Opportunistic Screening for Cirrhosis using Echocardiography: A Multi-Modal and Multi-Institutional Validation

Yuki Sahashi*[1], Dilek Yalcinkaya*[2], Alan C. Kwan[1], David Ouyang[1,2]
1.Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
2.Department of Cardiology, Kaiser Permanente Santa Clara Medical Center, Santa Clara, CA / *Equal contributor

Introduction

Chronic liver disease is frequently underdiagnosed

- An estimated **100 million people** in the United States have chronic liver disease, which can progress to malignancy and cirrhosis.
- While echocardiography is routinely performed for cardiovascular indications, the **liver images captured within the subcostal view** are not clinically leveraged for diagnosing chronic liver disease.
- We developed a **deep learning model to detect cirrhosis from standard echocardiography** for opportunistic screening, validated across three reference standards (AUS, VCTE, and liver biopsy) in two healthcare systems.

AUS: Abdominal Ultrasound
VCTE: Vibration-Controlled Transient Elastography

Methods

Study Design & Training

- An image-based deep learning model (DenseNet) was trained to identify cirrhosis using **AUS labels** from subcostal echocardiographic images at **Cedars-Sinai Medical Center (CSMC)**.
- Subcostal frames were automatically extracted from echocardiographic studies; images with visible liver parenchyma were selected as model input.
- Cirrhosis labels were derived from AUS reports performed within **30 days** of the index echocardiogram.

Training Hyperparameters.

A DenseNet-121 backbone was fine-tuned on single-frame subcostal images (640 × 480) using AdamW (lr = 1 × 10^{-4}), batch size 48, and binary cross-entropy with logits loss. The model was trained for up to 100 epochs with early stopping (patience = 10) on validation AUC. Standard echo augmentation (random affine: ± 10° rotation, ± 10% translation, 0.9 – 1.1× scaling) was applied during training. Patients were split at the MRN level into training, validation, and test sets (80 / 10 / 10).

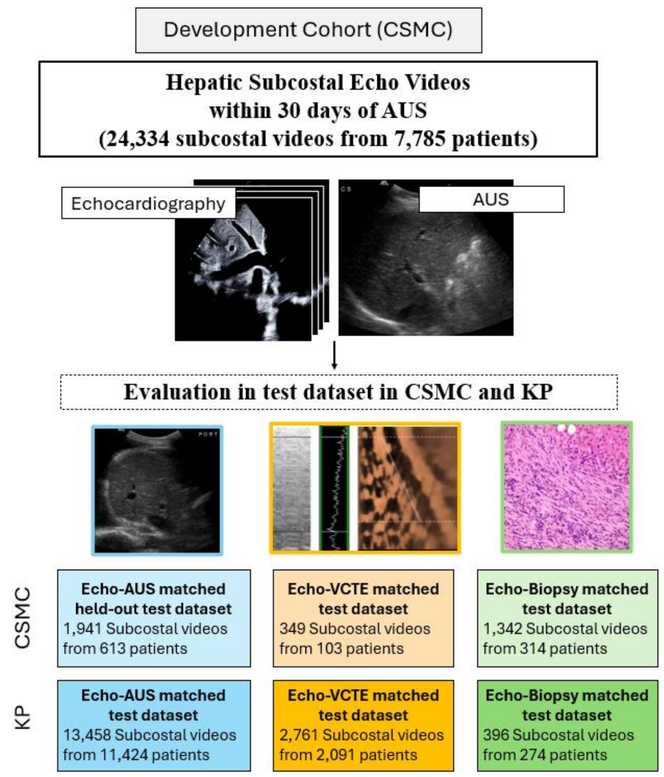
Validation

- Validated on independent test sets with diagnosis defined by three reference standards:
  - (1) **Abdominal ultrasound (AUS)**
  - (2) **Vibration-controlled transient elastography (VCTE)**
  - (3) **Liver biopsy** (histopathology)
- To assess robustness, **external validation** across all three modalities was performed at **Kaiser Permanente (KP)**.

Clinical Deployment Analysis

- Performance was analyzed at a high-specificity (95%) operating threshold to simulate a real-world screening deployment scenario.
- Number needed to screen (NNS) and positive predictive value (PPV) were calculated across a range of thresholds to identify the optimal operating point for downstream clinical decision-making.
- The trade-off between sensitivity and specificity was evaluated to optimize the balance between detecting true cirrhosis cases and minimizing unnecessary follow-up procedures.

Methods



Results

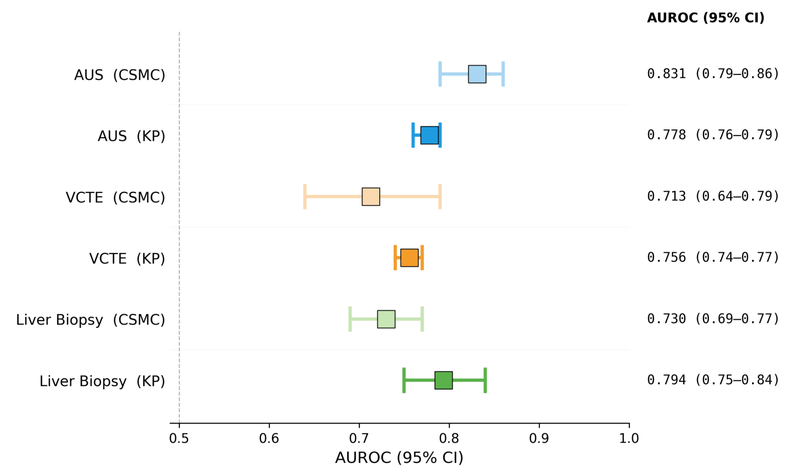
Study Population

- For development, **7,785 patients** underwent both transthoracic echocardiography and paired AUS within 30 days at CSMC.
- **14,379 echo studies** (**24,334 subcostal images**) paired with **9,529 AUS studies**.

Model Performance

- Robust AUROC across institutions and modalities:
  - **AUS:** 0.831 (CSMC); 0.778 (KP)
  - **VCTE:** 0.713 (CSMC); 0.756 (KP)
  - **Biopsy:** 0.730 (CSMC); 0.794 (KP)

Figure. Model Performance across reference standards and institutions



Deployment Simulation

- Clinical threshold of 0.807 was defined at a threshold where the specificity equal to 95%.
  - **VCTE:** specificity 98.7%, sensitivity 5.4%
  - **Liver biopsy:** specificity 95.5%, sensitivity 26.3%
- Assuming cirrhosis prevalence of **1.0–4.0%** in a general echo population:
  - **NNS:** 466–1,866 (VCTE); 95–381 (liver biopsy)
  - **PPV:** 4.1–15.0% (VCTE); 5.0–17.7% (liver biopsy)

Results

Table. Model Performance in CSMC Test Cohorts at Clinical Threshold (0.807).

Modality	AUROC	Sensitivity	Specificity	PPV	NPV	Accuracy
AUS	0.831 (0.79–0.86)	31.7%	95.0%	38.8%	93.3%	89.3%
VCTE	0.713 (0.64–0.79)	5.4%	98.7%	85.7%	42.4%	44.0%
Liver Biopsy	0.730 (0.69–0.77)	26.2%	95.5%	75.0%	71.5%	71.9%

Table. Number Needed to Screen (NNS) and 1/PPV at Assumed General-Population Cirrhosis Prevalence, using threshold = 0.807.

Cohort	Assumed Prevalence	PPV	NPV	NNS	1/PPV
AUS	1.0%	6.1%	99.3%	**315**	**16.5**
	2.0%	11.5%	98.6%	**158**	**8.7**
	3.0%	16.5%	97.8%	**105**	**6.1**
	4.0%	21.0%	97.1%	**79**	**4.8**
VCTE	1.0%	4.1%	99.1%	**1,866**	**24.4**
	2.0%	7.9%	98.3%	**933**	**12.6**
	3.0%	11.6%	97.4%	**622**	**8.6**
	4.0%	15.0%	96.5%	**466**	**6.7**
Liver Biopsy	1.0%	5.0%	99.2%	**381**	**20.2**
	2.0%	9.5%	98.5%	**190**	**10.5**
	3.0%	13.8%	97.7%	**127**	**7.3**
	4.0%	17.7%	96.9%	**95**	**5.7**

NNS = 1 / (Sensitivity × Prevalence). 1/PPV = 1 / Positive Predictive Value, interpreted as the number of test-positive individuals required to identify one true cirrhosis case at the given population prevalence.

Sensitivity values are taken from each modality test cohort at threshold 0.807.

Discussion

- Model robustness was confirmed across multiple institutions and against independent reference cohorts.
- Given the opportunistic screening context, a high-specificity threshold was selected to minimize false positives and avoid unnecessary downstream workup.
- A prospective trial is currently being planned to validate real-world clinical utility — specifically whether the model improves the diagnostic yield of cirrhosis.

Conclusion

Our findings demonstrate the potential of an echocardiography-guided AI model for the opportunistic detection of cirrhosis. By validating against multi-modal diagnosis standards across multi-institutional datasets, we confirmed the model's generalizability for real-world clinical use.

GitHub | Twitter/X: @Yuki_Sahashi | Linkedin

Code: https://github.com/echonet/liver