L, Laouenan C, Vallet-Pichard A, et al. High Prevalence of NASH and
 410 Advanced Fibrosis in Type 2 Diabetes: A Prospective Study of 330 Outpatients
 411 Undergoing Liver Biopsies for Elevated ALT, Using a Low Threshold. *Diabetes*412 *Care*. 2023;46(7):1354-1362.
- 413 6. Sterling RK, Patel K, Duarte-Rojo A, et al. AASLD Practice Guideline on
 414 blood-based non-invasive liver disease assessments of hepatic fibrosis and steatosis.
 415 *Hepatology*. Published online March 15, 2024.
 416 doi:10.1097/HEP.00000000000845
- 417 7. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, et al. AASLD Practice
 418 Guidance on the clinical assessment and management of nonalcoholic fatty liver
 419 disease. *Hepatology*. 2023;77(5):1797-1835.
- 420 8. Virnig BA, Shippee ND, O'Donnell B, Zeglin J, Parashuram S. *Trends in the Use of*421 *Echocardiography, 2007 to 2011.* Agency for Healthcare Research and Quality
 422 (US); 2014.
- Yuan N, Duffy G, Dhruva SS, et al. Deep Learning of Electrocardiograms in Sinus
 Rhythm From US Veterans to Predict Atrial Fibrillation. *JAMA Cardiol*. Published
 online October 18, 2023. doi:10.1001/jamacardio.2023.3701

- 426 10. Hughes JW, Yuan N, He B, et al. Deep learning evaluation of biomarkers from
 427 echocardiogram videos. *EBioMedicine*. 2021;73:103613.
- 428 11. Duffy G, Cheng PP, Yuan N, et al. High-Throughput Precision Phenotyping of Left
 429 Ventricular Hypertrophy With Cardiovascular Deep Learning. *JAMA Cardiol*.
 430 2022;7(4):386-395.
- 431 12. Vrudhula A, Duffy G, Vukadinovic M, Liang D, Cheng S, Ouyang D. High
 432 throughput deep learning detection of mitral regurgitation. *bioRxiv*. Published
 433 online February 12, 2024. doi:10.1101/2024.02.08.24302547
- 434 13. Goto S, Mahara K, Beussink-Nelson L, et al. Artificial intelligence-enabled fully
 435 automated detection of cardiac amyloidosis using electrocardiograms and
 436 echocardiograms. *Nat Commun.* 2021;12(1):2726.
- 437 14. Oikonomou EK, Holste G, Yuan N, et al. A Multimodal Video-Based AI Biomarker
 438 for Aortic Stenosis Development and Progression. *JAMA Cardiol*. Published online
 439 April 6, 2024. doi:10.1001/jamacardio.2024.0595
- Lau Emily S., Di Achille Paolo, Kopparapu Kavya, et al. Deep Learning–Enabled
 Assessment of Left Heart Structure and Function Predicts Cardiovascular
 Outcomes. J Am Coll Cardiol. 2023;82(20):1936-1948.
- 443 16. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of
 444 cardiac function. *Nature*. 2020;580(7802):252-256.
- 445 17. Kwan AC, Chang EW, Jain I, et al. Deep Learning-Derived Myocardial Strain.
 446 JACC Cardiovasc Imaging. Published online March 12, 2024.
 447 doi:10.1016/j.jcmg.2024.01.011
- 18. Tromp J, Seekings PJ, Hung CL, et al. Automated interpretation of systolic and
 diastolic function on the echocardiogram: a multicohort study. *The Lancet Digital Health*. 2022;4(1):e46-e54.
- 451 19. He B, Kwan AC, Cho JH, et al. Blinded, randomized trial of sonographer versus AI
 452 cardiac function assessment. *Nature*. 2023;616(7957):520-524.
- 20. Harrison Stephen A., Bedossa Pierre, Guy Cynthia D., et al. A Phase 3, Randomized,
 Controlled Trial of Resmetirom in NASH with Liver Fibrosis. *N Engl J Med.*2024;390(6):497-509.

456	21. Newsome Philip N., Buchholtz Kristine, Cusi Kenneth, et al. A Placebo-Controlled
457	Trial of Subcutaneous Semaglutide in Nonalcoholic Steatohepatitis. N Engl J Med.
458	2021;384(12):1113-1124.

- 459 22. Sanyal AJ, Bedossa P, Fraessdorf M, et al. A phase 2 randomized trial of
 460 survodutide in MASH and fibrosis. *N Engl J Med*. Published online June 7, 2024.
 461 doi:10.1056/NEJMoa2401755
- 462 23. Leung PB, Davis AM, Kumar S. Diagnosis and Management of Nonalcoholic Fatty
 463 Liver Disease. *JAMA*. 2023;330(17):1687-1688.
- 464 24. Tapper EB, Parikh ND. Diagnosis and Management of Cirrhosis and Its
 465 Complications: A Review. *JAMA*. 2023;329(18):1589-1602.
- 25. Caussy C, Alquiraish 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(4):1348-1359.
- 469 26. Huang G, Liu Z, Weinberger KQ. Densely connected convolutional networks. *Proc*470 *IEEE Comput Soc Conf Comput Vis Pattern Recognit*. Published online August 25,
 471 2016:2261-2269.
- 472 27. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated
 473 guidance for reporting clinical prediction models that use regression or machine
 474 learning methods. *BMJ*. 2024;385. doi:10.1136/bmj-2023-078378
- 28. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM:
 Visual Explanations from Deep Networks via Gradient-Based Localization. In:
 2017 IEEE International Conference on Computer Vision (ICCV). IEEE;
 2017:618-626.
- 479 29. Pickhardt PJ. Value-added Opportunistic CT Screening: State of the Art. *Radiology*.
 480 2022;303(3):E41.
- 30. 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(9):703-714.