The Stroop Smartphone App is A Short and Valid Method to Screen for Minimal Hepatic Encephalopathy

Jasmohan S Bajaj, MD[1], **Leroy R Thacker, PhD**[2], **Douglas M Heuman, MD**[1], **Michael Fuchs, MD**[1], **Richard K Sterling, MD**[1], **Arun J Sanyal, MD**[1], **Puneet Puri, MD**[1], **Mohammad S Siddiqui, MD**[1], **Richard T Stravitz, MD**[1], **Iliana Bouneva, MD**[1], **Velimir Luketic, MD**[1], **Nicole Noble, BS**[1], **Melanie B White, RN**[1], **Pamela Monteith, RN**[1], **Ariel Unser, BS**[1], and **James B Wade, PhD**[3]

[1]Division of Gastroenterology, Hepatology and Nutrition, Virginia Commonwealth University and McGuire VA Medical Center, Richmond, Virginia

[2]Department of Biostatistics, Virginia Commonwealth University and McGuire VA Medical Center, Richmond, Virginia

[3]Department of Psychiatry, Virginia Commonwealth University and McGuire VA Medical Center, Richmond, Virginia

Abstract

Minimal hepatic encephalopathy(MHE) detection is difficult due to the unavailability of short screening tools. Therefore MHE patients can remain undiagnosed and untreated.

Aim—To use a Stroop smartphone application(EncephalApp_Stroop) to screen for MHE.

Methods—The App and standard psychometric tests(SPT; 2/4 abnormal is MHE, gold standard), psychometric hepatic encephalopathy score(PHES) and Inhibitory control(ICT) tests were administered to cirrhotics [with/without prior-overt HE(OHE)] and age-matched controls from two centers; a subset underwent re-testing. A separate validation cohort was also recruited. Stroop has "off"state with neutral stimuli and "on"state with incongruent stimuli. Outcomes: time to complete five correct runs, number of trials needed in On(Ontime) and Off(Offtime) states. Stroop results were compared between controls and cirrhotics with/without OHE, and those with/without MHE(using SPT,ICT,PHES). ROC analysis was performed to diagnose MHE in cirrhotics with/without prior OHE.

Results—125 cirrhotics (43 prior OHE) and 134 controls were included in the original cohort. App times were correlated with MELD (Offtime:r=0.57,Ontime:r=0.61,p<0.0001) and were worst in prior-OHE patients compared to the rest and controls. Stroop performance was also significantly impaired in those with MHE compared to no-MHE according to SPT,ICT and PHES (all p<0.0001). A cut-off of>274.9 seconds(OnTime+OffTime) had an AUC=0.89 in all patients and 0.84 in patients without prior OHE for MHE diagnosis using SPT as the gold standard. The validation cohort showed 78% sensitivity and 90% specificity with the >274.9 seconds OnTime +OffTime cut-off. App result patterns were similar between the centers. Test-retest reliability in controls and those without prior OHE was good; a learning effect on Ontime in cirrhotics without prior OHE was seen. In <u>conclusion</u>, the Stroop smartphone app is a short, valid and reliable tool for screening of MHE.

<u>Corresponding Author:</u> Jasmohan S Bajaj, MD, MS Division of Gastroenterology, Hepatology and Nutrition, Virginia Commonwealth University and McGuire VA Medical Center, 1201 Broad Rock Boulevard, Richmond VA 23249 Phone: (804) 675-5802, Fax: (804) 675-5816 jsbajaj@vcu.edu.

Keywords

Spectrum of neurocognitive impairment in cirrhosis; complications of cirrhosis; validity; cirrhosis; covert hepatic encephalopathy

The spectrum of neuro-cognitive impairment in cirrhosis (SONIC), that ranges from unimpaired, minimal (MHE) to overt hepatic encephalopathy (OHE), can adversely affect the daily life of affected patients and caregivers(1, 2). MHE is associated with impaired quality of life (QOL), employment, driving capability and a higher risk of progression to overt HE (OHE)(3–6). MHE treatment can improve QOL, driving capability and progression to OHE(7–10). However, this is not standard of care partly because routine MHE testing is often not feasible in the United States(11). In order to increase testing and subsequently treatment rates in MHE, a high-sensitivity, easily applicable screening test, which can be administered in the clinic within a few minutes, is required(12). More definitive cognitive and electro-physiological testing could then only be reserved for patients who perform poorly on this task(12). The Stroop task is a test of psychomotor speed and cognitive flexibility that evaluates the functioning of the anterior attention system and has been found to be sensitive for the detection of cognitive impairment in MHE(13–15). There is also a smartphone application for the Stroop task (EncephalApp_Stroop) which is an attractive option for a point-of-care testing strategy for the diagnosis of MHE and cognitive dysfunction in cirrhosis (www.encephalapp.com).

We aimed to validate the use of this Stroop smartphone app for the screening of cognitive dysfunction in cirrhosis.

Methods

Healthy controls and patients with cirrhosis were recruited prospectively from two independent hepatology centers: VCU Medical Center and McGuire VA Medical Center, after obtaining written informed consent. Cirrhosis was diagnosed by compatible laboratory features of thrombocytopenia and AST/ALT reversal with radiological findings of cirrhosis or endoscopic evidence of varices in the setting of chronic liver disease. We excluded patients who were not able to consent, had uncontrolled neuro-psychiatric diagnoses, were on psychoactive medications apart from stable antidepressants, were abusing alcohol or illicit drugs in the past 3 months, had red-green color blindness and had uncontrolled OHE (defined as mini-mental status exam<25) at the time of the examination.

Subjects then underwent a battery of recommended cognitive tests(1, 12, 16): (a) Psychometric hepatic encephalopathy score (PHES), (b) block design test (BDT: subjects are required to replicate standardized designs with given blocks in a timed manner. The score is based on the designs correctly copied) and (c) Inhibitory control test [This is a 15 minute computerized test. Subjects are instructed to respond to alternating presentations of X and Y on the screen (targets) while inhibiting response when X and Y are not alternating (lures) or responding to letters other than X or Y (random)](17). The PHES consists of 5 tests: number connection test-A/B (NCT-A/B: subjects are asked to "join the dots" between numbers or numbers and letter in a timed fashion and the number of seconds required is the outcome), digit symbol (DST: subjects are required to pair numbers with special symbols. An individual's score reflects the number of correct pairs achieved within a 120 second time frame), line tracing [LTTtime: subjects are required to trace a line between two parallel lines and the time required is noted and LTTerrors: number of times the subject strays outside the lines] and serial dotting (SDT, subjects are asked to dot the center of a group of blank circles and the time required is the outcome)]. The PHES is a validated battery for cognitive

dysfunction in cirrhosis and tests for psychomotor speed, visuo-motor coordination, attention and set-shifting. The BDT requires an individual to use colored blocks to reproduce a 2 dimensional visual design which tests visual-motor coordination and nonverbal problem-solving. The ICT is a validated computerized test of attention, psychomotor speed, response inhibition and working memory. Weighted lures (WL) are lures/square of target accuracy and are a composite outcome of ICT in impaired patients(18). A high score on BDT, DST and ICT targets and a low score on the rest of the tests indicates good cognitive performance.

Age-balanced healthy controls without chronic medical illness or alcohol and illicit drug abuse were recruited from the community through word-of-mouth referrals and advertisements. A portion of these controls had been recruited for ongoing cognitive dysfunction studies within the last two years and only had performance on NCT-A, NCT-B, DST, BDT and ICT while the remainder had all tests performed (PHES, ICT and Stroop).

We used three complementary modalities to diagnose MHE in our population. As recommended by Ferenci *et al*, we used impaired performance in any two of NCT-A/B, DST or BDT of two standard deviations beyond healthy controls to be the gold standard(1). We also used two other modalities, the ICT and PHES to compared the App. We used the healthy control performances to score the PHES from +1 to −3 standard deviations (SD) and any score impaired beyond -4SD was considered MHE by PHES while we used WL to define poor ICT performance(19). Stroop (EncephalApp_Stroop) app: the application was downloaded from the Apple app store (EncephalApp Stroop) and used on the apple iPod platforms. The iPod screens were used to administer the task to all subjects. The task has two components: "Off" and "On" state depending on the discordance or concordance of the stimuli. Both components were administered after 2 training runs were given for each state. In the easier "Off" state, the subject views a neutral stimulus, pound signs (###) presented in red, green or blue, one at a time and has to respond as quickly as possible by touching the matching color of the stimulus to the colors displayed at the bottom of the screen. The colors at the bottom of the screen are also randomized and not fixed to their respective positions. This continues until a total of 10 presentations which is one run and the total time taken for the run as well as the individual responses. If the subject makes a mistake, i.e. presses a wrong color, the run stops and has to restart again. Therefore the number of runs required to make 5 correct runs also indicates the number of mistakes. We continued the off state till the subject had achieved 5 correct runs. The "On" state is more challenging from a cognitive standpoint in that incongruent stimuli are presented in nine of the ten stimuli. In this portion, the subject has to accurately touch the color of the word presented which is actually the name of the color in discordant coloring, i.e., the word "RED" is displayed in blue color and the correct response is blue, not red. Similar to the "off" state, we gave two training runs and then continued the task till 5 correct runs were achieved.

The specific outcomes at the end of the Stroop app were: a. total time for 5 correct runs in the "off" state (OffTime), b. number of runs needed to complete the 5 correct "Off" runs, c. total time for 5 correct runs in the "on" state (OnTime) and d. number of runs needed to complete the 5 correct "On" runs. The test of cognitive processing controlling for psychomotor speed was subtracting the Offtime from the Ontime and this was performed for all groups. All administrations and psychometric tests were supervised by the psychologist (JBW). We compared the App results between the two independent study sites (VCU and VAMC). One administrator (AU) started working on this trial recently and her findings were compared to the previously obtained data.

Prospective validation cohort

A further group of cirrhotic patients were prospectively recruited to validate the findings on the ROC generated with the first cohort.

Longitudinal study

A subgroup of cirrhotic patients (with/without prior OHE) and controls, whose clinical status remained unchanged, underwent Stroop testing twice within 6 months of the prior testing.

Statistical analysis

We compared demographics, MELD score, venous ammonia, serum sodium across study groups. Cirrhotic patients were studied as a whole while the subgroup without prior OHE was studied separately as well. All cognitive tests including Stroop outcomes were compared between controls and the cirrhotic groups. App results were correlated with age, education status, ammonia, sodium, MELD score and other cognitive tasks and were compared in patients with and without an alcoholic and hepatitis C etiology of cirrhosis. An ROC analysis was performed with all Stroop results comparing it to MHE diagnosed using Standard tests as well as by PHES and weighted lures in all cirrhotics and in patients without prior OHE. The inflection point of the variable with the highest sensitivity/specificity was chosen as the cut-off for determination of cognitive dysfunction using the Stroop app. This inflection point was then used to evaluate the results of the prospective cohort. Baseline variables, age, education and MELD score along with App results were entered into a logistic regression for MHE diagnosis for all three modalities. Kappa statistics were used to study agreement between the three modes of MHE diagnosis and with the App results. We compared the App and cognitive results between the two study centers and used the sites in a regression analysis for prediction of MHE and cognitive dysfunction. The longitudinal results were studied using paired t-tests of Offtime and Ontime between the first and second time separately in the control, cirrhotic with prior OHE and those without prior OHE.

The Institutional Review Boards at Virginia Commonwealth University and McGuire VA Medical Center approved the protocols.

Results

We recruited 126 patients with cirrhosis (43 with prior OHE) and 51 age-balanced healthy controls for the initial cohort and 43 additional cirrhotic patients (12 with prior OHE) for the prospective validation cohort. One male cirrhotic patient without prior OHE could not perform Stroop due to color-blindness and therefore was not considered further. Of the 43 patients with prior OHE, the median time before the last episode was 2.5 months and the median prior OHE episodes was two. The majority (n=29) were controlled on lactulose alone while the remainder was on additional rifaximin for OHE therapy. The demographics and cognitive performance of the controls compared to cirrhotics with and without prior overt HE as shown in table 1.

Control group

While 51 age-matched controls were recruited for this study specifically, the gold standard definition of MHE using NCT-A/B, DST and BDT were obtained by pooling these 51 controls with an 83 additional controls who had been previously tested with these four tests and ICT but had not received the remaining PHES tests or Stroop. The mean age of the pooled 134 controls was 52±4 years and the years of education was 13±5 years (74 High School/15 some college-education/45 College degree or higher). The mean±SD for the 134 controls was NCT-A: 25±5, NCT-B: 90±15, DST: 83±16, BDT: 44±17, Weighted lures

8±7. Therefore the mean for standard MHE diagnosis was two of the following: NCT-A>35 seconds, NCT-B>120 seconds, DST<51 or BDT<10, while that using weighted lures was 22 or higher.

Based on these cut-offs, we found that in the whole cohort (n=125), the prevalence of cognitive dysfunction by all three modalities was significantly higher in those with prior OHE (table 1). The prevalence of MHE in cirrhotics without prior OHE was 33% with standard tests, 22% using weighted lures and 52% using PHES (tables 2A, B and C). The kappa of MHE diagnosis between the three modalities was highest between Standard and PHES (0.63), then between Standard and ICT (0.41) and lowest between PHES and ICT (0.2). Despite these different prevalence rates and lack of agreement according the methods, the App results were consistently worse in those with poor performance on any of the three modalities.

Of the subjects enrolled 24% of the patients and 27% of the controls had operated and were familiar with an iPod or smartphone. The Stroop Ontime was significantly higher than the Offtime in both healthy controls (72±10 vs 59±8 seconds, p<0.0001) and cirrhotic patients (100±32 vs 80±21 seconds, p<0.0001) while the number of runs needed to complete 5 runs were similar between off and on states for controls and cirrhotic patients. Patients with prior OHE had a significantly higher Stroop Ontime and Offtime and number of trials needed to complete 5 correct runs in the off and on state (table1, Figure 1a), and similar findings were seen in the MHE group compared to the no MHE and control groups (table 2). In patients with alcoholic etiology of cirrhosis (including those with concomitant hepatitis C) there was a significantly higher Offtime (159±43 vs. 128±32 seconds, p=0.004), Ontime (196±65 vs. 161±50 seconds, p=0.028), number of trials for 5 off runs (median 8 vs. 5, p=0.002) but similar number of trials for 5 on runs and On minus offtime (37±38 vs. 33±27 seconds, p=0.66), compared to the remainder of the cirrhotic patients. Patients with alcoholic etiology had similar MELD score (13±6 vs 12±6, p=0.33) to non-alcoholic patients. When patients with only hepatitis C (not with concomitant alcohol) were compared to the rest of the cirrhotics, no significant difference in Offtime (133±33 vs. 133±37 seconds, p=0.99), Ontime (166±55 vs. 1.63±52 seconds, p=0.74), On minus Offtime (35±30 vs. 32±27 seconds, p=0.49) number of trials to achieve 5 off (median 6 vs. 5, p=0.65) and 5 on runs (6 vs. 6, p=0.85) was observed; their MELD scores were similar (12±6 vs 12±5, p=0.67).

There was no significant difference between those who had prior exposure to smartphones compared to the rest on Offtime (141±37 vs. 131±60 seconds, p=0.75), Ontime (173±52 vs. 161±61 seconds, p=0.67), On minus Offtime (33±31 vs. 33±28 seconds, p=0.54) number of trials to achieve 5 off (median 5 vs. 5, p=0.78) and 5 on runs (6 vs. 6, p=0.76).

In the cirrhosis group as a whole, a significant correlation was seen between MELD score and OffTime (0.57, p<0.001) and OnTime (0.61, p<0.0001). There was a modest correlation between age and OffTime (0.38, p<0.001), and OnTime (0.31, p=0.001) but not between these variables and education, ammonia and serum sodium. No consistent correlations between number of on/off trials was found with the above variables. In patients, Stroop Offtime was significantly correlated with Ontime (0.87, p<0.0001), Number connection-A (0.62, p<0.0001), number connection-B (0.65, p<0.0001), digit symbol (−0.79, p<0.0001), block design (−0.54, p<0.0001), serial dotting (0.57, p<0.0001), line tracing (0.48, p<0.0001), ICT Lures (0.4, p<0.0001) and targets (−0.62, p<0.0001). Similarly Ontime was significantly correlated with number connection-A (0.57, p<0.0001), number connection-B (0.67, p<0.0001), digit symbol (−0.71, p<0.0001), block design (−0.51, p<0.0001), serial dotting (0.52, p<0.0001), line tracing (0.4, p<0.0001), ICT Lures (0.4, p<0.0001) and targets (−0.65, p<0.0001). An overall weaker correlation between Ontime minus Offtime and cognitive tests was seen; number connection-A (0.30, p=0.001), number connection-B (0.43,

p<0.0001), digit symbol (−0.36, p<0.0001), block design (−0.29, p=0.001), serial dotting (0.26, p=0.004), line tracing (0.07, p=0.43), ICT Lures (0.21, p=0.02) and targets (−0.50, p<0.0001). When Kappa was performed between Stroop App OnTime+OffTime and the three modalities, the agreement between the standard tests and Stroop (0.7) and PHES and Stroop (0.7) were similar while it was comparatively lower for ICT weighted lures (0.53).

Comparison between centers

The two independent centers, VCU and VAMC were compared with respect to their results using a combination of the original and validation cohort (n=168; 62 VCU and 106 VAMC). We found that the spread of cognitive and Stroop app results were aligned with the total findings; i.e. the worse scores were at the VAMC since the patients there were more advanced in their liver disease and were older compared to the VCU patients (Supplementary table 1). Stroop results were significantly correlated with MELD and age in both centers independently. MELD score was positively correlated with OffTime(r=0.4, p=0.005 VCU, r=0.5, p<0.0001 VAMC) and OnTime (r=0.5, p<0.0001 VCU, r=0.5, p<0.0001 VAMC). Age had similar correlations between centers with OffTime(r=0.4, p=0.006 VCU, r=0.3, p=0.02 VAMC), OnTime (r=0.3, p=0.03 VCU, r=0.25, p=0.04). On regression analysis site was not an independent predictor when added to age and MELD score for OnTime, OffTime and Ontime+Offtime.

ROC analysis

Using the four standard tests, the App variables (Offtime, Ontime, number of trials for 5 on runs, number of trials for 5 off runs, product of Offtime and number of runs, product of Ontime and number of runs and sum of Offtime and Ontime (Tables 3a and b). The highest AUC and sensitivity was equivalent between Offtime and the sum of Offtime and Ontime in the total cirrhosis group and in cirrhotics without prior OHE. The cut-off for Stroop OnTime +OffTime was 274.9 seconds. A majority (56%, n=70) of the 125 cirrhotics took >274.9 seconds to complete the App while when those only without prior OHE were considered it was 43% (n=35). The ROC for Stroop to diagnose MHE in patients with and without prior OHE was similarly highest with Ontime+Offtime with AUC of 0.77 and 0.79 respectively using weighted lures as the gold standard and 0.87 and 0.82 respectively for PHES (Supplementary tables 2 and 3). We used logistic regression to assess the effect of age, MELD and Ontime+Offtime (education, ammonia and sodium were not significant on univariate analysis and were not considered further) on the division of the group into MHE or not using the three modalities. Using standard tests, age and MELD were significant predictors (p=0.032 and p=0.0007 respectively) on their own but when Ontime+Offtime was added, this prediction became non-significant (p=0.89 and p=0.32) while Ontime+Offtime had significance of p<0.0001. This pattern was again repeated when PHES (age p=0.02, MELD,p=0.01 alone and became p=0.94 and p=0.36 after Ontime+Offtime was introduced which had p<0.0001) or WL (age p=0.62, MELD p=0.003 before and became p=0.41 and p=0.69 after Ontime+Offtime was introduced which had p<0.0001) were used to diagnose MHE.

Prospective validation cohort

We recruited 43 additional cirrhotic patients (age 56±7 years, education 12±2 years, MELD score 11±5) for validation of which 12 had controlled OHE. The mean venous ammonia was 42±13 mg/dl and sodium was 137±4.6 meq/L. Using the cut-off of 2 abnormal tests, 22 patients (50%) had MHE diagnosed by standard criteria while using the ROC value of Ontime+Offtime >274.9 seconds, 19 (44%) were impaired. The sensitivity for this ROC value was 78% (17 of the 22 impaired by standard tests were impaired on the app) while the specificity was 90% (19 of the 21 unimpaired on standard tests were unimpaired on the app, Figure 3C, Table 4).

Comparison of the new test administrator to the remainder

Of the 43 patients in the validation cohort, 25 were administered the App by the new investigator. When these 25 patients were compared to the remaining cirrhotic patients in the study (n=143), there were no significant differences in age (57±7 vs. 59±5 years, p=0.1), MELD score (12±6 vs. 12±7, p=0.9) or in App Results Offtime (133±35 vs. 127±34 seconds, p=0.6), OnTime (166±53 vs. 154±55 seconds,p=0.5), trials in the off state (median 6 vs. 5.5, p=0.1) and on state (median 6 vs 5.5, p=0.5).

Longitudinal analysis

The Stroop tests were re-administered 40±14 days apart. The results of the Stroop Offtime and Ontime remained statistically similar between the two testing periods for controls and cirrhotics with prior OHE. While the Offtime remained similar, there was a significant reduction in Ontime in cirrhotics without OHE after retesting (Table 5).

Discussion

We found that the Stroop smartphone app is able to detect cognitive dysfunction, has good discriminative validity and test-retest reliability in cirrhosis. The EncephalApp_Stroop was easy to administer, quick to teach to subjects, and simple to score and interpret. Use of this convenient App may improve the screening process and subsequent treatment rates in potential patients with MHE, especially in the US where the testing is not routinely performed. Abnormalities in attention and psychomotor speed are the hallmark of cognitive impairment in MHE and negatively impact on patient's quality of life (2). The anterior attention system is hypothesized to modulate response inhibition, behavioral selection and executive control, which are cognitive skills necessary to perform the Stroop task, as well as other tests in our cognitive battery (e.g., the ICT and Digit Symbol Test) (20). Studies have shown that the Stroop paradigm is sensitive to cognitive change in MHE. Specifically, MHE patients are much more likely to show impairment on tasks assessing the integrity of the anterior, compared to the posterior attention system (13, 14). We indeed found that the Stroop task's psychomotor speed component (Offtime and Ontime) was highly correlated with a wide range of other cognitive domains; specifically response inhibition (ICT lures), visuo-motor coordination (block design), set shifting (number connection-B) and not just with tests of psychomotor speed and accuracy (serial dotting, line drawing, and digit symbol).

The "gold standard" for MHE/cognitive dysfunction diagnosis varies across populations, therefore we used three methods (Standard tests, ICT and PHES) to test the ability of the App to differentiate between groups(21). We found that the App performance was worse in impaired patients with or without OHE defined by any of the three techniques. This was further confirmed by applying the App cut-off in a prospective validation cohort with similar discriminative capability. This could potentially increase the applicability of this App regardless of the standard used for MHE/cognitive dysfunction.

The Stroop task, regardless of whether it is the, "on," or "off" condition, from a reaction time perspective, requires an individual to allocate attentional resources and focus on the visually presented stimulus, determine its color, and provide a motor response by pressing the appropriate choice on the screen(13). In our study the slowness seen in cirrhosis may be due to problems in cognitive processing and/or problems in motor speed(22). The cognitive processing demand is higher during the "On" state and correspondingly we found that Ontimes were always greater than Offtimes in all groups. While the accuracy of responses was lower and Ontime minus Offtime was greater in more affected patients, psychomotor speed variables (Offtime and Ontime themselves) were the best differentiators. This could

be due to the design of the task in which the App run stops as soon as a subject makes a mistake, providing fewer opportunities to get accuracy data. These mistakes would demonstrate to the subject that their error and could potentially reinforce the test rules, preventing them from making future errors. Overall there was a significant correlation between Ontime and Offtime with all cognitive tests spanning several domains. The common denominator linking achievement in all the cognitive tests with Stroop Ontime and Offtime is psychomotor speed(22). Psychomotor speed in cirrhosis can be affected by central and peripheral causes, including HE with the accompanying cognitive deficits, neuromuscular weakness and incoordination(5). While not specifically studied here, Stroop abnormalities have been associated with poor connectivity between anterior cingulate cortex, dorsolateral prefrontal cortex and posterior parietal lobes in MHE and prior overt HE(23). However it is possible that worsening psychomotor speed in advanced cirrhotic patients could have a peripheral, neuromuscular component, as is also reflected in the poor cognitive performance in tests other than Stroop. Therefore the Ontime minus Offtime variable was created to control for psychomotor speed and provide a measure of cognitive processing and this was also significantly worse in affected patients. However, this measure was specific (AUC 0.71–0.86), but not sensitive (AUC 0.44–0.54) in differentiating groups on ROC analysis. This could indicate that the psychomotor impairment rather than the "Stroop" component of the App may be responsible for its discriminating ability. However, regardless of the specific underlying cognitive deficit, the App was able to replicate the end-result of the other tests in our population. As expected, we found that advanced cirrhosis and patients with prior OHE, was associated with worse cognitive and App performance in both speed and accuracy(6). This replicates prior studies and proves that this app has discriminative ability for patients from the early to advanced stage of cirrhosis, provided they can understand the task(6, 12, 24, 25). It is also interesting that patients with prior OHE remain cognitively challenged and did not improve on repeated testing(25). This was in contrast to the practice effect learning demonstrated in patients without OHE in the Ontime (which is a measure of response inhibition and motor speed) but not in the Offtime condition (which primarily assesses psychomotor ability). This phenomenon has been noted with other studies of cirrhotics without prior OHE(24, 25). Therefore with this app, like other tests for MHE, the learning effect has to be considered for therapeutic trials(7, 22, 26). We also found that patients with alcoholic liver disease, despite similar MELD score, were more likely to perform worse on the App. The effect of alcohol on cognitive performance even after cirrhosis has set in has been described before and is additional evidence of the discriminative validity of the App; interestingly this difference was confined to psychomotor speed since cognitive processing (Ontime minus Offtime) was similar between groups(27). We did not find any differences in App performance in patients with hepatitis C compared to others. This could be due to the predominant pre-cirrhotic effect of hepatitis C on cognition which is overwhelmed by the impairment due to cirrhosis itself(28).

We found that all subjects, despite prior unfamiliarity with smartphones in the majority, were able to adapt to the iPod screen and follow instructions to complete the app in accordance with the research staff instructions. The administering staff was also able to explain the task and export the data easily after completion. We also found that the App results achieved by a new administrator of the App were in line with prior overall results and the App results between centers were individually and similarly related to age and cirrhosis severity and independent on regression, which could increase its generalizability. The app is limited in that it requires a compatible device to operate; however considering the high financial burden of untreated MHE, it could be a worthwhile upfront cost(29, 30). Our study is limited by the relatively high educational status of cirrhotic patients and these findings may change when applied to persons of lower educational status. In addition, the demonstration of cognitive impairment with any of the tests, included the App should be

interpreted in light of the subjects' clinical status since all these modalities are sensitive but not specific for MHE(21).

We conclude that the Stroop app is a valid and reliable method for screening for MHE. Further studies evaluating its validity in other populations are needed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

List of Abbreviations

|	
SONIC	spectrum of neuro-cognitive impairment in cirrhosis
MHE	Minimal hepatic encephalopathy
QOL	Quality of life
OHE	overt hepatic encephalopathy
PHES	Psychometric hepatic encephalopathy score
BDT	block design test
DST	digit symbol test
LTT	line tracing time
ICT	inhibitory control test
WL	weighted lures
NCT-A/B	number connection test A/B
SD	standard deviations

References

1. Ferenci P, Lockwood A, Mullen K, Tarter R, Weissenborn K, Blei AT. Hepatic encephalopathy--definition, nomenclature, diagnosis, and quantification: final report of the working party at the 11th World Congresses of Gastroenterology, Vienna, 1998. Hepatology. 2002; 35:716–721. [PubMed: 11870389]

2. Bajaj JS, Wade JB, Sanyal AJ. Spectrum of neurocognitive impairment in cirrhosis: Implications for the assessment of hepatic encephalopathy. Hepatology. 2009; 50:2014–2021. [PubMed: 19787808]

3. Bajaj JS, Saeian K, Schubert CM, Hafeezullah M, Franco J, Varma RR, Gibson DP, et al. Minimal hepatic encephalopathy is associated with motor vehicle crashes: the reality beyond the driving test. Hepatology. 2009; 50:1175–1183. [PubMed: 19670416]

4. Ortiz M, Jacas C, Cordoba J. Minimal hepatic encephalopathy: diagnosis, clinical significance and recommendations. J Hepatol. 2005; 42(Suppl):S45–53. [PubMed: 15777572]

5. Soriano G, Roman E, Cordoba J, Torrens M, Poca M, Torras X, Villanueva C, et al. Cognitive dysfunction in cirrhosis is associated with falls: a prospective study. Hepatology. 2012; 55:1922–1930. [PubMed: 22213000]

6. Dhiman RK, Kurmi R, Thumburu KK, Venkataramarao SH, Agarwal R, Duseja A, Chawla Y. Diagnosis and prognostic significance of minimal hepatic encephalopathy in patients with cirrhosis of liver. Dig Dis Sci. 2010; 55:2381–2390. [PubMed: 20508990]

7. Bajaj JS, Heuman DM, Wade JB, Gibson DP, Saeian K, Wegelin JA, Hafeezullah M, et al. Rifaximin improves driving simulator performance in a randomized trial of patients with minimal hepatic encephalopathy. Gastroenterology. 2011; 140:478–487. e471. [PubMed: 20849805]

8. Prasad S, Dhiman RK, Duseja A, Chawla YK, Sharma A, Agarwal R. Lactulose improves cognitive functions and health-related quality of life in patients with cirrhosis who have minimal hepatic encephalopathy. Hepatology. 2007; 45:549–559. [PubMed: 17326150]

9. Sharma P, Sharma BC, Agrawal A, Sarin SK. Primary prophylaxis of overt hepatic encephalopathy in patients with cirrhosis: an open labeled randomized controlled trial of lactulose versus no lactulose. J Gastroenterol Hepatol. 2012; 27:1329–1335. [PubMed: 22606978]

10. Sidhu SS, Goyal O, Mishra BP, Sood A, Chhina RS, Soni RK. Rifaximin improves psychometric performance and health-related quality of life in patients with minimal hepatic encephalopathy (the RIME Trial). Am J Gastroenterol. 2011; 106:307–316. [PubMed: 21157444]

11. Bajaj JS, Etemadian A, Hafeezullah M, Saeian K. Testing for minimal hepatic encephalopathy in the United States: An AASLD survey. Hepatology. 2007; 45:833–834. [PubMed: 17326210]

12. Weissenborn K, Ennen JC, Schomerus H, Ruckert N, Hecker H. Neuropsychological characterization of hepatic encephalopathy. J Hepatol. 2001; 34:768–773. [PubMed: 11434627]

13. Amodio P, Schiff S, Del Piccolo F, Mapelli D, Gatta A, Umilta C. Attention dysfunction in cirrhotic patients: an inquiry on the role of executive control, attention orienting and focusing. Metab Brain Dis. 2005; 20:115–127. [PubMed: 15938130]

14. Felipo V, Ordono JF, Urios A, El Mlili N, Gimenez-Garzo C, Aguado C, Gonzalez-Lopez O, et al. Patients with minimal hepatic encephalopathy show impaired mismatch negativity correlating with reduced performance in attention tests. Hepatology. 2012; 55:530–539. [PubMed: 21953369]

15. Company L, Zapater P, Perez-Mateo M, Jover R. Extrapyramidal signs predict the development of overt hepatic encephalopathy in patients with liver cirrhosis. Eur J Gastroenterol Hepatol. 2010; 22:519–525. [PubMed: 20010298]

16. Randolph C, Hilsabeck R, Kato A, Kharbanda P, Li YY, Mapelli D, Ravdin LD, et al. Neuropsychological assessment of hepatic encephalopathy: ISHEN practice guidelines. Liver Int. 2009; 29:629–635. [PubMed: 19302444]

17. Bajaj JS, Hafeezullah M, Franco J, Varma RR, Hoffmann RG, Knox JF, Hischke D, et al. Inhibitory control test for the diagnosis of minimal hepatic encephalopathy. Gastroenterology. 2008; 135:1591–1600. e1591. [PubMed: 18723018]

18. Amodio P, Ridola L, Schiff S, Montagnese S, Pasquale C, Nardelli S, Pentassuglio I, et al. Improving the inhibitory control task to detect minimal hepatic encephalopathy. Gastroenterology. 2010; 139:510–518. 518, e511–512. [PubMed: 20470775]

19. Romero-Gomez M, Cordoba J, Jover R, del Olmo JA, Ramirez M, Rey R, de Madaria E, et al. Value of the critical flicker frequency in patients with minimal hepatic encephalopathy. Hepatology. 2007; 45:879–885. [PubMed: 17393525]

20. Posner MI, Petersen SE. The attention system of the human brain. Annu Rev Neurosci. 1990; 13:25–42. [PubMed: 2183676]

21. Cordoba J. New assessment of hepatic encephalopathy. J Hepatol. 2011; 54:1030–1040. [PubMed: 21145874]

22. Weissenborn K, Giewekemeyer K, Heidenreich S, Bokemeyer M, Berding G, Ahl B. Attention, memory, and cognitive function in hepatic encephalopathy. Metab Brain Dis. 2005; 20:359–367. [PubMed: 16382346]

23. Zhang LJ, Yang G, Yin J, Liu Y, Qi J. Neural mechanism of cognitive control impairment in patients with hepatic cirrhosis: a functional magnetic resonance imaging study. Acta Radiol. 2007; 48:577–587. [PubMed: 17520437]

24. Riggio O, Ridola L, Pasquale C, Nardelli S, Pentassuglio I, Moscucci F, Merli M. Evidence of persistent cognitive impairment after resolution of overt hepatic encephalopathy. Clin Gastroenterol Hepatol. 2010; 9:181–183. [PubMed: 20951832]

25. Bajaj JS, Schubert CM, Heuman DM, Wade JB, Gibson DP, Topaz A, Saeian K, et al. Persistence of cognitive impairment after resolution of overt hepatic encephalopathy. Gastroenterology. 2010; 138:2332–2340. [PubMed: 20178797]

26. Bajaj JS, Cordoba J, Mullen KD, Amodio P, Shawcross DL, Butterworth RF, Morgan MY. Review article: the design of clinical trials in hepatic encephalopathy--an International Society for Hepatic Encephalopathy and Nitrogen Metabolism (ISHEN) consensus statement. Aliment Pharmacol Ther. 2011; 33:739–747. [PubMed: 21306407]

27. Sorrell JH, Zolnikov BJ, Sharma A, Jinnai I. Cognitive impairment in people diagnosed with end-stage liver disease evaluated for liver transplantation. Psychiatry Clin Neurosci. 2006; 60:174–181. [PubMed: 16594941]

28. Bajaj JS, Ahluwalia V, Wade JB, Sanyal AJ, White MB, Noble NA, Monteith P, et al. Asymmetric dimethylarginine is strongly associated with cognitive dysfunction and brain MR spectroscopic abnormalities in cirrhosis. J Hepatol. 2012

29. Bajaj JS, Pinkerton SD, Sanyal AJ, Heuman DM. Diagnosis and treatment of minimal hepatic encephalopathy to prevent motor vehicle accidents: a cost-effectiveness analysis. Hepatology. 2012; 55:1164–1171. [PubMed: 22135042]

30. Huang E, Esrailian E, Spiegel BM. The cost-effectiveness and budget impact of competing therapies in hepatic encephalopathy - a decision analysis. Aliment Pharmacol Ther. 2007; 26:1147–1161. [PubMed: 17894657]

A: Time required for Stroop App completion in the On and Off states in all cirrhotic patients
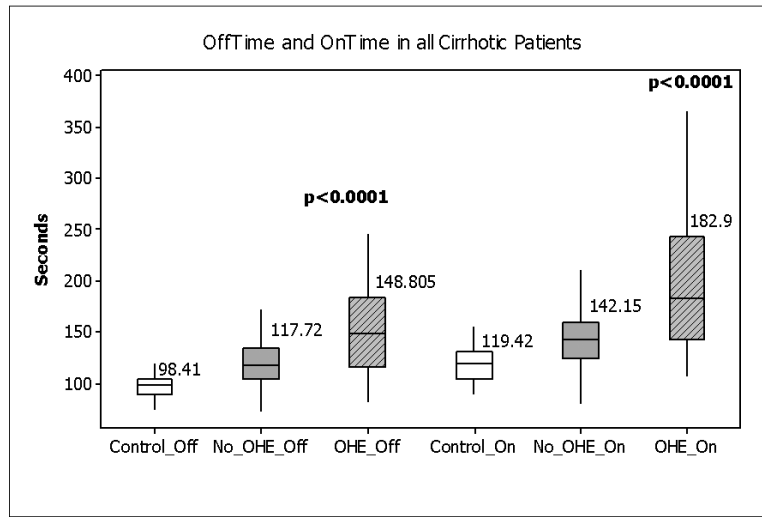
B: Time required for Stroop App completion in the On and Off states in patients without prior OHE (MHE defined by Standard Tests
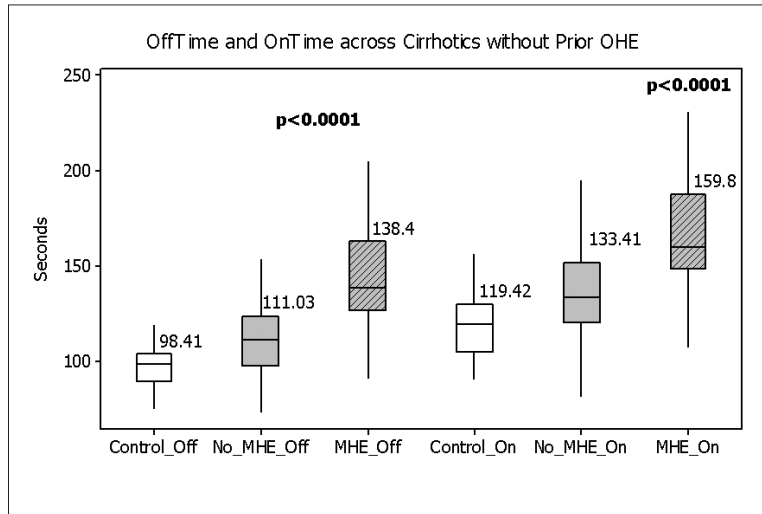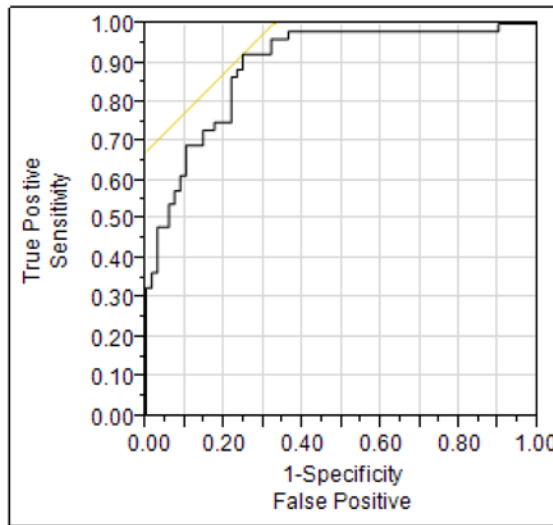

Figure 1.
Median and IQR of time required to complete the App in the "Off" and "On" states. Figure 1A shows the distribution of values in the entire group in which control values are compared to cirrhotics without prior OHE and those with prior OHE. Figure 1B shows the comparison between controls and cirrhotics with MHE and no MHE based on standard tests. No_OHE: no prior overt hepatic encephalopathy, OHE: prior overt HE, No_MHE: no MHE and MHE: minimal hepatic encephalopathy. All comparisons were highly statistically significant

A: ROC curve of diagnosis of MHE using Standard tests as the gold standard against Stroop OnTime+OffTime in all cirrhotic patients (AUC 0.89)

B: ROC curve of diagnosis of MHE using Standard tests as the gold standard against Stroop OnTime+OffTime in cirrhotic patients without prior OHE (AUC 0.84)
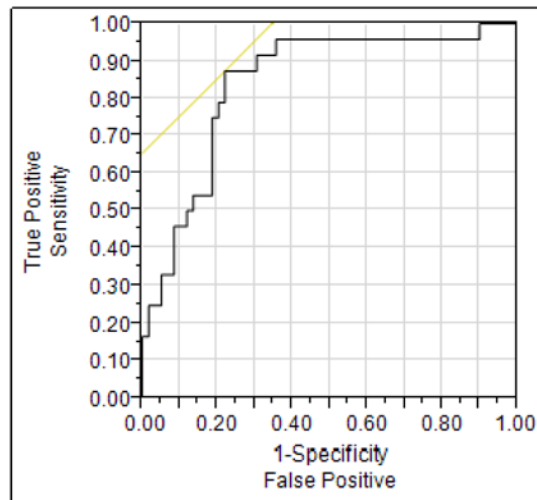
Figure 2.
ROC curve of OnTime+OffTime for all cirrhotic patients (2A) and in those without prior OHE (2B) showing a high AUC for diagnosis when standard tests are used as gold standard. The cut-off was a 274.9 seconds for both populations.

A: Individual value plot of OnTime+OffTime in all cirrhotic patients showing the ROC-generated cut-off (MHE defined by standard tests)
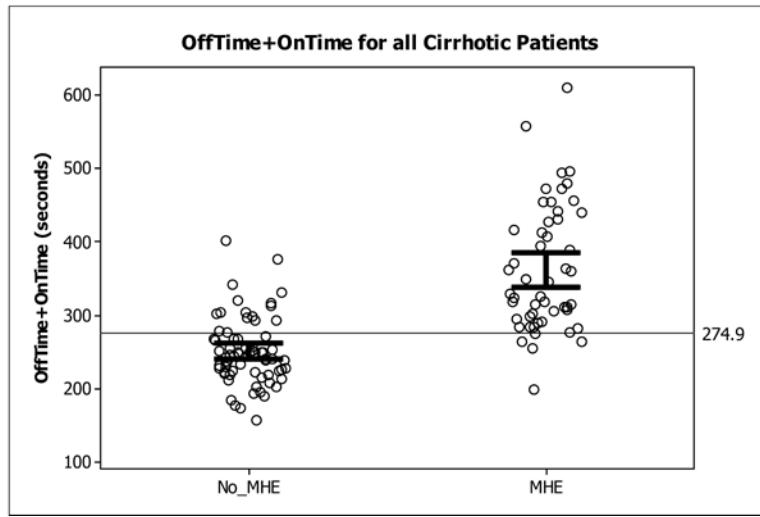
B: Individual value plot of OnTime+OffTime in cirrhotic patients without prior OHE showing the ROC-generated cut-off (MHE defined by standard tests)
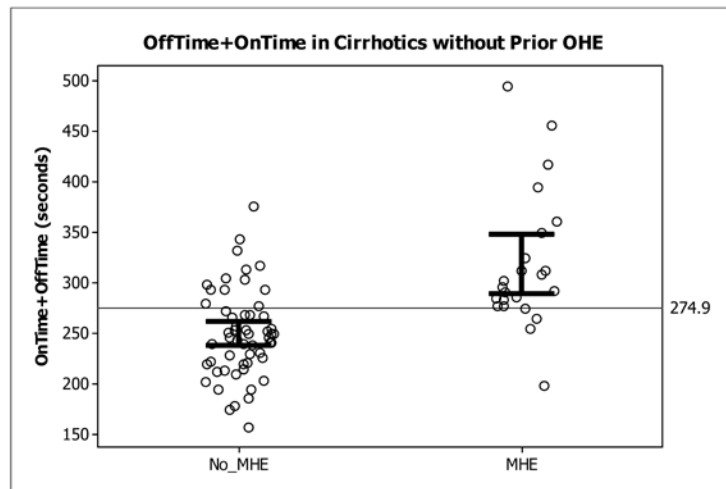
C: Validation of the cut-off generated by the original cohort in a prospectively
collected additional group of cirrhotic patients. (MHE defined by standard tests)
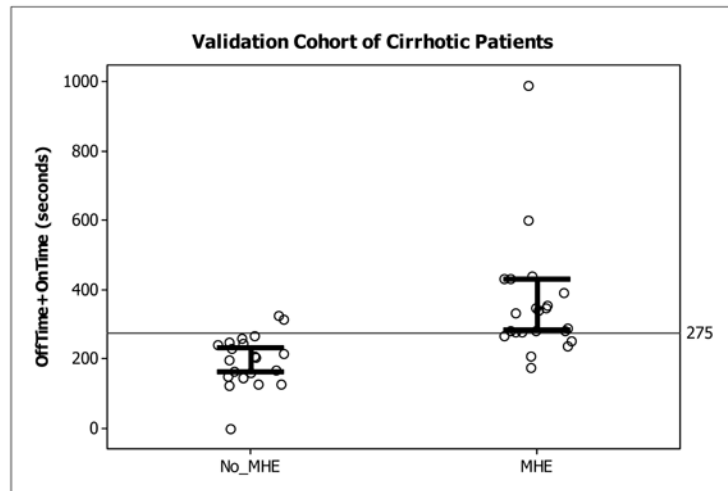
Figure 3.
Individual plots of OnTime+OffTime in the original and prospective validation cohort; the
open circles are individual patients while the bars represent 95% confidence interval for the
mean. Figure 3A shows the cut-off generated by ROC analysis dividing the original 125
cirrhotic patients, Figure 3B shows similar values in the 82 cirrhotics of the original cohort
without prior OHE, and Figure 3C shows the performance of this cut-off in prospective
cirrhotic validation cohort.

Table 1

Cross-sectional comparison of healthy controls[†] compared to cirrhotic patients

	Healthy Controls (n=51)	Cirrhosis (n=125)	
		Without prior OHE (n=82)	With prior OHE (n=43)
Age (years)	55 ± 5	56 ± 6	57 ± 7
Gender (Male/Female)	30/21	51/31	28/15
Race (Caucasian/African American/Hispanic/Other)	34/11/5/1	54/21/7/0	31/11/2/0
Education (years)	14±2[*]	13±2	13±2
Etiology of cirrhosis (HCV/Alc/HCV+Alc/NASH/Other)	–	47/8/4/16/8	21/3/6/6/7
Venous ammonia (mg/dl)	–	43 ± 23	48 ± 19
Serum sodium (mmol/l)	–	137 ± 16	137 ± 4
MELD score	–	9 ± 3	16 ± 7[***]
Number connection-A (sec)	27±7	37±19	52±25[***]
Number connection-B (sec)	69±29	105±77	168±110[***]
Digit symbol (raw score)	73±11	67±29	94±40[***]
Serial Dotting (sec)	51±12	57±16	45±17[***]
Line Tracing (seconds)	80±25	100±37	133±71[***]
Line tracing errors (number)	34±38	37±32	40±33
Block Design (raw score)	36±13	31±15	24±14**
ICT lures (number)	7±5	10±8	14±8[***]
ICT targets (% right)	97±6	96±7	89±18[***]
ICT random (number)	6±2	8±4	14±15[*]
Weighted Lures (number)	9±7	12±12	23±18[***]
MHE using standard tests	–	24 (29%)	31 (72%)[***]
MHE using weighted lures>22	–	15 (18%)	18 (42%)[***]
MHE using PHES >4 SD	–	44 (54%)	34 (79%)[***]
Stroop App results			
Total OffTime (sec)	98±13	121±27	153±40[***]
Median trials for 5 off correct runs (range)	5 (5–9)	5 (5–19)	6 (5–17)[*]
Total OnTime (sec)	119±17	148±38	198±63[***]
Median trials for 5 on correct runs (range)	5 (5–13)	6 (5–16)	6 (5–16)[*]
Total Ontime minus Offtime(sec)	22±13	27±22	47±37[***]
Total Ontime plus Offtime(sec)	217±27	271±60	365±98[***]

ICT: inhibitory control test, OHE: overt hepatic encephalopathy, PHES: psychometric hepatic encephalopathy score.

[*]:p=0.05 – 0.01,

***
:p<0.0001.

†the healthy control results here are those who underwent the Stroop and PHES. The overall results of an additional 83 controls who underwent only the four standard tests and ICT are in the text. Standard Tests are the following: number connection-A/B, digit symbol and block design; two of which need to be abnormal to be considered MHE.

Table 2A

Cognitive test performance between healthy controls and subgroups of cirrhotic patients without prior OHE based on Standard Psychometric tests

MHE defined as 2 of the 4 of NCT-A, NCT-B, DST or BDT abnormal	Cirrhosis without prior OHE (n=82)	
	No MHE (n=55)	**MHE (n=27)**
Number connection-A (sec)	29±8	57±23 ***
Number connection-B (sec)	73±20	182±105 ***
Digit symbol (raw score)	64±13	40±10 ***
Block Design (raw score)	37±15	17±8 ***
Serial Dotting (sec)	51±12	81±29 ***
Line Tracing (seconds)	91±31	120±42 ***
Line tracing errors (number)	35±32	41±33
ICT lures (number)	8±7	16±10 ***
ICT targets (% right)	98±3	89±11 ***
ICT random	7±3	10±6 *
Weighted lures (number)	9±7	22±17 ***
Stroop App Results		
Total OffTime (sec)	112±18	145±28 ***
Median trials for 5 correct off runs (range)	5 (5–12)	6 (5–19)
Total OnTime (sec)	138±28	174±43 ***
Median trials for 5 correct on runs (range)	6 (5–13)	7 (5–16) *
Total Ontime minus Offtime(sec)	22±15	32±22 *
Total Ontime plus Offtime(sec)	249±43	319±68 ***

ICT: inhibitory control test, OHE: overt hepatic encephalopathy. MHE diagnosed on the basis of 134 controls' performance.

*
:p=0.05 – 0.01,

***
:p<0.0001.

Table 2B

Cognitive test performance between healthy controls and subgroups of cirrhotic patients without prior OHE based on PHES

Based on PHES impaired >4 SD	Cirrhosis without prior OHE (n=82)	
	No MHE (n=39)	MHE (n=43)
Number connection-A (sec)	26±7	47±20 ***
Number connection-B (sec)	65±16	138±91 ***
Digit symbol (raw score)	68±14	49±13 ***
Serial Dotting (sec)	51±12	81±29 ***
Line Tracing (seconds)	81±23	117±39 ***
Block Design (raw score)	40±15	25±12 ***
ICT lures (number)	7±6	13±9 ***
ICT targets (% right)	97±5	93±9 ***
ICT random (number)	7±3	8±4
Weighted lures (number)	8±7	16±14**
Stroop App Results		
Total OffTime (sec)	108±18	133±27 ***
Median trials for 5 correct off runs (range)	5 (5–8)	6 (5–19) *
Total OnTime (sec)	130±22	164±40 ***
Median trials for 5 correct on runs (range)	6 (5–12)	6 (5–16)
Total Ontime minus Offtime(sec)	22±13	32±25 *
Total Ontime plus Offtime(sec)	237±37	297±61 ***

ICT: inhibitory control test, OHE: overt hepatic encephalopathy, PHES: psychometric hepatic encephalopathy score. PHES abnormalities are based on data of 51 controls.

*
:p=0.05 – 0.01,

***
:p<0.0001.

Table 2C

Cognitive test performance between healthy controls and subgroups of cirrhotic patients without prior OHE based on Weighted lures

Based on WL >22	Cirrhosis without prior OHE (n=82)	
	No MHE (n=64)	MHE (n=18)
Number connection-A (sec)	33±12	57±31***
Number connection-B (sec)	85±34	196±137**
Digit symbol (raw score)	61±15	41±15***
Serial Dotting (sec)	61±24	95±29**
Line Tracing (seconds)	93±34	129±38***
Block Design (raw score)	34±15	19±13***
ICT lures (number)	7±5	25±5***
ICT targets (% right)	97±5	89±11*
ICT random (number)	7±3	12±5***
Weighted lures (number)	7±6	34±12***
Stroop App Results		
Total OffTime (sec)	117±22	142±36*
Median trials for 5 correct off runs (range)	5 (5–12)	5.5 (5–19)
Total OnTime (sec)	140±27	187±52***
Median trials for 5 correct on runs (range)	6 (5–14)	7 (5–16)
Total Ontime minus Offtime(sec)	22±15	44±23*
Total Ontime plus Offtime(sec)	257±45	329±86***

ICT: inhibitory control test, OHE: overt hepatic encephalopathy; MHE diagnosed on the basis of 134 controls' performance.

*
:p=0.05 – 0.01,

***
:p<0.0001.

Table 3A

: Receiver operating curve results for Stroop for diagnosis of cognitive dysfunction in all cirrhotic patients using standard tests as the gold standard

Variable	AUC	Cut-point	Sensitivity	Specificity
Total time off (Offtime)	0.91	125.84	0.94	0.79
Total time on (Ontime)	0.87	148.7	0.88	0.74
Offtime × no. of runs off	0.85	644.7	0.94	0.66
Ontime × no. of runs on	0.79	982.8	0.71	0.76
Ontime minus Offtime	0.64	41.8	0.44	0.86
Sum of Offtime and Ontime	0.89	274.9	0.92	0.75

Table 3B

Receiver operating curve results for Stroop for diagnosis of cognitive dysfunction in cirrhotic patients without prior OHE using standard tests as the gold standard

Variable	AUC	Cut-point	Sensitivity	Specificity
Total time off (Offtime)	0.87	125.84	0.92	0.79
Total time on (Ontime)	0.80	148.7	0.79	0.76
Offtime × no. of runs off	0.80	631.5	0.96	0.63
Ontime × no. of runs on	0.72	1070.3	0.54	0.83
Ontime minus Offtime	0.58	29.8	0.54	0.71
Sum of Offtime and Ontime	0.89	274.9	0.92	0.75

Table 4

Validation cohort

Validation cohort (n=43)	Ofttime+Ontime <274.9 seconds	Ofttime+Ontime >274.9 seconds
Age (years)	54.8±6.7	57.4±6.8
Education (years)	12.9±1.7	11.8±2.9
MELD score	10.6±3.9	11.9±5.9
OHE (%)	14%	50%**
Number connection-A (sec)	35.7±9.8	58.5±31.0***
Number connection-B (sec)	97.5±67.1	190.0±112.0***
Digit symbol (raw score)	59.1±13.7	32.7±13.6***
Serial Dotting (sec)	59.3±17.1	110.5±69.1***
Line Tracing (seconds)	107.1±27.7	119.0±52.1
Line tracing errors (number)	23.6±29.0	73.2±45.4**
Block Design (raw score)	34.0±13.7	32.7±13.6***
ICT lures (number)	8.6±7.4	18.2±10.2***
ICT targets (% right)	96.4±5.5	84.5±15.4***
ICT random (number)	7.2±4.2	16.7±10.6***
Weighted Lures (number)	10.1±10.1	28.6±18.6**

Using the cut-off developed by the original cohort, the validation cohort was significantly impaired on the cognitive tests.

Table 5

: Re-testing of Stroop between baseline and second testing

Controls (n=10)	Baseline (seconds)	Second test (seconds)
Total Offtime	73 ± 11	68 ± 12
Total Ontime	59 ± 7	59 ± 7
Cirrhotic without OHE (n=21)		
Total Offtime	70 ± 12	70 ± 9
Total Ontime	89 ± 13	84 ± 12 [*]
Cirrhotic with prior OHE (n=9)		
Total Offtime	84 ± 20	81 ± 20
Total Ontime	108 ± 46	91 ± 29

[*] : p<0.05 on paired t-test