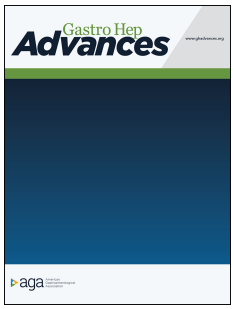
Journal Pre-proof



At-Home Critical Flicker Frequency Testing in Cirrhosis: A Feasibility Pilot Study Using the Novel Beacon Device

Philip Vutien, MD, MS, Richard Li, PhD, Ravi Karkar, PhD, James Fogarty, PhD, Sean A. Munson, PhD, Sabrina Omer, MS, Michael Yacoub, MS, James Kashima, BS, George N. Ioannou, BMBCh, MS

PII: S2772-5723(26)00228-1

DOI: https://doi.org/10.1016/j.gastha.2026.101107

Reference: GASTHA 101107

To appear in: *Gastro Hep Advances*

Received Date: 30 November 2025

Revised Date: 19 August 2026

Accepted Date: 24 August 2026

Please cite this article as: Vutien P, Li R, Karkar R, Fogarty J, Munson SA, Omer S, Yacoub M, Kashima J, Ioannou GN, At-Home Critical Flicker Frequency Testing in Cirrhosis: A Feasibility Pilot Study Using the Novel Beacon Device, *Gastro Hep Advances* (2026), doi: https://doi.org/10.1016/j.gastha.2026.101107.

Title: At-Home Critical Flicker Frequency Testing in Cirrhosis: A Feasibility Pilot Study Using the Novel Beacon Device

Philip Vutien MD, MS[1,2*], Richard Li PhD[3*], Ravi Karkar, PhD [4,3], James Fogarty PhD[3], Sean A. Munson PhD [5], Sabrina Omer MS[1], Michael Yacoub MS[1], James Kashima BS[1], George N. Ioannou, BMBCh, MS[1,2]
*These authors contributed equally to this manuscript.

Authors' Affiliations:
[1]University of Washington, Division of Gastroenterology and Hepatology, Seattle, WA, USA
[2]Department of Medicine Veterans Affairs Puget Sound Healthcare System and University of Washington, Seattle, WA, USA
[3]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA
[4]College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA, USA
[5]Human Centered Design and Engineering, University of Washington, Seattle, WA, USA

Guarantors of the article:
George Ioannou, BMBCh, MS, FAASLD
Veterans Affairs Puget Sound Health Care System
Gastroenterology
S-111-Gastro
1660 S. Columbian Way
Seattle, WA 98108
Tel 206-277-3136
Fax 206-764-2232
georgei@medicine.washington.edu

Philip Vutien, MD, MS
University of Washington
Division of Gastroenterology and Hepatology
1959 NE Pacific Street, Box 356175
Seattle, WA 98195-6175
pvutien@uw.edu

Declaration of funding interests:
This study was funded in part by NIH/NIDDK award number R21DK117431 and by NIH/NLM award number R01LM012810, and by the University of Washington Population Health Initiative. The investigators own the data, performed all statistical analysis and prepared the entire manuscript independently.

Author e-mails
1) Philip Vutien: pvutien@uw.edu
2) Richard Li: lichard@cs.washington.edu
3) Ravi Karkar: rkarkar@umass.edu
4) Sean A. Munson: smunson@uw.edu
5) James Fogarty: jfogarty@cs.washington.edu
6) Michael Yacoub: myacoub@pnwu.edu

7) Sabrina Omer: somer11@uw.edu
8) James Kashima: kashij@uw.edu
9) George N. Ioannou: George.ioannou@va.gov

Authorship Contributions:

1) Philip Vutien: Study concept and design, analysis and interpretation of data, drafting and critical review of the manuscript
2) Richard Li: Study concept and design, analysis and interpretation of data, drafting and critical review of the manuscript
3) Ravi Karkar: Study concept and design, analysis and interpretation of data, critical review of the manuscript
4) Sean A. Munson: Study concept and design, analysis and interpretation of data, critical review of the manuscript
5) James Fogarty: Study concept and design, analysis and interpretation of data, critical review of the manuscript
6) Kara Walter: Study concept and design, acquisition of data, critical review of the manuscript
7) Sabrina Omer: acquisition of data, critical review of the manuscript
8) Michael Yacoub: acquisition of data, critical review of the manuscript
9) James Kashima: acquisition of data, critical review of the manuscript
10) Debby Tsuang: acquisition of data, critical review of the manuscript
11) George N. Ioannou: Study concept and design, analysis and interpretation of data, drafting and critical review of the manuscript

Keywords: Flicker fusion, covert encephalopathy, Flicker-App, Beacon, home testing

Abbreviations:
CFF – Critical Flicker Frequency
HE – Hepatic Encephalopathy
MOL-D – Method of Limits–Descending
MASLD – Metabolic Dysfunction Associated Steatotic Liver Disease
MELD – Model for End-Stage Liver Disease
SUS – System Usability Scale
TLX – NASA Task Load Index
UBS – User Burden Scale
MHE – Minimal Hepatic Encephalopathy
OHE – Overt Hepatic Encephalopathy
CTP – Child–Turcotte–Pugh
TIPS – Transjugular Intrahepatic Portosystemic Shunt
PBC – Primary Biliary Cholangitis
PSC – Primary Sclerosing Cholangitis
INR – International Normalized Ratio
FFS – Flicker Fusion System
API – Application Programming Interface

Conflicts of interest: All authors disclose no conflicts of interests in relation to this manuscript.
Data request: Data and analytic methods will be made available to other researchers upon reasonable request.

Reporting guidelines/checklist: This prospective feasibility study was reported in accordance with the STROBE guidelines.
Ethical considerations: This study was reviewed and approved by the University of Washington Institutional Review Board (IRB) under study number STUDY00004521.
Short title: At-Home CFF Testing for HE in Cirrhosis

Title count: 69 characters
Manuscript word count (excluding references): 3,341
Abstract word count: 275
Tables: 2**, Figures:** 5

Abstract

Background

Critical flicker frequency (CFF) is a well-validated neurophysiologic screening test for minimal hepatic encephalopathy (MHE) in cirrhosis. We evaluated the feasibility and reliability of repeated at-home, self-administered CFF testing using a novel device (Beacon).

Methods

We enrolled 21 participants with cirrhosis. Each received a Beacon device and smartphone preloaded with an application to measure CFF using two protocols: 1) method of limits–descending (MOL-D) and 2) Forced Choice. Participants were instructed to complete ≥5 tests per week for 2 weeks, followed by ≥1 test per week for an additional 4 weeks. Outcomes included adherence, test duration, and longitudinal CFF trends to assess potential learning effects from repeated testing.

Results

Participants were predominantly male (76%) with a mean age of 52.2 years; 43% had alcohol-related liver disease and 38% had prior overt HE. Adherence (completing ≥70% of recommended measurements) was observed in 19 (90.5%) for daily and 16 participants (76.2%) for weekly testing. Median CFF was 42.9 Hz (IQR: 38.4-44.4) by MOL-D and 43.7 Hz (IQR: 41.7-45.8) by Forced Choice. Forced Choice CFF was significantly lower in those with prior HE (median 40.7 Hz vs. 44.9 Hz; $p = 0.02$), while MOL-D CFF did not differ by HE status (42.7 Hz vs. 42.9 Hz; $p = 0.44$). CFF remained stable over time for both protocols (MOL-D: +0.01 Hz/day of testing, $p=0.67$; Forced Choice: +0.03 Hz/day, $p=0.09$). Forced Choice testing required more time than MOL-D to complete (median 4.67 vs. 2.48 minutes; $p < 0.001$).

Conclusion

Beacon enabled reliable, self-administered CFF testing at home with high adherence and no evidence of learning effects. Further studies incorporating other validated MHE tools are needed to determine whether at-home CFF monitoring can detect MHE early and prevent overt HE episodes and hospitalizations.

Introduction

Cirrhosis commonly causes a spectrum of neurocognitive impairments known as hepatic encephalopathy (HE) [1]. HE fluctuates greatly over time and can range from minimal HE (MHE) to overt HE, which can be life-threatening. The most widely studied neurophysiologic screening test for MHE is critical flicker frequency (CFF) [2-9]. Recognizing MHE is important because it is linked to driving accidents, falls, future overt HE, and increased mortality.[10-14] Early recognition allows treatment with medications such as lactulose or rifaximin to prevent complications.[15, 16]

Current commercially available devices that measure CFF are large, expensive, and non-portable [17]. They are not designed for self-administration but instead require a technician to administer the test. As a result, they are rarely used in routine clinical practice or for at-home monitoring and instead have been used primarily in research settings. To address this limitation, we developed Beacon, a device that includes a stand-alone, portable, inexpensive flickering light source that communicates via Bluetooth to a smartphone application that administers protocols to measure CFF [18].

To date, no CFF test, including Beacon, has been evaluated for at-home use in patients with cirrhosis. This is important since ultimately any screening test for MHE needs to be able to be executed reliably by patients at home in an unsupervised manner. We aimed to evaluate the feasibility and usability of at-home CFF testing using the Beacon device in patients with cirrhosis. Specifically, we sought to: 1) determine whether patients could reliably self-administer CFF measurements and assess adherence to daily or weekly testing protocols; 2) evaluate for potential learning effects over time, which commonly occur with many psychometric tests resulting in improved scores without any actual improvement in the underlying condition; and 3) compare performance and completion time between two CFF testing protocols—Method of Limits-Descending (MOL-D) and Forced Choice.

Methods

Development of Beacon and Flicker-App system for measuring CFF

Beacon **(Figure 1a)** consists of a light source that communicates via Bluetooth with a smartphone application (**Figure 1b**) [18]. Beacon generated a milky white, flickering light at frequencies controlled by

the smartphone application, which executes a protocol to show different frequencies to determine an observer's CFF threshold (i.e., the highest frequency at which an observer can distinguish that a light source is flickering). The formative design of Beacon has been described in detail in prior publications [18-20].

Study participants and baseline characteristics

We recruited participants with cirrhosis (n=21) from the hepatology clinics of the University of Washington between 7/2022-4/2024. Cirrhosis was diagnosed by clinical, laboratory, and radiological features or by liver biopsy. We intentionally included patients with a history of prior overt HE, as this population is at highest risk for recurrent HE episodes and represents the group in whom at-home CFF monitoring would be most clinically relevant; we sought to determine whether self-administered testing was feasible even in this more impaired population. We excluded study participants with a history of dementia, severe vision impairment, epilepsy, or benzodiazepine, alcohol, or antiepileptic drug use within the week prior to recruitment. We collected demographic, clinical, and laboratory data from patient electronic health records. History of overt HE was defined by the prescription of medications (rifaximin or lactulose) specifically documented in the medical notes to address prior HE. This study was reviewed and approved by the University of Washington Institutional Review Board (IRB) under study number STUDY00004521. All procedures performed in this study were in accordance with the ethical standards of the institutional review board and with the Helsinki Declaration. Informed consent was obtained from all individual participants included in the study.

Study procedures

Each participant was provided with a Beacon device and a study smartphone preloaded with the Flicker-App to enable at-home CFF testing. Prior to beginning home testing, participants received in-person or video-based training on device setup, test initiation, and interpretation of the interface. Printed quick-start instructions were included with each kit. Participants were instructed to place Beacon in an at-home location with minimal visual and auditory distractions; however, they were not explicitly advised on a specific environmental setup (e.g., ambient lighting levels or room configuration). Devices were mailed directly to participants' homes, and follow-up support was available as needed. Study participants executed two CFF protocols:

1) Method-of-limits-descending (MOL-D) protocol.[21] In this protocol, the light stimulus begins at 55 Hz—a frequency above the threshold of human flicker perception—and gradually decreases at a rate of 0.5

Hz per second **(Figure 2a)**. Participants were instructed to tap the smartphone screen as soon as they perceive flickering. This process was repeated eight times. The CFF threshold is calculated as the average of the eight responses. The app has a lower bound of 25 Hz and automatically stops at this frequency unless the participant chooses to retake the measurement. Retaking a measurement was always presented as an option.

2) Forced Choice protocol.[21] In this protocol **(Figure 2b)**, participants were sequentially shown two flickering lights via the smartphone app. One light flickers at 120 Hz (should be perceived as fused by all observers), and the other—the "target"—starts at 25 Hz. Participants indicated which light appears to flicker. Based on their responses, the target frequency is adjusted: it increases by 2 Hz after three consecutive correct "target" identifications or decreases by 2 Hz after an incorrect response. This continues until eight reversals (i.e., changes in direction of frequency adjustment) are recorded. The CFF threshold is defined as the mean of the final six reversal values.

At-home flicker CFF measurements with Beacon using both daily and weekly testing

Participants were asked to complete both the MOL-D and Forced Choice protocols each time at home:

1) During weeks 1–2: at least five measurement sessions, once daily per week
2) During weeks 3–6: at least one measurement session per week

Patients were monetarily compensated up to 150 US dollars (30 dollars for week 1, 40 dollars for week 2, and then 20 dollars for weeks 3-6) for participating in this at-home study.

The Flicker-app automatically uploaded test results to a secure server in real time, enabling remote study monitoring.

Qualitative survey instruments – System Usability Scale and exit survey

At the end of the six-week period, participants completed an exit interview that included the System Usability Scale (SUS) questionnaire **(Supplementary Figure 1)** to assess the usability of the Beacon system.[22] Questions on the SUS instrument are presented with alternating positive and negative tone, so each individual response must be adjusted and summed and scaled to obtain a final score. SUS scores above 84.1 correspond to a letter grade of A+, indicating high usability. During this exit-interview, patients were asked about their motivation for using Beacon, their understanding of CFF measurements, and challenges they encountered when using this device. Further details and discussion of the survey findings are available in a prior publication.[19]

Statistical Analysis

Descriptive statistics were reported using medians and interquartile ranges (IQRs) for continuous variables, and differences between groups were assessed using the Wilcoxon rank-sum test. Differences in test completion time and critical flicker frequency (CFF) between protocols (MOL-D and Forced Choice) were evaluated per participant using within-subject medians and paired differences. To assess changes in critical flicker frequency (CFF) values and test duration over time — i.e., to evaluate the presence of a learning effect — we fit linear mixed-effects models and separately for CFF values derived from MOL-D and Forced Choice protocols. Each model included a fixed effect for measurement day (representing the chronological day of at-home self-administered testing) and random intercepts and slopes for participants to account for repeated measures within individuals. All statistical analyses were conducted in R version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria), using RStudio version 1.4.1717. This study was approved by the Institutional Review Board at the University of Washington (Study ID 4521), and all research was conducted in accordance with both the Declarations of Helsinki and Istanbul. Written consent was given in writing by all study participants.

Results

Baseline characteristics of at-home study participants

We recruited 21 study participants with cirrhosis with mean age 52.2 ± 12.3 years **(Table 1)**, of whom 16 (76%) were male, 7 (38%) had a prior history of overt hepatic encephalopathy and were on lactulose and/or rifaximin prior to testing. The etiology of liver disease was hepatitis C in 4 (19%), alcohol-related in 9 (42.9%), and metabolic dysfunction-associated steatotic liver disease (MASLD) in 4 participants (19%). The mean MELD score was 14.2 ± 4.4.

During the at-home testing period for each participant, there were no documented medication initiations of lactulose, rifaximin, or L-Ornithine L-Aspartate or changes in these medications. No participants developed overt HE or were hospitalized during the at-home testing period.

CFF measurement values by protocol and by prior HE status

The median CFF values per-participant over the at-home study period ranged from 35.7 to 50.6 Hz (overall median: 42.9 Hz) using the MOL-D protocol, and from 37.5 to 53.3 Hz (overall median: 43.7 Hz) using the Forced Choice protocol **(Table 2, Figure 3)**. When comparing protocols within individuals, Forced Choice CFF values were a median of 2.4 Hz higher than those obtained with the MOL-

D protocol (IQR: 0.9 to 4.2 Hz, *p*-value= 0.032). CFF values derived from the MOL-D protocol were similar between participants with and without a history of HE (median 42.7 Hz vs. 42.9 Hz, respectively; p = 0.44). In contrast, CFF values obtained using the Forced Choice protocol were lower in participants with prior HE compared to those without (median 40.7 Hz vs. 44.9 Hz; p = 0.02) **(Figure 4)**.

Impact of repeated testing on CFF values according to protocol

In the MOL-D protocol, CFF values remained stable over time, with no significant change observed across repeated measurements within individuals (average increase of 0.01 Hz per day of testing; p = 0.67) (**Table 2, Supplementary Figure 2**) and with a median increase of 1 Hz (IQR: -0.46 to 3.09) from the first (median 41.5 Hz) to last measurement (median 42.4 Hz) per participant. There was also not a significant change in CFF values over time for the Forced Choice protocol (average increase of 0.03 Hz per day of testing; p = 0.09) and with a median increase of 1.33 Hz (IQR: -1.34 to 3.33) from the first (median 42.7 Hz) to last (median 43 Hz) measurement per participant.

Length of time to complete MOL-D and Forced Choice testing

The median time to complete the test was 2.48 minutes (IQR: 2.28-3.27) for the MOL-D protocol and 4.67 minutes (IQR: 3.87-6.52) for the Forced Choice protocol. On average, the Forced Choice protocol took 1.85 minutes longer to complete than the MOL-D protocol ($p < 0.001$) **(Figure 5)**. When stratified by HE history, there were no significant differences in test duration for either protocol. Among participants without HE, the median duration for MOL-D was 2.65 minutes (IQR: 1.06), compared to 2.35 minutes (IQR: 0.86) in those with HE ($p = 0.74$). For the Forced Choice protocol, the median duration was 4.98 minutes (IQR: 2.75) in those without HE and 4.19 minutes (IQR: 1.03) in those with HE ($p = 0.46$).

Decreasing test duration across repeat sessions

For the MOL-D protocol, completion time declined by an average of 1.39 seconds per day of testing (p < 0.001). Overall, participants showed a median reduction of 1.25 minutes (IQR: 3.02 to 0.8) in test duration from their first (median 4.22 minutes) to last session (median 2.35 minutes). In the Forced Choice protocol, the reduction was more pronounced, with an average decrease of 4.6 seconds per day (p < 0.001) and a median decrease of 3.47 minutes (IQR: 6.44 to 2.55) between the first (median 10.4 minutes) and last session (median 3.97 minutes) (**Supplementary Figure 3**).

Participant adherence to testing schedule

According to the study protocol, participants were expected to complete at least 10 measurements during weeks 1–2 (i.e. at least 5/week) and 4 additional measurements during weeks 3–6 (i.e. at least 1/week) **(Table 2)**. Among the 21 participants, 19 (90.5%) completed at least 7 measurements during the initial two-week period, and 12 (57%) met the full target of 10 daily measurements. During weeks 3–6, 16 participants (76.2%) completed at least 3 weekly measurements, while 10 (48%) were fully adherent and provided 4 weekly measurements.

Participant qualitative reports of Beacon ease of use and perceived value of testing

For patients who completed the exit interview (n = 15), the mean score for the System Usability Scale was 88.5 (standard deviation 1) for the MOL-D and 86 (standard deviation 0.71) for the Forced Choice protocol, indicating excellent usability for both protocols. Most participants saw the Beacon system as a meaningful tool for personal health monitoring, with 14 of 15 interviewed patients reporting that they were motivated to participate in the study to advance science and see how this technology can help patients with hepatic encephalopathy. For example, **participant #8** noted, "It gives them an idea of how careful they need to be. Unfortunately, we did not hear about HE when we first had it. My husband thought I had a stroke. People who are going in the beginning of their journey will find this extremely helpful." **Participant #16** stated "It's nice that the measures are better than I thought I was, but I know I'm still below. I know I have liver issues. It is what it is. There's nothing I wanna do or should do. I wanna know what's going on with my body." Similarly, **Participant #15** reflected, "It helped me realize that you know there's more the liver has so many functions in the body, not so much struggling with but the liver affects this and that so many things, nice to learn what all it's affecting." These responses show that participants found Beacon usable and CFF results empowering in helping them understand their liver disease.

Discussion

Beacon is currently the only CFF test[3, 6, 9, 10] designed for self-administration at home, and in this pilot feasibility study, patients with cirrhosis reliably used it to obtain daily and weekly CFF measurements. Unlike psychometric tests for HE, Beacon demonstrated minimal to no learning effects, supporting its reliability for repeated longitudinal monitoring. Thus, Beacon can complement existing tools including the Psychometric Hepatic Encephalopathy Score (PHES) and EncephalApp[1, 15, 23-26], and additional studies

utilizing these tests are needed to determine whether longitudinal at-home CFF monitoring enables early detection of MHE and prevention of overt HE episodes and hospitalizations[15, 16, 27].

There is a critical need for at-home screening tests for HE, since patients can deteriorate and develop HE at home, which is often unrecognized or diagnosed late. The most widely used tool is the EncephalApp which administers the Stroop test, a psychometric test of cognitive flexibility and response inhibition [15, 23-25]. Stroop test scores can predict future adverse outcomes including episodes of overt HE [28]. However, as a *psychometric test*, test performance is potentially confounded by effort, literacy, education, and social determinants of health. There may also be a learning effect, whereby repeated testing leads to improved scores independent of actual changes in mental status or clinical acuity. Beacon, as a *neurophysiologic test* that measures CFF, is unlikely to be affected by these factors. As such, CFF testing with Beacon may serve as an alternative or complementary screening test to EncephalApp and can improve HE screening rates, which are currently low[29, 30]. The estimated cost of manufacturing Beacon at scale would be low (estimated at under 50 US dollars per unit), and we hope to pursue opportunities to make Beacon available to patients, clinical providers, and investigators for future research.

Prior studies evaluating CFF in patients with cirrhosis have used protocols based on the method-of-limits, which is prone to a response bias that arises from test habituation and familiarity[21, 31]. This bias may be particularly relevant with repeated measurements over time, such as with at-home testing. To address this issue, we implemented and piloted an alternative protocol using a 'Forced Choice' method, which is not susceptible to a response bias[21, 31]. In this study, we found that CFF values obtained using the Forced Choice protocol were reliable and there was no evidence for a learning response over time. However, as expected, the Forced Choice protocol also took significantly more time to complete (approximately 4 minutes for Forced Choice vs. 2.4 minutes for MOL-D). The longer duration required for the Forced Choice protocol may be inconvenient for 'routine' at-home use. To reduce the overall time needed for the Forced Choice protocol, potential approaches could include (a) decreasing the number of reversals required; (b) determining whether a person can detect flicker at a target frequency instead of determining their actual CFF threshold; or (c) using historical information about a person's CFF to run a narrower sweep (e.g., starting the target light at 35 Hz instead of 25 Hz). Additional studies are needed to optimize the Forced Choice paradigm for longitudinal at-home CFF monitoring.

We note potential study limitations. First, only 21 participants completed the at-home portion of this study and none of the participants were hospitalized or developed overt HE during the study period. Furthermore, validated psychometric or neurophysiologic assessments of hepatic encephalopathy, such as PHES, were not collected in this pilot study, and therefore we are unable to assess diagnostic accuracy or clinical correlation. Additional studies with more participants, a longer follow-up time, and concurrent assessment of PHES and other validated HE tests are planned to evaluate if decreases in at-home CFF measurements are associated with the development of overt HE. Although we intentionally enrolled patients with prior OHE to test feasibility in a higher-risk population, we did not have data on the timing of each participant's most recent OHE episode, which limits our ability to characterize the relationship between OHE recency and CFF values. Additionally, home testing environments were not standardized and variability in ambient lighting and testing conditions may have introduced measurement noise. Despite this, per-participant CFF values remained stable over time across repeated measurements. In addition, participants were aware that the investigators had developed and manufactured the Beacon. This awareness may have biased participant responses on the usability of Beacon to be more favorable. Finally, the study cohort partially overlaps with a prior publication focused on deployment logistics and usability design of the Beacon system for a human-computer interaction audience.[19] The current manuscript is distinct in its focus on quantitative CFF measurement data including protocol comparison, assessment of learning effects over time, and adherence to daily and weekly testing schedules.

Conclusions:

In summary, Beacon enables fast, reliable, at-home measurement of CFF in patients with cirrhosis, with high adherence and no evidence of learning effects. Further studies with other validated MHE tools are needed to assess whether at-home CFF monitoring can detect MHE early and prevent overt HE episodes and HE-related hospitalizations.

FIGURE LEGENDS

Figure 1: Devices used in this study to measure critical flicker frequency.

1a) Beacon device and smartphone application. The Beacon system consists of two components: (A) an Application running on a smartphone for user input and to record results and (B) (overall dimensions: 8.9 cm x 30 cm), the light stimulus source with wireless controller and battery base.

1b) Beacon smartphone application (Flicker-App). An observer uses the Flicker-App to pair with the Beacon device over Bluetooth.

Figure 2a-2b: Schematic diagram for the CFF protocols run in this study.

2a) Method-of-limits, descending (MOL-D) protocol. Participants are shown a light source that flickers at a high frequency (55 Hz) at which all observers should perceive as a fused light. The frequency decreases over time at a rate of 0.5 Hz/sec. The observer ends a "run" when they perceive the light to be flickering. Eight runs are repeated, and the mean of the eight runs are calculated as the CFF.

2b) Forced Choice protocol. Participants were sequentially shown two flickering lights via the smartphone app. One light flickers at 120 Hz (should be perceived as fused by all observers), and the other—the "target"—starts at 25 Hz. Participants indicate which light appears to flicker. Based on their responses, the target frequency is adjusted: it increases by 2 Hz after three consecutive correct "target" identifications or decreases by 2 Hz after an incorrect response. This continued until eight reversals (i.e., changes in direction of frequency adjustment) are recorded. The CFF threshold is defined as the mean of the final six reversal values.

Figure 3. At-home Beacon CFF test results obtained over time and by protocol. According to the study protocol, participants were expected to complete at least 10 measurements during weeks 1–2 and 4 additional measurements during weeks 3–6. Among the 21 participants, 19 (90.5%) completed at least 7 measurements during the initial two-week period, and 12 (57%) met the full target of 10. During weeks 3–6, 16 participants (76%) completed at least 3 measurements, while 10 (48%) were fully adherent with the 4-session target.

Figure 4. At-home Beacon CFF results for those without a history of overt hepatic encephalopathy (n = 14) vs those with prior hepatic encephalopathy (n = 7). CFF values derived from the MOL-D protocol were similar between participants with and without a history of HE (median 42.7 Hz vs. 42.9 Hz, respectively; $p = 0.44$). In contrast, CFF values obtained using the Forced Choice protocol were lower in participants with prior HE compared to those without (median 40.7 Hz vs. 44.9 Hz; $p = 0.02$).

Figure 5. Length of time to complete CFF testing by protocol. Across all participants, the median time to complete the test was 2.48 minutes (IQR: 2.28-3.27) for the MOL-D protocol and 4.67 minutes (IQR: 3.87-6.52) for the Forced Choice protocol. On average, the Forced Choice protocol took 1.85 minutes longer to complete than the MOL-D protocol ($p < 0.001$).

References

1. Vilstrup H, Amodio P, Bajaj J, et al. Hepatic encephalopathy in chronic liver disease: 2014 Practice Guideline by the American Association for the Study of Liver Diseases and the European Association for the Study of the Liver. Hepatology 2014;60:715-35.
2. Ampuero J, Montoliu C, Simon-Talero M, et al. Minimal hepatic encephalopathy identifies patients at risk of faster cirrhosis progression. J Gastroenterol Hepatol 2018;33:718-725.
3. Barone M, Shahini E, Iannone A, et al. Critical flicker frequency test predicts overt hepatic encephalopathy and survival in patients with liver cirrhosis. Dig Liver Dis 2018;50:496-500.
4. Dhiman RK, Kurmi R, Thumburu KK, et al. Diagnosis and prognostic significance of minimal hepatic encephalopathy in patients with cirrhosis of liver. Dig Dis Sci 2010;55:2381-90.
5. Greinert R, Ripoll C, Hollenbach M, et al. Stepwise diagnosis in covert hepatic encephalopathy: critical flicker frequency and MELD-score as a first-step approach. Aliment Pharmacol Ther 2016;44:514-21.
6. Kircheis G, Bode JG, Hilger N, et al. Diagnostic and prognostic values of critical flicker frequency determination as new diagnostic tool for objective HE evaluation in patients undergoing TIPS implantation. Eur J Gastroenterol Hepatol 2009;21:1383-94.
7. Kircheis G, Hilger N, Haussinger D. Value of critical flicker frequency and psychometric hepatic encephalopathy score in diagnosis of low-grade hepatic encephalopathy. Gastroenterology 2014;146:961-9.
8. Romero-Gomez M, Cordoba J, Jover R, et al. Value of the critical flicker frequency in patients with minimal hepatic encephalopathy. Hepatology 2007;45:879-85.
9. Torlot FJ, McPhail MJ, Taylor-Robinson SD. Meta-analysis: The diagnostic accuracy of critical flicker frequency in minimal hepatic encephalopathy. Aliment Pharmacol Ther 2013;37:527-36.
10. Ampuero J, Simon M, Montoliu C, et al. Minimal hepatic encephalopathy and critical flicker frequency are associated with survival of patients with cirrhosis. Gastroenterology 2015;149:1483-9.
11. Bajaj JS, Saeian K, Schubert CM, et al. Minimal hepatic encephalopathy is associated with motor vehicle crashes: the reality beyond the driving test. Hepatology 2009;50:1175-83.
12. Roman E, Cordoba J, Torrens M, et al. Minimal hepatic encephalopathy is associated with falls. Am J Gastroenterol 2011;106:476-82.
13. Bajaj JS, Lai JC, Tandon P, et al. Role of Oral Health, Frailty, and Minimal Hepatic Encephalopathy in the Risk of Hospitalization: A Prospective Multi-Center Cohort of Outpatients With Cirrhosis. Clin Gastroenterol Hepatol 2023;21:1864-1872 e2.
14. Gairing SJ, Mangini C, Zarantonello L, et al. Minimal hepatic encephalopathy is associated with a higher risk of overt hepatic encephalopathy and poorer survival. J Intern Med 2024;295:331-345.
15. Bajaj JS, Heuman DM, Wade JB, et al. Rifaximin improves driving simulator performance in a randomized trial of patients with minimal hepatic encephalopathy. Gastroenterology 2011;140:478-487 e1.

16. Prasad S, Dhiman RK, Duseja A, et al. Lactulose improves cognitive functions and health-related quality of life in patients with cirrhosis who have minimal hepatic encephalopathy. Hepatology 2007;45:549-59.
17. Instruments L. Flicker Fusion System, 2021.
18. Karkar R, Kocielnik R, Zhang X, et al. Beacon: Designing a Portable Device for Self-Administering a Measure of Critical Flicker Frequency. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2018;2:1-27.
19. Li R, Vutien P, Omer S, et al. Deploying and Examining Beacon for At-Home Patient Self-Monitoring with Critical Flicker Frequency. Proc SIGCHI Conf Hum Factor Comput Syst 2025;2025.
20. Vutien P, Li R, Karkar R, et al. Evaluating a Novel, Portable, Self-Administrable Device ("Beacon") That Measures Critical Flicker Frequency as a Test for Hepatic Encephalopathy. Am J Gastroenterol 2023;118:1096-1100.
21. Eisen-Enosh A, Farah N, Burgansky-Eliash Z, et al. Evaluation of Critical Flicker-Fusion Frequency Measurement Methods for the Investigation of Visual Temporal Resolution. Sci Rep 2017;7:15621.
22. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale Benchmarking for Digital Health Apps: Meta-analysis. JMIR Mhealth Uhealth 2022;10:e37290.
23. Allampati S, Duarte-Rojo A, Thacker LR, et al. Diagnosis of Minimal Hepatic Encephalopathy Using Stroop EncephalApp: A Multicenter US-Based, Norm-Based Study. Am J Gastroenterol 2016;111:78-86.
24. Acharya C, Wade JB, Fagan A, et al. Overt hepatic encephalopathy impairs learning on the EncephalApp stroop which is reversible after liver transplantation. Liver Transpl 2017;23:1396-1403.
25. Labenz C, Gairing SJ, Kaps L, et al. QuickStroop for screening for minimal hepatic encephalopathy in patients with cirrhosis. JHEP Rep 2025;7:101298.
26. Weissenborn K, Ennen JC, Schomerus H, et al. Neuropsychological characterization of hepatic encephalopathy. J Hepatol 2001;34:768-73.
27. Sidhu SS, Goyal O, Mishra BP, et al. Rifaximin improves psychometric performance and health-related quality of life in patients with minimal hepatic encephalopathy (the RIME Trial). Am J Gastroenterol 2011;106:307-16.
28. Duarte-Rojo A, Allampati S, Thacker LR, et al. Diagnosis of covert hepatic encephalopathy: a multi-center study testing the utility of single versus combined testing. Metab Brain Dis 2019;34:289-295.
29. Bajaj JS. Diagnosing minimal hepatic encephalopathy: from the ivory tower to the real world. Gastroenterology 2015;149:1330-3.
30. Bajaj JS, Etemadian A, Hafeezullah M, et al. Testing for minimal hepatic encephalopathy in the United States: An AASLD survey. Hepatology 2007;45:833-4.
31. Clark WC, Brown JC, Rutschmann J. Flicker sensitivity and response bias in psychiatric patients and normal subjects. Journal of Abnormal Psychology 1967;72:35-42.

Table 1. Baseline characteristics of the study participants who underwent at-home testing

Baseline characteristics	***All Participants (N=21)***
Age, years (mean ± SD)	*52.2 ± 12.3*
Male	*16 (76%)*
Race/ethnicity	
White, non-Hispanic	*16 (76%)*
Hispanic or Latino	*1 (4.8%)*
Native American	*2 (9.5%)*
Pacific Islander	*1 (4.8%)*
Other/Prefer not to answer	*1 (4.8%)*
Underlying liver disease	
Alcohol-related liver disease	*9 (43%)*
Chronic hepatitis C infection	*4 (19%)*
Metabolic dysfunction-associated steatohepatitis	*4 (19%)*
Autoimmune, PBC, PSC	*2 (9.5%)*
Other liver disease	*2 (9.5%)*
Presence of gastro-esophageal varices	*10 (47.6%)*
Prior history of hepatic encephalopathy	*7 (38%)*
Baseline lab studies	
Albumin	*3.8 ± 0.5*
Platelet count	*121.9 ± 62.6*
Total bilirubin	*1.7 ± 1.6*
International Normalized Ratio (INR)	*1.3 ± 0.3*
Serum creatinine	*1.2 ± 1.0*
MELD score (mean ± SD)	*14.2 ± 4.4*
CTP Class	
A	*11 (52.4%)*
B	*9 (42.9%)*
C	*1 (4.8%)*

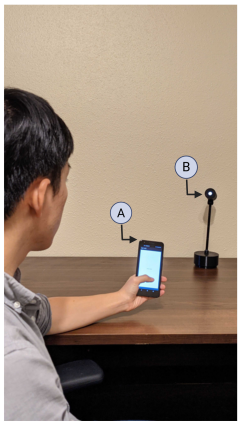
MELD: Model for end stage liver disease, CTP: Child-Turcotte-Pugh, TIPS: transjugular intrahepatic portosystemic shunt, PBC: Primary biliary cholangitis, PSC: Primary sclerosing cholangitis

Table 2. Summary of Beacon testing performance and adherence

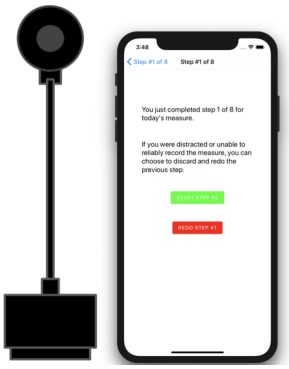
	Descending Method	Forced Choice Method
A. Critical Flicker Frequency (CFF)		
Median CFF (Hz, [IQR])	42.9 (38.4-44.4)	43.7 (41.7–45.8)
Median CFF at first test (Hz, [IQR])	41.5 (37.1-44.8)	42.7 (39.7-44)
Median CFF at last test (Hz, [IQR])	42.4 (39-44.5)	43 (41-47.7)
Average change in CFF per day of testing (Hz, [95% CI])	0.01 (-0.03 to 0.05)	0.03 (-0.004 to 0.06)
B. Time to Completion		
Median test duration in minutes (median, [IQR])	2.48 (2.28-3.27)	4.67 (3.87-6.52)

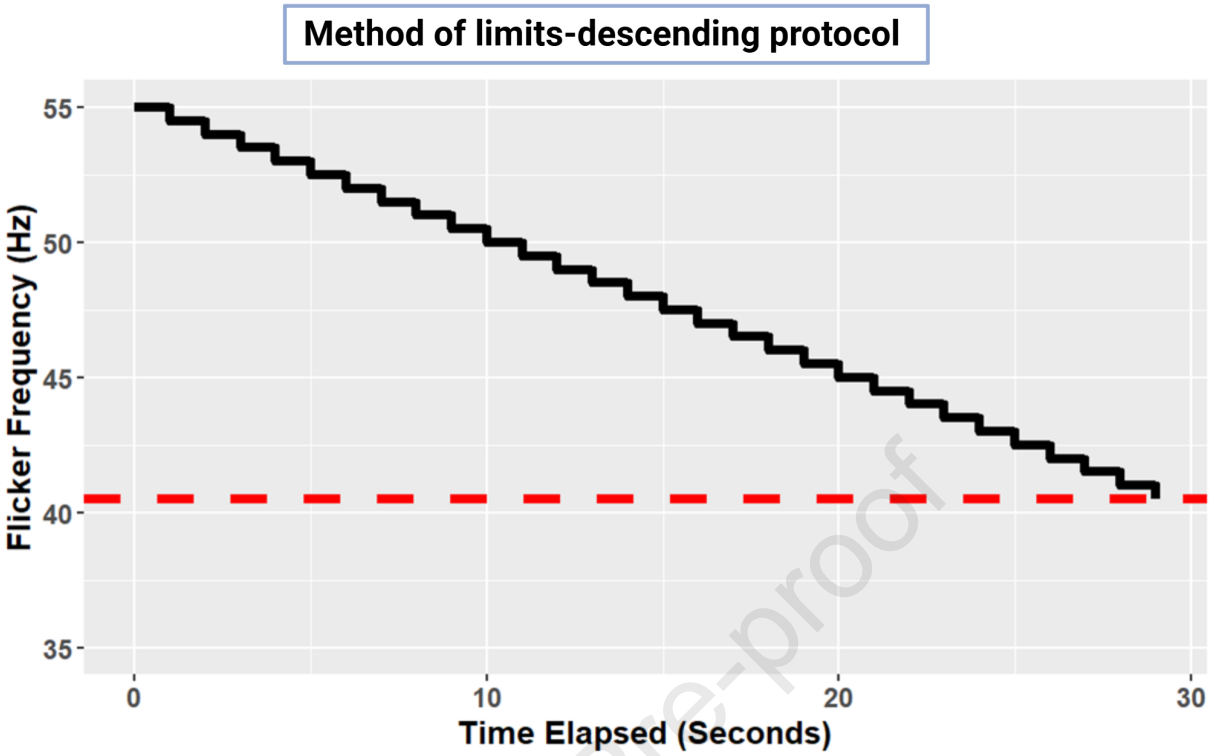
First test duration in minutes (median, [IQR])	4.22 (3.20-5.24)	10.40 (6.26-12.2)
Last test duration in minutes (median, [IQR])	2.35 min (2.13-3.17)	3.97 (3.15-5.70)
Difference in timing to complete first and last test in minutes (median, [IQR])	-1.25 (-3.02 to -0.8)	–3.47 (-6.44 to -2.55)
C. Adherence to daily and weekly testing protocol		
Median (IQR) number of daily tests completed (goal: 10 tests)	10 (8-10)	
Proportion completing at least 7 daily testing	19 (90.5%)	
Median (IQR) number of tests completed during weekly testing (goal: 4 tests)	3 (IQR: 3-4)	
Proportion completing at least 3 weekly testing	16 (76.2%)	

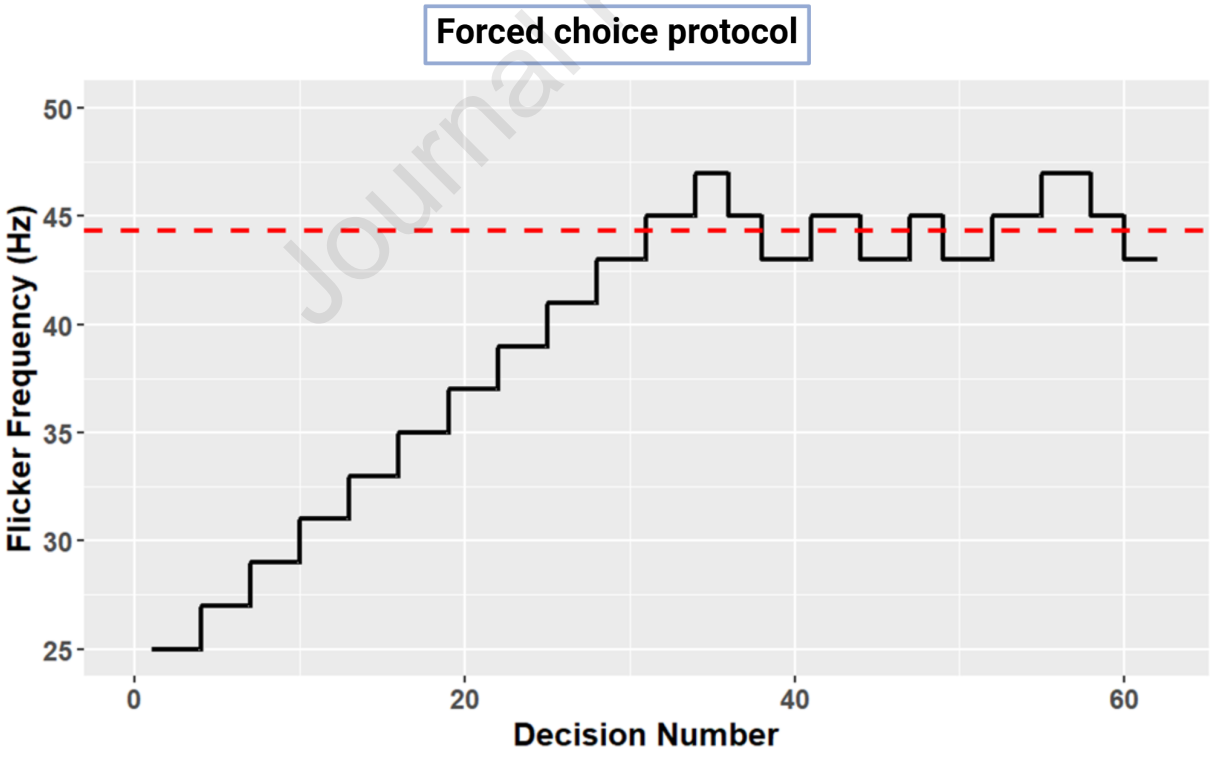
(A) Smartphone and
(B) Beacon Device

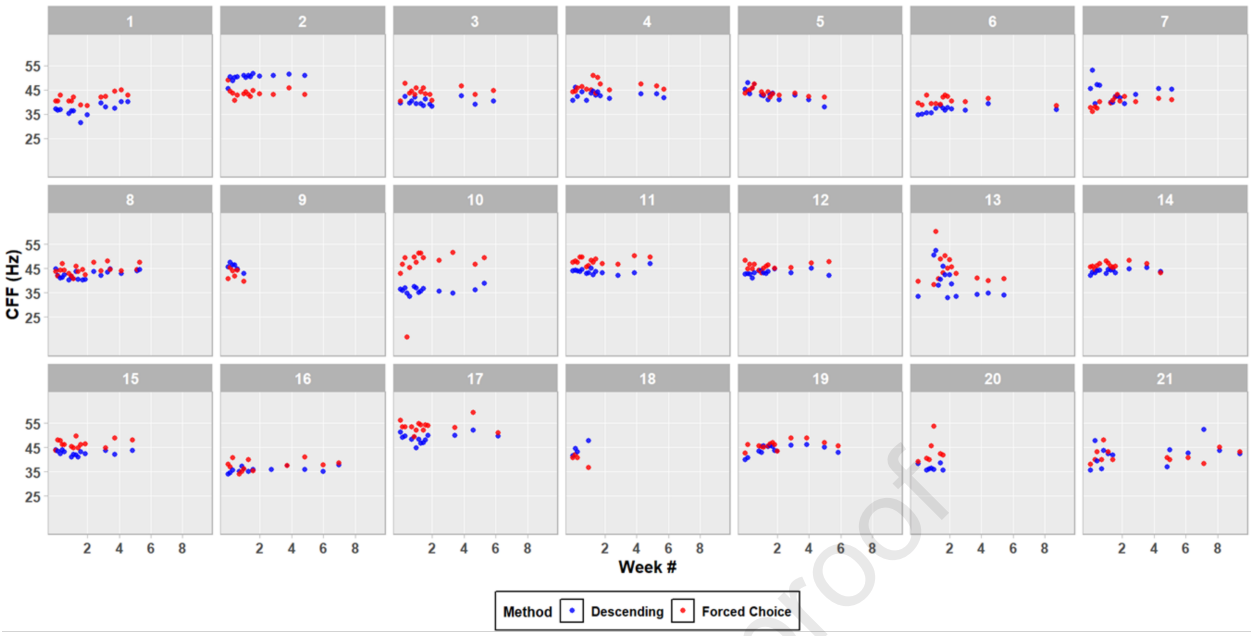


Smartphone application (Flicker-App)

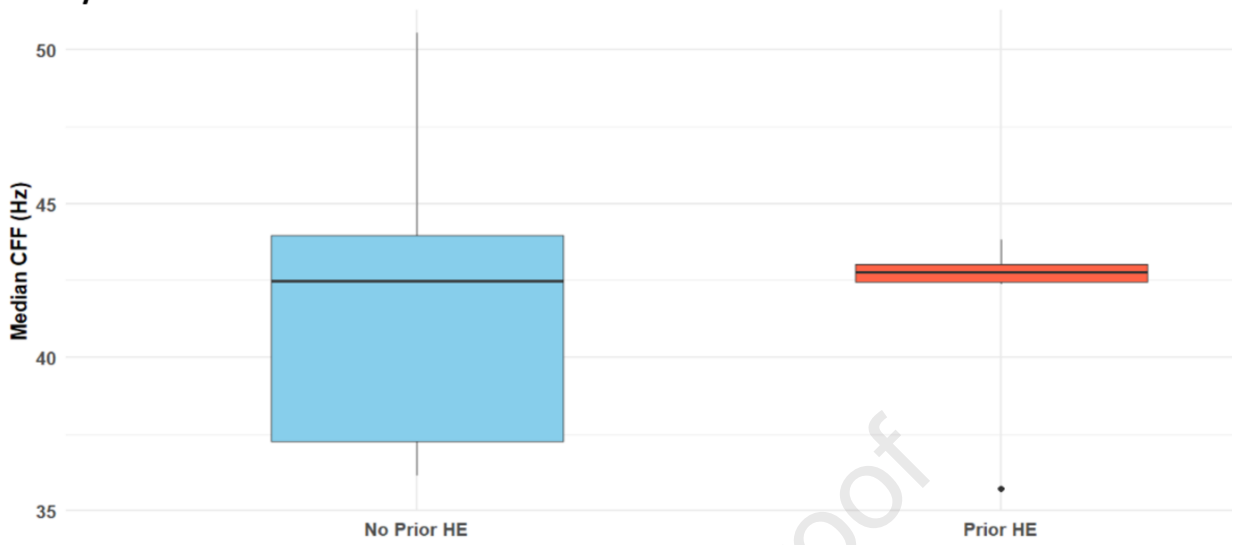








CFF by MOL-D



CFF by Forced Choice