

Early Detection of Advanced Liver Disease Via
Artificial Intelligence-Enabled
Electrocardiogram (Advance): A Pragmatic,
Cluster-Randomized Clinical Trial

NCT05782283

3March2023

**DETECTION OF UNDIAGNOSED LIVER CIRRHOSIS VIA ARTIFICIAL
INTELLIGENCE-ENABLED ELECTROCARDIOGRAM (DULCE):A
PRAGMATIC, CLUSTER-RANDOMIZED CLINICAL TRIAL**

Version: 2

Date: March 3, 2023

Table of Contents

Abstract.....	3
1. Background and Objectives	5
1.1 Screening for Cirrhosis	5
1.2 Current Screening and Diagnosis Methods.....	5
1.3 Development of an Artificial Intelligence (AI)-Enabled ECG-Based Screening Tool	6
1.4 Study Objectives.....	8
2. Trial Design.....	8
2.1 Cluster Randomization in Primary Care Setting	8
2.2 Inclusion and Exclusion Criteria	9
2.3 Randomization, Recruitment, and Consenting Procedures	10
2.4 Delivery of Intervention	10
2.5 Sample Size Consideration.....	11
3. Data Collection	12
3.1 Patient-Level Measures	12
3.2 Provider-Level Measures	14
3.3 Data Monitoring.....	15
3.4 Clinician Interview and Survey	15
4. Statistical Analysis	17
5. Discussion and Conclusion	18

Abstract

Background Worldwide, complications of decompensated cirrhosis and portal hypertension, including ascites, variceal hemorrhage, and hepatic encephalopathy, lead to approximately 2 million deaths per year and a rapidly rising number of hospitalizations. Cirrhosis insidiously develops over many years of liver injury, and it remains asymptomatic until late stages of disease. Approximately 2-5% of the general population have significant undetected liver fibrosis or established cirrhosis, with rates as high as 18-27% in populations with risk factors for liver disease. Thus, high index of suspicion is required for early diagnosis and to allow for effective therapeutic interventions. Currently available laboratory-based markers of advanced liver fibrosis have only modest performance in the general population with area under the receiver operating curve (AUC) between 0.54-0.71, and thus, novel screening tools with greater accuracy are urgently needed.

Cirrhosis is associated with cardiovascular changes and distinct abnormalities on electrocardiogram (ECG). The application of an inexpensive and widely available tool such as ECG for detection of unsuspected cirrhosis is of significant clinical value, particularly in low resource areas with high prevalence of chronic liver disease. Through an unsupervised machine learning approach, we have developed and internally validated a convoluted neural network using digitized ECG images of patients with cirrhosis. This original model, AI-Cirrhosis-ECG (ACE) score, demonstrated excellent performance with an AUC of 0.90 in identifying patients with advanced cirrhosis. These early findings led to the development of a new model, the Detection of Undiagnosed Liver Cirrhosis via Electrocardiogram (DULCE) score, trained and internally validated using 12-lead ECGs of patients with early asymptomatic cirrhosis stages. Thus, the DULCE score may potentially be used for screening of asymptomatic and unsuspected cirrhosis in primary care settings.

Objectives

The primary objective of this pragmatic trial is to validate a deep learning-based artificial intelligence (AI) model for early detection of cirrhosis-associated signals on digitized ECG and apply this newly developed neural network for primary care screening of chronic advanced liver disease. The secondary objective is to assess barriers for adoption of an AI-enabled algorithm for detection of liver disease in routine community clinical practice.

Design

A pragmatic, cluster randomized trial in 45 Mayo Clinic primary care practices will be conducted over a period of 6 months with 6 months of follow up. Care teams will be randomized 1:1 to intervention or usual care, stratified by region and patient volume. In the intervention arm, the DULCE score will be used to alert consenting providers to the likelihood of advanced liver disease with a recommendation for a FibroTest-ActiTest. The primary endpoint will be detection of advanced liver disease. Secondary outcomes will include completion of noninvasive fibrosis assessment tests and hepatology referral within 180 days of ECG, new diagnosis of liver disease stratified by etiology (nonalcoholic fatty liver disease, alcohol-associated liver disease, hepatitis C, and others) and severity (compensated with and without clinically-significant portal hypertension, and decompensated disease), initiation of prophylactic nonselective beta-blockers and imaging for hepatocellular carcinoma surveillance, according to published society guidelines. Post-study surveys to participating clinicians will be applied.

Summary

Preliminary findings suggest outstanding potential for the use of a novel and inexpensive ECG-enabled machine learning algorithm for detection of advanced liver disease in the primary care community.

1. Background and Objectives

1.1 Screening for cirrhosis

Early recognition of unsuspected cirrhosis remains an unmet clinical need despite significant advances in the field. Cirrhosis is among the leading causes of death in the United States and results from a wide range of infectious, metabolic, hereditary, toxic and autoimmune conditions.¹ Globally, complications of cirrhosis are responsible for 1.16 million deaths, and cirrhosis-associated liver cancer 788,000 deaths, making them the 11th and 16th most common causes of death, respectively, each year. Progression from asymptomatic or compensated liver disease to decompensated stages, defined by the development of variceal bleeding, ascites, jaundice or hepatic encephalopathy, occurs at a relatively slow rate of 4-10% per year depending on the etiology.¹ Thus, an opportunity for early detection of cirrhosis during which effective preventative and/or therapeutic interventions can be implemented exists. For instance, treatment of chronic liver diseases such as viral hepatitis C or B, alcohol use disorder, metabolic syndrome, primary biliary cholangitis and others, may successfully delay or prevent the progression of cirrhosis to decompensated stages or the development of hepatocellular carcinoma. In addition, nonselective beta-blockers have been shown to significantly decrease the incidence of gastroesophageal variceal bleeding² and development of ascites in patients with compensated cirrhosis³ and their use in this population is recommended by both national and international society guidelines.⁴

1.2 Current Screening and Diagnosis Methods

Currently, the diagnosis of asymptomatic compensated cirrhosis relies on a high index of suspicion and availability of abdominal imaging, noninvasive fibrosis assessment imaging tools (e.g. transient elastography or magnetic resonance elastography) or liver biopsy. Studies in the general population or in areas of high disease prevalence are being conducted with the use of

transient elastography⁵ or application of laboratory-based scoring systems (e.g. APRI, FIB-4).^{6,7} However, special equipment and trained personnel are required for the former, whereas the performance of laboratory tests alone remains suboptimal. Furthermore, transient elastography is not readily available in community settings outside European countries and thus cannot be currently used for screening of liver disease in United States. In addition, these noninvasive scoring systems have been developed from select populations with a high prevalence of advanced liver fibrosis. Such scores might not be suitable to use in populations in which the prevalence of significant liver disease is low, such as in a primary care setting. The application of an ECG-derived score for early detection of advanced liver disease is highly attractive given the low cost and wide availability of this test in community-based practices, as well as the potential for large scale population screening through increasingly popular wearable devices offering ECGs.

1.3 Development of an Artificial Intelligence (AI)-Enabled ECG-Based Screening Tool

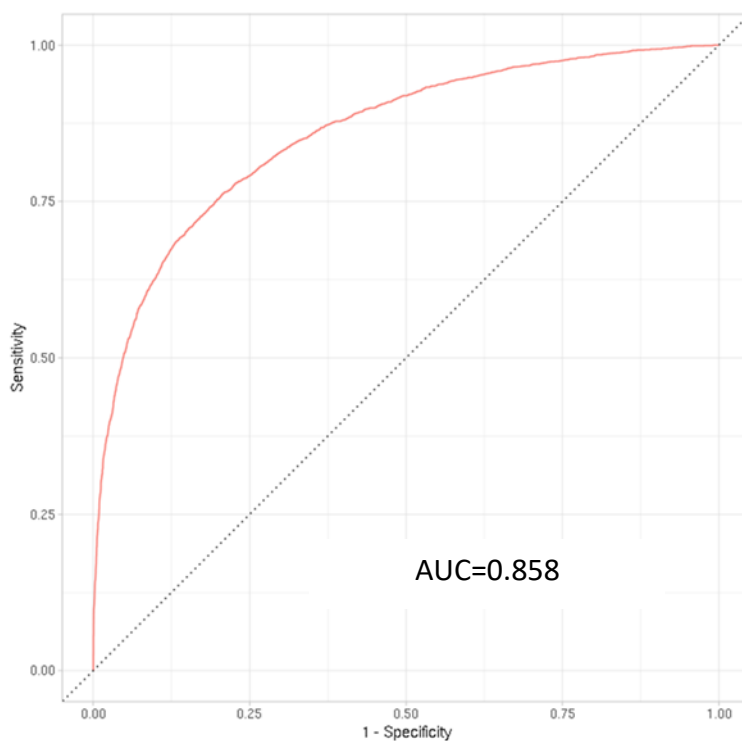
Development and application of ECG-enabled machine learning algorithms has been highly successful for screening of cardiac dysfunction in asymptomatic patients and detection of atrial fibrillation in otherwise normal ECGs.^{8,9} Cirrhosis results in significant hemodynamic changes that may impact cardiac function and electrical conductivity. Thus, it is plausible that additional unrecognized electrical cardiac signals of liver disease exist and could be leveraged for early detection of advanced liver disease. Through an unsupervised machine learning approach, our group developed and internally validated a convoluted neural network using digitized ECG images of 5,212 patients with advanced decompensated cirrhosis, paired to 20,728 controls without liver disease. This original model, AI-Cirrhosis-ECG (ACE) score, demonstrated excellent performance with an AUC of 0.90 in the testing subset.¹⁰ Based on these initial findings, our group has re-trained the deep learning model to include patients with earlier and asymptomatic stages of disease. The DULCE model retained excellent performance on internal

validation (AUC=0.86) with 79.4% sensitivity, 76.1% specificity, and 94.3% negative predictive value. (Figure 1) Additionally, sequential application of DULCE and FIB-4 models, revealed that DULCE thresholds can be optimized to maximize sensitivity (95%) and when combined with FIB-4 resulted in improved performance with 80-84% sensitivity and 84-89% specificity. This demonstrates that DULCE may be used more broadly as a screening tool, particularly when followed by more specific tests.

Figure 1. Model Performance

The receiver operating characteristic (ROC) curve for the model performance using 10 seconds of 12 leads.

Abbreviation: AUC = area under the curve



1.4 Study Objectives

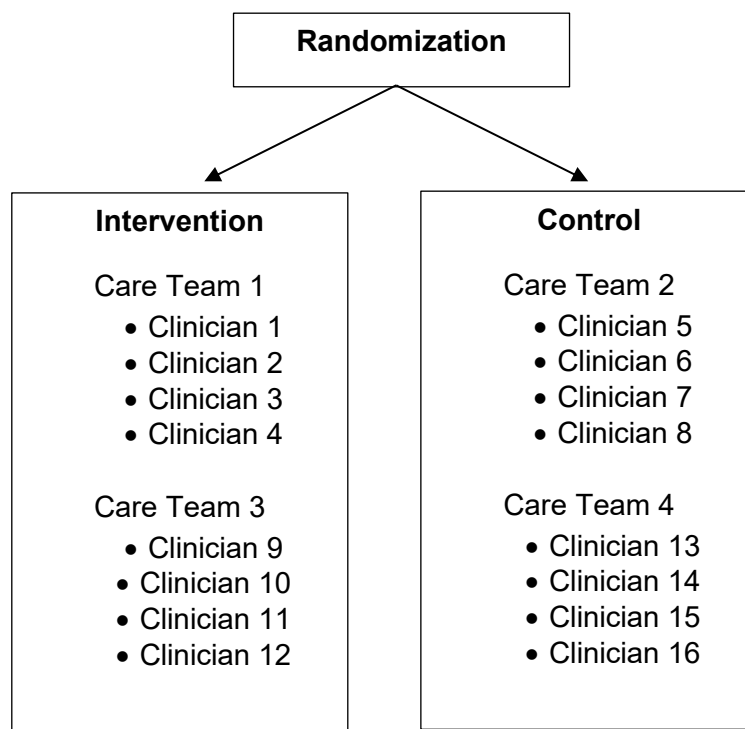
The objective of this study is to evaluate an ECG-based artificial intelligence-powered tool for detection of undiagnosed cirrhosis in primary care practices. Secondary objectives include assessing barriers for adoption of an AI-enabled algorithm for detection of liver disease in routine community clinical practice.

2. Trial Design

2.1 Cluster Randomization in Primary Care Setting

Randomization will be conducted at the care team level, with each care team serving as a cluster. Care teams will be randomized to intervention or usual care, stratified by region (Rochester, Southeast Minnesota, Southwest Minnesota, Northwest Wisconsin, and Southwest Wisconsin), and patient volume (i.e., the number of potential trial-eligible patients within the first 6 months of 2022).

Figure 2: Illustration of Cluster RCT



If clinicians agree to participate, they will have access to the screening report in Epic and will receive communications from the study team, e.g., a reminder when there is a positive screening result. If a clinician refuses to participate, he/she will not have access to the report and will not be contacted again, but his/her colleagues in the same care team can still participate.

2.2 Inclusion and Exclusion Criteria

The study participants will be primary care clinicians who are part of a care team that cares for adult patients (≥ 18 years) and who have the ability to order ECG, including physicians, nurse practitioners, and physician assistants. Among the internal medicine and family medicine care teams across Mayo Clinic's practice in Rochester and Mayo Clinic Health System, we excluded pediatrics, acute, nursing home, and resident care teams.

A patient's data will be included in the analysis if the patient is ≥ 18 years old and undergoing an ECG for any indication over a period of 6 months. Patients with known cirrhosis based on noninvasive fibrosis assessment tests, liver biopsy or complications of decompensated disease, or with a documented history of cirrhosis identified by clinical notes will be excluded. Patient data will only be collected if permitted (based on the use of the Minnesota Research Authorization Retrieval Tool).

The trial will run for approximately 6 months, during which a small number of patients may receive ECGs multiple times. For these patients, only data associated with the first encounter will be included in the analysis.

2.3 Randomization, Recruitment, and Consenting Procedures

Care teams will be randomized to intervention or usual care, stratified by site and patient volume. We will send clinicians at eligible care teams a recruitment email that includes a description of the intervention and the trial and will ask them to reply to indicate consent for participation. The email will explain that participation is voluntary and choosing to participate or not will not affect any aspect of their employment or work performance evaluation. We will send another three follow-up email reminders to clinicians who do not respond, and the study personnel will contact them via phone/pager or in person if necessary. If members of a care team decline to participate, the care team will still be eligible for randomization if their colleagues in the same team decide to participate. However, these clinicians who decline to participate will have not access to the report nor be contacted by the investigative team.

2.4 Delivery of Intervention

The intervention will be a report containing the screening result. Primary care clinicians in the intervention group will have access to the report, sent via an email alert, which will display whether the DULCE result is positive or negative. When the DULCE result is negative, the recommendation will be for *'no testing for advanced liver disease recommended unless indicated by other clinical or laboratory findings'*. When the ACE result is positive, the recommendation will be to 'consider ordering a FibroTest-ActiTest. The email alert will contain a link where clinicians can document a reason for choosing not to proceed with further testing if that is their decision. Clinicians in the usual care group will not receive any contact from the investigators. Clinical data will be abstracted from EMR for 6 months after ECG completion.

The intervention will be activated in the electronic health record system during the training period before the start of the trial. Clinicians in the intervention group will receive training materials via email. They will also receive 30-60 minutes of training via videoconferences, which will contain an introduction of the screening tool and the trial procedures and a session to answer clinicians' questions. Training videos will be available for clinicians who are unable to attend a group training session. All clinicians will be encouraged to contact the investigative team if they encounter any questions when using the tool on their own. This one-week pre-trial period will allow clinicians to get familiar with the tool before the start of the trial.

2.5 Sample Size Consideration

Based on a prior study conducted at the proposed sites, we estimate a total of 22,000 ECGs in unique patients will be completed during the study period (11,000 each arm). With an estimated prevalence of 2-5% in the general population, we should be adequately powered for detection of 50% difference in the diagnosis of cirrhosis in the intervention arm compared to usual care during the proposed period. A potential pitfall is that we may not observe a significant overall difference due to potential lower prevalence of undiagnosed cirrhosis in this community. However, due a significantly higher prevalence of advanced liver disease in individuals with underlying risk factors (obesity, metabolic syndrome, high risk alcohol use), we anticipate successfully achieving our secondary outcomes after stratifying by risk factors and etiology.

3. Data Collection

3.1 Patient-Level Measures

There will be no contact with patients. All the patient-level data will be collected from the EHR. Index date is defined as the date when the screening result is available for a patient. The time period prior to the index date is defined as this patient's baseline period, which will be used to collect medical history data. The following patient baseline characteristics will be abstracted (Table 1).

Table 1. Patient Baseline Characteristics

Patient Characteristics	Definition
1. Socio-demographic	
Age	
Sex	
Race	
Ethnicity	
Rural/urban status	
Marital Status	
Insurance	
Income	
Education	
Limited English Proficiency	
2. Medical History	
Hypertension	Diagnosis codes
Diabetes	Diagnosis codes
Obesity	BMI
MI	Diagnosis codes
Atrial fibrillation	Diagnosis codes
Other arrhythmias	Diagnosis codes
Obstructive Sleep Apnea	Diagnosis codes
Viral hepatitis	Diagnosis codes
Alcohol liver disease	Diagnosis codes
Nonalcoholic liver disease	Diagnosis codes
Other chronic liver disease	Diagnosis codes
Chronic kidney disease	Diagnosis codes
Cancer	Type and treatment if possible
Prior hepatology consult	
Smoking	Structured field in Epic
Alcohol use	Structured field in Epic
Previous liver ultrasound	Most recent US date, findings
Serum creatinine	Most recent measure

Serum sodium	Most recent measure
Albumin	Most recent measure
Prothrombin time/INR	Most recent measure
Total bilirubin	Most recent measure
AST	Most recent measure
ALT	Most recent measure
Platelet count	Most recent measure
3. Medication Use	
ACEi	Drug names
ARB	Drug names
Beta Blocker	Drug names and dose
Other hypertension medication	will need drug names, may group them in different ways
Antidiabetic medications	Classified into different classes
Statins	Classified into high, medium, low doses
Diuretics	Classified into different classes
Anti-craving medications	

Table 2. Patient Outcomes

Outcome	Definition	Analysis
Clinical Outcome		
Diagnosis of chronic advanced liver disease	Abnormal liver findings (coarse echotexture, nodular appearance), or splenomegaly on ultrasound	Stratified by risk factors or etiology of liver disease
Healthcare Process/Quality Measures		
Abdominal ultrasound	Within 180 days of the initial ECG	Stratified by screening result
Serum AST, ALT, Platelet count		Stratified by screening result
HCV Antibody		
Incidental liver mass		
Detection of hepatic steatosis		
Prescription of NSBB		
Prescription of antidiabetic medications		
Prescription of statins		
Primary care appointment		
Hepatology consult		
Hepatology econsult		
Upper endoscopy		

We will also pull the data for patients cared by the participating clinicians during the three months prior to the start of the trial. We will collect the outcome data, such as the use of abdominal ultrasound, to examine if the intervention group and the control group have similar practice patterns before the introduction of the intervention. If any difference is found, the baseline practice patterns will be adjusted in the regression.

3.2 Provider-Level Measures

Clinicians' information, e.g., age, sex, specialty (e.g., internal medicine and family medicine), type (e.g., physician and advanced practice provider), and years of practice, will be captured at the time of informed consent via a link to a Redcap survey. Two emails will be sent and the study personnel will follow up and send a paper survey via mail if no response is obtained from emails.

We will also count the number of clinicians in a care team and the number of patients cared by the team during the trial period.

3.3 Data Monitoring

Reports will be generated to track the proportion of patients with a positive screening result, the proportion of patients receiving the recommended tests, and the proportion of patients receiving guideline-recommended management. A report will be sent to each care team randomized to the intervention every month. Additionally, participating clinicians will receive an e-mail alert for each patient who has a positive screening result. The email will contain a link to a survey to document the reasons if clinicians choose not to order the recommended laboratory tests. Follow-up reminders will be sent via email, In Basket message or phone call if no tests are ordered and no reason is documented. The goal of the reminders is to better understand the barriers in the implementation. We may also follow up with clinicians to inquire about clinical cases in which the algorithm may have made a positive clinical impact.

3.4 Clinician Interview and Survey

During or right after the trial, the investigators will contact care teams with the lowest compliance with the recommendations and those with the highest compliance rates. This approach, called “positive deviance”, is commonly used in health services research to understand how some people, with similar resources and circumstances, outperform others.^{17, 18} The successful strategies can be then applied to help improve the performance of other teams.

Contact will be made with operations managers and/or clinician champions to schedule interviews via with either individual clinicians or a group of clinicians. Interviews may occur by phone, video conference, or in-person depending on what works best for participating clinicians/care teams. A topic guide was developed based on the normalization process theory

(NPT), a commonly used tool to understand the dynamics of implementing, embedding, and integrating new technology or complex interventions (Table 3).^{19, 20} We will not reveal to the teams whether they are good or bad outliers on performance. The blinding will enable participants to answer questions more forthrightly and in a non-defensive manner.²¹

Table 3. Mapping of Overarching Themes and Subthemes of NPT Framework

NPT construct	Theme	Subtheme
Coherence	Making sense of the new screening tool	Does the new tool provide valuable information that cannot be obtained elsewhere? Are you able to interpret the results of the tool? Do you trust the information? Is there a shared understanding of the value of this new tool in this team?
Cognitive participation	Participants' commitment and investment in using the intervention	Do you use the new tool to inform your decision making? Are you well prepared to discuss with the patients in the event of a positive screening result?
Collective action	Resources available to facilitate the intervention	Do leadership and management adequately support the new screening tool? Are the resources provided to you sufficient for using the new tool, e.g., training, education materials, and the contact information?
Reflexive monitoring	Appraising the impact of the new screening tool	Does the new tool improve the care provided to patients? What are the pros and cons of using this tool when fully integrated into your workflow? Does it affect your and your team's workload? Are there concerns related to costs, either from the department, payers, or patients? Would you like to continue using the tool when the trial ends? Who would you recommend use or not use this tool?

At the end of the trial, a short survey will be sent to all clinicians in the intervention group to thank them for their participation, inform clinicians of what to expect during the final data collection period, and assess their experience with the intervention. The survey questions are also based on NPT and are similar to the interview topic guide. Three emails with a link to the REDCap survey will be sent. Additionally, we will send a final e-mail to all clinicians in the control group to thank them for their participation.

4. Statistical Analysis

We will compare the intervention and control groups in terms of patient characteristics. We will also compare the practice patterns during the three months prior to the trial (e.g., use of abdominal ultrasound and liver tests). Two-sample t-test will be used for continuous variables, and Chi-square test will be used for binary and categorical variables. If any significant imbalance is detected for any baseline characteristics, a sensitivity analysis will be performed including the unbalanced characteristics in the regression models for outcomes. We will examine whether socioeconomic status played a role in the outcome of an AI-powered clinical decision support tool. The socioeconomic status will be measured by the area deprivation index, which captures area-level (e.g., census block group) social determinants of health.

For each outcome, we will use mixed-effect logistic regression to compare the intervention group to the control group with the care team as a random effect. Subgroup analyses will be performed stratified by age, sex, race, region (Rochester, Southwest Minnesota, Southeast Minnesota, Northwest Wisconsin, and Southwest Wisconsin), patient baseline comorbidities, and clinician/care team characteristics (e.g., mean age of clinicians, size of the team, proportion of physicians, and mean years of practice).

Exploratory analyses will be performed to assess relationships between each item in the clinician questionnaire and the compliance with the recommendations (i.e., the proportion of

patients with positive screening results receiving a FibroTest-ActiTest. For example, logistic regression will be performed to assess the relationship between the proportion of physicians answering “strongly disagree or disagree” to an item in a care team and the FibroTest-ActiTest. Such analyses will identify items that are most strongly related to compliance, and thus, help refine the tool and guide future implementation strategies.

For the qualitative feedback collected from the interview (which will be recorded and transcribed verbatim) and surveys, we will have two analysts independently analyze the data using the principle of constant comparison via a combination of deductive and inductive analyses. All the quotes will be first deductively categorized into different topic areas, and then further analyzed using a process of open coding (inductive analysis), followed by axial and selective coding (deductive analysis). After each phase, the findings will be compared between the two analysts and discussed to reach a consensus. If a consensus cannot be reached, the issue will be discussed with the PIs and the rest of the team.

5. Discussion and Conclusion

Many AI algorithms have been developed over the past decade. However, few studies have rigorously evaluated whether the algorithms translate to improved patient care in routine practice settings. This pragmatic RCT aims to answer this key question – does this new AI algorithm with good performance measured by c statistic, sensitivity, specificity, and accuracy, lead to an improvement in the diagnosis and management of advanced liver disease in everyday practice? The qualitative assessment will also explore how to best facilitate human-AI interaction and refine the implementation strategies, and thus, maximize the potential of AI in improving care delivery and patient outcomes.

The level of pragmatism is very high for this trial in that there are few inclusion/exclusion criteria for either clinicians or patients and the setting of the trial is almost the same as routine

clinical practice. The only difference is that email alerts, reminders, and monthly reports will be sent to clinicians, which may or may not continue during the enterprise-wide implementation.

In conclusion, we will conduct a pragmatic cluster RCT to test a newly developed AI-enabled ECG-based screening tool for improving the diagnosis and management of chronic advanced liver disease. The experience learned from this trial can pave the way for the evaluation and implementation of many AI algorithms that are currently under development.

REFERENCES

1. Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. *J Hepatol* 2019;70:151-71.
2. Sharma M, Singh S, Desai V, et al. Comparison of Therapies for Primary Prevention of Esophageal Variceal Bleeding: A Systematic Review and Network Meta-analysis. *Hepatology* 2019;69:1657-75.
3. Villanueva C, Albillos A, Genesca J, et al. beta blockers to prevent decompensation of cirrhosis in patients with clinically significant portal hypertension (PREDESCI): a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet* 2019;393:1597-608.
4. Garcia-Tsao G, Abraldes JG, Berzigotti A, Bosch J. Portal hypertensive bleeding in cirrhosis: Risk stratification, diagnosis, and management: 2016 practice guidance by the American Association for the study of liver diseases. *Hepatology* 2017;65:310-35.
5. Roulot D, Costes JL, Buyck JF, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. *Gut* 2011;60:977-84.
6. Hagstrom H, Talback M, Andreasson A, Walldius G, Hammar N. Ability of Noninvasive Scoring Systems to Identify Individuals in the Population at Risk for Severe Liver Disease. *Gastroenterology* 2020;158:200-14.
7. Hagstrom H, Talback M, Andreasson A, Walldius G, Hammar N. Repeated FIB-4 measurements can help identify individuals at risk of severe liver disease. *J Hepatol* 2020;73:1023-9.
8. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med* 2019;25:70-4.
9. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet* 2019;394:861-7.
10. Ahn JC, Attia ZI, Rattan P, et al. Development of the AI-Cirrhosis-ECG Score: An Electrocardiogram-Based Deep Learning Model in Cirrhosis. *Am J Gastroenterol* 2022;117:424-32.