Validation of EncephalApp, Smartphone-Based Stroop Test, for the Diagnosis of Covert Hepatic Encephalopathy

Jasmohan S. Bajaj,* Douglas M. Heuman,* Richard K. Sterling,* Arun J. Sanyal,* Muhammad Siddiqui,* Scott Matherly,* Velimir Luketic,* R. Todd Stravitz,* Michael Fuchs,* Leroy R. Thacker,‡ HoChong Gilles,* Melanie B. White,* Ariel Unser,* James Hovermale,* Edith Gavis,* Nicole A. Noble,* and James B. Wade§

*Department of Gastroenterology, Hepatology and Nutrition, ‡Department of Biostatistics, and §Department of Psychiatry, Virginia Commonwealth University and McGuire VA Medical Center, Richmond, Virginia

BACKGROUND & AIMS: **Detection of covert hepatic encephalopathy (CHE) is difficult, but point-of-care testing could increase rates of diagnosis. We aimed to validate the ability of the smartphone app EncephalApp, a streamlined version of Stroop App, to detect CHE. We evaluated face validity, test-retest reliability, and external validity.**

METHODS: **Patients with cirrhosis (n = 167; 38% with overt HE [OHE]; mean age, 55 years; mean Model for End-Stage Liver Disease score, 12) and controls (n = 114) were each given a paper and pencil cognitive battery (standard) along with EncephalApp. EncephalApp has Off and On states; results measured were OffTime, OnTime, OffTime+OnTime, and number of runs required to complete 5 off and on runs. Thirty-six patients with cirrhosis underwent driving simulation tests, and EncephalApp results were correlated with results. Test-retest reliability was analyzed in a subgroup of patients. The test was performed before and after transjugular intrahepatic portosystemic shunt placement, and before and after correction for hyponatremia, to determine external validity.**

RESULTS: **All patients with cirrhosis performed worse on paper and pencil and EncephalApp tests than controls. Patients with cirrhosis and OHE performed worse than those without OHE. Age-dependent EncephalApp cutoffs (younger or older than 45 years) were set. An OffTime+OnTime value of >190 seconds identified all patients with CHE with an area under the receiver operator characteristic value of 0.91; the area under the receiver operator characteristic value was 0.88 for diagnosis of CHE in those without OHE. EncephalApp times correlated with crashes and illegal turns in driving simulation tests. Test-retest reliability was high (intraclass coefficient, 0.83) among 30 patients retested 1–3 months apart. OffTime+OnTime increased significantly (206 vs 255 seconds, $P = .007$) among 10 patients retested 33 ± 7 days after transjugular intrahepatic portosystemic shunt placement. OffTime+OnTime decreased significantly (242 vs 225 seconds, $P = .03$) in 7 patients tested before and after correction for hyponatremia (126 ± 3 to 132 ± 4 meq/L, $P = .01$) 10 ± 5 days apart.**

CONCLUSIONS: **A smartphone app called EncephalApp has good face validity, test-retest reliability, and external validity for the diagnosis of CHE.**

Keywords: Minimal Hepatic Encephalopathy; TIPS; Screening; Stroop Test.

The diagnosis of covert hepatic encephalopathy (CHE) is complicated by the lack of convenient and reliable tools suitable for use in daily practice.[1] Therefore, despite being aware of the negative impact of CHE on quality of life, driving capability, and overall progression to overt hepatic encephalopathy (OHE), most practitioners are not able to detect it.[2] The Stroop test, which evaluates cognitive flexibility and psychomotor speed, has been used to diagnose CHE in

Abbreviations used in this paper: AUC, area under the curve; CHE, covert hepatic encephalopathy; MELD, Model for End-Stage Liver Disease; NCT, number connection test; OHE, overt hepatic encephalopathy; ROC, receiver operating characteristic; SD, standard deviation; SONIC, spectrum of neurocognitive impairment in cirrhosis; TIPS, transjugular intrahepatic portosystemic shunting.

Most current article

1542-3565/$36.00
http://dx.doi.org/10.1016/j.cgh.2014.05.011

paper-pencil and electronic formats.[3–5] The recent use of an app based on Stroop testing by our group showed good differentiation between groups with and without CHE.[6] However, this app was not specifically designed for CHE diagnosis, was not checked for external validity, and needed streamlining to increase its ease of use and reduce time required for completion. Our aim was to validate the new streamlined test, EncephalApp_Stroop, for the diagnosis of CHE with emphasis on face validity, test-retest reliability, and external validity.

Methods

The new Stroop App has the same principles as the one initially used.[6] However, the modifications are to streamline it for user friendliness and speed by reducing the time between runs, inserting language to standardize the instructions to the subjects, increasing the size of the stimuli, and to make it feasible to e-mail results in an Excel (Microsoft Corp, Redmond, WA) sheet. The format includes an easier "Off" state where subjects have to touch the appropriate color of presented ### signs from "green," "blue," and "red," and a difficult "On" state presents discordant-colored stimuli, eg, the word *Green* will be presented in blue color, and the correct answer is blue. The order of runs that need to be completed in turn before the app moves forward is (1) two practice Off runs, (2) five correct test Off runs, (3) two practice On runs, and (4) five correct test On runs. Each run stops when a mistake is made. If the subject requires >20 attempts at any stage, the app stops automatically. We have also created an iPad (Apple Inc, Cupertino, CA) version that was tested against the iPod/iPhone version.

We tested the new app (EncephalApp_Stroop) in terms of (1) face validity: (a) cross-sectional comparison varying severity of cirrhosis and (b) correlation with driving simulation; (2) test-retest reliability; and (3) external validity: (a) before and after transjugular intrahepatic portosystemic shunting (TIPS) and (b) before and after hyponatremia correction.

Face Validity

We recruited outpatient cirrhotic patients (histology or evidence of decompensation, radiologic/endoscopic evidence of cirrhosis, or chronic liver disease with reversed aspartate aminotransferase/alanine aminotransferase ratio) who could consent, were able to understand cognitive tests, were without neurocognitive disorders (apart from being prescribed antidepressants), obtained a Mini-Mental State score of >25, absence of red-green color blindness, no history of alcohol or illicit drug use for 6 months, or focal neurologic deficits. In this group, we included outpatients with prior OHE controlled on lactulose/rifaximin and patients without prior OHE. We then recruited age-balanced healthy people without chronic diseases to serve as controls.

All subjects underwent the following procedures: paper-pencil tests (number connection test [NCT]-A/B, digit symbol, and block design) and then EncephalApp. The CHE gold standard was defined as performance impaired ≥2 standard deviations (SDs) on 2 paper-pencil tests on the basis of our local norms.[6,7] For EncephalApp we chose the OffTime+OnTime >2 SDs beyond the controls on the basis of age groups to decide impaired performance. We performed receiver operating characteristic (ROC) analysis of the cutoffs by using the Youden index generated by OffTime+OnTime compared with the gold standard in all patients and those without prior OHE.

Driving simulation. A group of cirrhotic patients without OHE underwent a validated driving simulation with 3 components[8]: (1) a 10-minute acclimatization run, (2) testing run in which the subject drives through several scenarios, with the outcomes being speeding tickets and crashes, and (3) navigation through a virtual city by using a paper map in which the first turn off the marked path is the outcome termed an illegal turn. Pearson correlations between driving outcomes and standard tests as well as EncephalApp outcomes were performed.

Test-Retest Reliability

A subset of patients whose clinical course remained stable were tested at least 1 month apart with the EncephalApp. Correlation and intraclass coefficient analyses were performed between the first and second administrations to gauge the test-retest reliability of the EncephalApp.

External Validity

Transjugular intrahepatic portosystemic shunting placement. A subset of the patients underwent elective TIPS placement and were tested with the standard tests and EncephalApp at least 1 month after TIPS as an outpatient. No preemptive HE therapy was started after TIPS.

Hyponatremia correction. A subset with hyponatremia and refractory ascites who underwent controlled diuretic withdrawal and fluid restriction for 7 days were tested before and after hyponatremia correction.

Paired *t* tests were used to analyze changes in tests before and after TIPS and hyponatremia correction.

Operational Study

We have developed an identical iPad version of the EncephalApp, and patients were tested with both versions in randomized order one after the other. Correlations between the scores were evaluated by using Pearson correlations.

Statistical analysis for all tests was performed by using SPSS (Chicago, IL) and SAS (Cary, NC) software. This study was approved by the Institutional Review

Board at Virginia Commonwealth University and McGuire VA Medical Centers.

Results

Face Validity

One hundred sixteen healthy controls were compared with 167 patients with cirrhosis; 64 patients had prior OHE (all on lactulose, 20 on additional rifaximin), and 103 patients were without prior OHE. Patients performed significantly worse on most paper-pencil tests and EncephalApp times (Table 1). Finally, EncephalApp times were significantly worse in patients with prior OHE compared with those without (Table 2 and Figure 1*A*, *B*). Individual paper-pencil tests were correlated with the EncephalApp (OffTime all, $P < .0001$ [NCT-A/B, both 0.7; block, −0.5; digit, −0.8], OnTime [NCT-A/B, both 0.7; block, −0.5; digit, −0.7], OffTime+OnTime [NCT-A/B, both 0.7; block, −0.5; digit, −0.8], number of Off runs [NCT-A/B, both 0.4, $P < .0001$; block, −0.3, $P = .002$; digit, −0.4, $P < .0001$], and number of On runs [NCT-A, 0.3, NCT-B, 0.4, both $P < .0001$; block, −0.23, $P = .004$; digit, −0.3, $P < .0001$]).

There was also a significant positive correlation between age and EncephalApp times (all $r = 0.6$, $P < .0001$) but not with number of runs or with education. Model for End-Stage Liver Disease (MELD) score was positively linked with OffTime (0.4, $P < .0001$), OnTime (0.3, $P = .001$), OnTime-OffTime (0.2, $P = .02$), and OffTime+OnTime (0.4, $P < .0001$) but not with number of runs. Because of this dependence on age, we divided the healthy control group into 2 equal halves (<45 and ≥45 years of age) and used separate values in each age group to determine the age-based cutoffs (mean ± 2 SDs) (Table 3 and Figure 1*C*).

Of the 116 controls, 2 were excluded because of their outlier performance on all tests (performed >3 SDs different than the rest of the group). The remaining 114 continued in the study; 65 were ≥45 years of age, and the remaining 49 were <45 years old. Control subjects in the 2 age groups significantly differed in terms of their performance on all paper-pencil tests as well as on EncephalApp. The OffTime+OnTime cutoff value (+2 SDs beyond controls) generated for group <45 years was ≥145 seconds, whereas for those ≥45 years it was 190 seconds.

ROC curves were generated separately for all cirrhotic patients and for those without prior OHE by using paper-pencil tests as the gold standard. In the cohort of all patients, the highest area under the curve (AUC) was with OffTime+OnTime (0.91; 95% confidence interval, 0.86–0.96), followed closely by OffTime (0.89) and OnTime (0.90) alone. AUC was comparatively lower for OnTime-OffTime (0.73) and number of runs Off (0.65) and On (0.68). The best separation was achieved with OffTime+OnTime cutoff >190 seconds, with 89.1% sensitivity and 82.1% specificity in all cirrhotic patients. When only patients without prior OHE were considered, AUC for OffTime+OnTime (0.88) was narrowly higher than that of OnTime (0.88) and OffTime (0.86) alone. AUCs for OffTime-OnTime (0.73) and runs to complete 5 off (0.66) and on runs (0.66) were lower (Figure 2). A cutoff of OffTime+OnTime >190 seconds was associated with 80% sensitivity and 81% specificity.

Although psychomotor speed underlies all the paper-pencil and EncephalApp times, the cognitive flexibility or "set-shifting" aspect of the test was analyzed. To account for this, we calculated NCT-B minus NCT-A and OnTime minus OffTime in EncephalApp. The NCT-B–NCT-A value for the cirrhosis group was 70.1 ± 60.8 seconds,

Table 1. Comparison Between Controls and Cirrhotic Patients

	Cirrhotic patients (n = 167)	Controls (n = 114)	*P* value
Age (*y*)	55 ± 7	54 ± 6	.20
Education (*y*)	14 ± 2	14 ± 3	1.0
Sex (men/women)	119/48	77/37	.51
MELD score	12 ± 4	—	
Etiology, % (HCV/alcohol/alcohol + HCV/NASH/others)	35/17/9/25/14		
OHE (%)	38	—	
Standard tests			
NCT-A	42 ± 24	26 ± 10	<.0001
NCT-B	112 ± 76	63 ± 31	<.0001
Digit symbol	54 ± 18	77 ± 15	<.0001
Block design	30 ± 16	41 ± 17	<.0001
EncephalApp			
OffTime	87 ± 24	62 ± 12	<.0001
OnTime	106 ± 39	73 ± 16	<.0001
No. of runs for Off state	5 (5–15)	5 (5–9)	.34
No. of runs for On state	6 (5–41)	5 (5–29)	.46
OffTime+OnTime	193 ± 61	136 ± 27	<.0001
OnTime–OffTime	21 ± 20	11 ± 7	<.0001

HCV, hepatitis C virus; NASH, nonalcoholic steatohepatitis.

Table 2. Comparison Between OHE and No-OHE Patients

	Without prior OHE (n = 103)	With prior OHE (n = 64)	*P* value
Age (*y*)	57 ± 7	57 ± 4	1.0
Education (*y*)	14 ± 3	13 ± 4	.08
% Men	70	75	.92
MELD score	10 ± 4	15 ± 5	<.0001
Sodium (*mmol/L*)	139 ± 4	136 ± 6	.001
Standard tests			
NCT-A	36 ± 15	52 ± 31	<.0001
NCT-B	96 ± 65	139 ± 86	<.0001
Digit symbol	58 ± 17	46 ± 16	<.0001
Block design	33 ± 16	23 ± 13	<.0001
CHE according to standard tests (%)	28 (27)	37 (58)	<.0001
EncephalApp			
OffTime	80 ± 18	97 ± 28	<.0001
OnTime	97 ± 24	121 ± 54	.001
No. of runs for Off state	5 (5–11)	5 (5–15)	.45
No. of runs for On state	5 (5–17)	6 (5–41)	.63
OffTime+OnTime	177 ± 41	220 ± 79	<.0001
OnTime–OffTime	18 ± 11	27 ± 30	.02

which was significantly correlated with OnTime-OffTime ($P = .5$, $P < .0001$).

Driving simulation. Thirty-six patients without OHE underwent driving simulation after cognitive testing. There was a significant positive correlation between crashes and OffTime (0.4, $P = .03$), OnTime (0.5, $P = .009$), number of runs for the Off (0.4, $P = .03$) and On states (0.4, $P = .05$) and OffTime+OnTime (0.5, $P = .02$). Illegal turns on navigation were also significantly correlated with OffTime (0.5, $P = .009$), OnTime (0.5, $P = .007$), number of runs On (0.4, $P = .02$), and with OffTime+OnTime (0.5, $P = .006$). No other significant correlations including speeding were found.

Test-Retest Reliability

Thirty patients (age, 55.4 ± 10 years; education, 14 ± 2 years; etiology: 10 hepatitis C virus, 4 alcohol, 3 hepatitis C virus + alcohol, 9 nonalcoholic steatohepatitis, and rest others; MELD score, 10.6 ± 3.3; 10 with OHE all on lactulose) underwent EncephalApp twice 2 ± 1 months apart by using paired *t* tests. During this time, the patients did not have any change in their liver disease or HE course. As a whole, the group did not show significant change in standard or EncephalApp results (Supplementary Table 1; baseline OffTime+OnTime 182 ± 56 vs repeat OffTime+OnTime 188 ± 69, $P = .45$). There was a significant correlation between the 2 OffTime+OnTime results ($r = 0.941$, $P < .0001$; Supplementary Figure 1), and the intraclass coefficient was 0.832 (95% confidence interval, 0.65–0.92; $P < .0001$).

External Validity

Before and after transjugular intrahepatic portosystemic shunting. Ten patients (age, 56 ± 7 years; education, 12 ± 2 years; 5 hepatitis C virus, 3 alcohol, and 2 nonalcoholic steatohepatitis, 2 OHE pre-TIPS on lactulose) underwent an elective TIPS. Eight patients underwent TIPS for refractory ascites, and 2 had hepatic hydrothorax. Patients were tested 12 ± 3 days before and retested 33 ± 7 days after TIPS as outpatients with

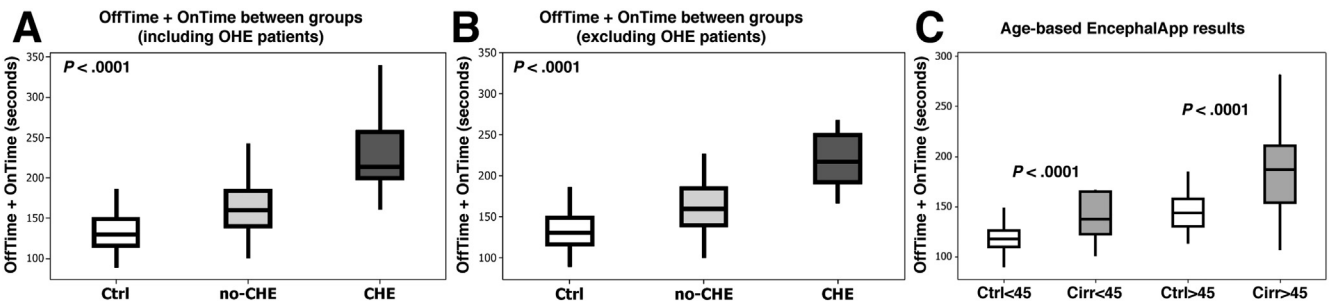


Figure 1. Comparison of OffTime+OnTime between groups. CHE on paper-pencil tests. Mean ± 95% confidence interval depicted in the *bars*. (*A*) Significantly higher OffTime+OnTime in patients with CHE than those without CHE in entire cirrhosis (Cirr) group compared with controls. (*B*) Significantly higher OffTime+OnTime in patients with CHE than those without CHE in patient group without prior OHE compared with controls. (*C*) Significantly higher OffTime+OnTime in subjects ≥45 years compared with those <45 years in both control and patient groups. Ctrl, controls.

Table 3. Analysis of Age-Divided Responses Between Groups

	Age <45 y			Age ≥45 y		
	Control (n = 49)	Cirrhotic (n = 24)	*P* value, control vs cirrhotic	Control (n = 65)	Cirrhotic (n = 142)	*P* value, control vs cirrhotic
NCT-A	22 ± 7	26 ± 5	.007	31 ± 11	43 ± 24	<.0001
NCT-B	58 ± 30	71 ± 23	.04	64 ± 24	114 ± 77	<.0001
Digit symbol	83 ± 14	72 ± 16	.004	71 ± 14	53 ± 18	<.0001
Block design	48 ± 17	35 ± 14	<.0001	36 ± 13	30 ± 16	.12
EncephalApp						
OffTime	55 ± 6	61 ± 8	.002	66 ± 9	88 ± 24	<.0001
OnTime	65 ± 11	78 ± 15	.001	78 ± 11	108 ± 41	<.0001
No. of runs for Off state	5	5	.42	6	5	.12
No. of runs for On state	5	5	.57	5	6	.04
OffTime+OnTime	120 ± 15	140 ± 22	<.0001	144 ± 23	197 ± 63	<.0001
OnTime–OffTime	10 ± 6	17 ± 9	.002	12 ± 5	21 ± 21	.05

NOTE. All tests were impaired in age group ≥45 y compared with those <45 y in both control and patient groups.

all the cognitive tests. MELD score did not change (12 ± 1 before and 13 ± 2 after TIPS), and none of the patients developed OHE episodes before the testing. There was a significant worsening of cognitive performance (NCT-A: 45 vs 53 seconds, $P = .03$; NCT-B: 125 vs 189 seconds, $P = .004$; Digit symbol test: 53 vs 39, $P = .01$; Block design: 21 vs 11, $P = .02$) after TIPS. There was also a significant worsening in EncephalApp performance after TIPS (OffTime 87 vs 110 seconds, $P = .001$; OnTime 124 vs 140, $P = .002$; and OffTime+OnTime 206 vs 255, $P = .007$) without change in runs required to achieve 5 correct runs (Figure 3*A*).

Pre- and post-hyponatremia correction. Seven cirrhotic patients with hyponatremia (Na <130 meq/L) were studied with EncephalApp before and after improvement of sodium. Patients underwent withdrawal of diuretics (median spironolactone dose 150 to 50 mg and furosemide 80 to 20 mg/day) and were advised fluid restriction (1.5 L per day) for at least 7 days (10 ± 5 days between testing). There was a significant increase in Na from 126 ± 3 to 132 ± 4 meq/L ($P = .01$). EncephalApp performance significantly improved from a time perspective (OffTime 101 vs 92 seconds, OnTime 137 vs 121 seconds, OffTime+OnTime, 242 vs 225 seconds; $P = .03$) after Na correction but not in terms of median runs needed for 5 correct runs in Off (6 vs 6) or On states (6 vs 6) (Figure 3*B*).

Operational Analysis

iPad vs iPod. Twenty-seven subjects (8 controls and 19 cirrhotic patients) underwent sequential testing with these 2 modes of administration in random order (15 had iPad first and 12 had iPod first). There were no differences in EncephalApp results, and the OffTime+OnTime were significantly correlated (Supplementary Table 2 and Supplementary Figure 2).

Discussion

CHE is a part of the spectrum of neurocognitive impairment in cirrhosis (SONIC) that is difficult to diagnose by using simple clinical examination.[1,9] In our

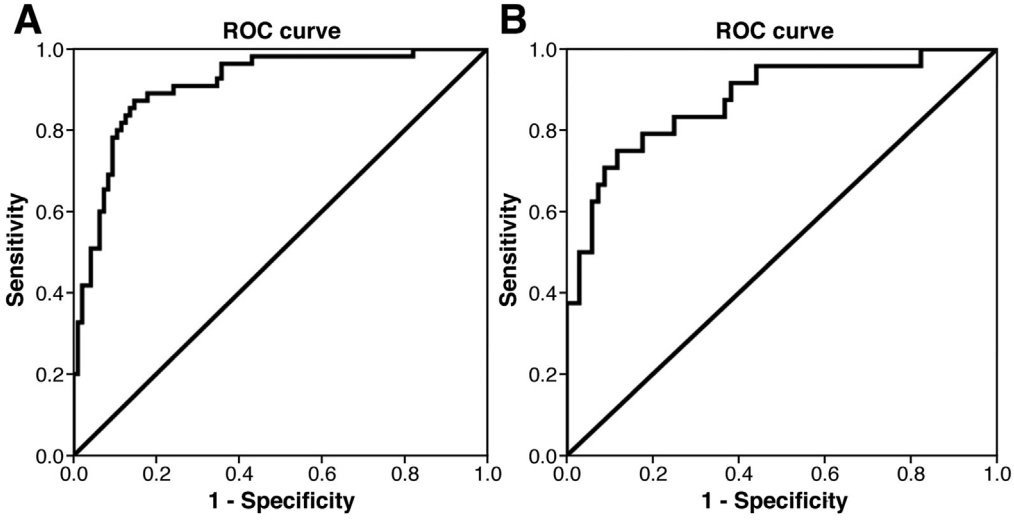


Figure 2. ROC analysis of OffTime+OnTime against paper-pencil tests. (*A*) In all patients (AUC, 0.91); (*B*) in patients without prior OHE (AUC, 0.88).

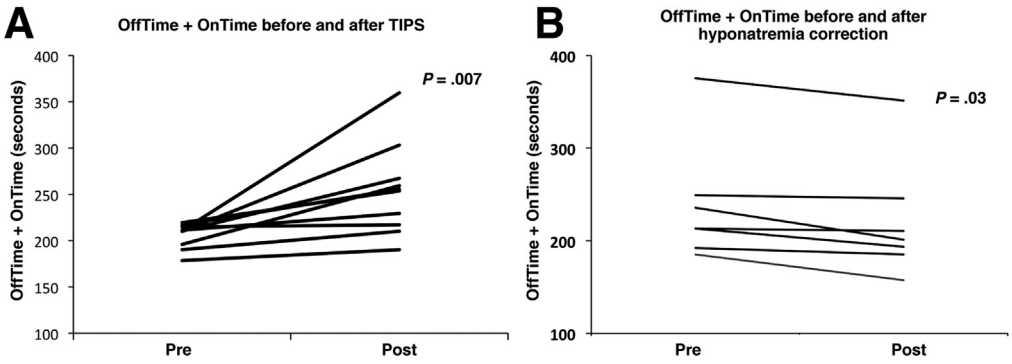


Figure 3. Change in EncephalApp results with TIPS and hyponatremia correction. (*A*) OffTime+ OnTime before/after TIPS placement. (*B*) OffTime+ OnTime before/after hyponatremia correction.

study, we found that the EncephalApp is able to detect CHE on the basis of the gold standard, paper-pencil tests.[10] Although these measures are used to diagnose CHE in research studies, they have not found traction in clinical practice because of their copyrighted nature, need for psychological expertise, and time requirements.[11] Therefore, a point-of-care strategy with or without cognitive tests, which could be used by a medical assistant before clinic, can increase testing rates.[12] The diagnosis of CHE has clinical relevance because it is associated with poor health-related quality of life, increased progression to OHE, difficulty in driving and work, and negatively affects socioeconomic status.[11,13–15] Recent studies have also demonstrated that CHE treatment can improve cognitive outcomes; however, treatment of CHE without testing is not necessarily cost-effective.[16,17] Our results are in sync with the prior experience with the old Stroop App in that in similar prior HE studies, we found that the psychomotor function, ie, the time required to complete tasks, was much more predictive than the errors committed (number of runs required) and the cognitive flexibility measures (OnTime-OffTime) in differentiating groups.[18–20] We found good test-retest reliability, suggesting the app is potentially suitable for repeated testing. We extended prior knowledge by correlating the results with driving simulation and testing the external validity.

We found that despite adequate treatment, patients with prior OHE had worse cognitive performance on most paper-pencil tests and on EncephalApp. EncephalApp performance was also correlated with MELD score. The results reiterate that even patients with treated OHE remain cognitively impaired despite having a clear mental state, which was adequately detected by the app.[21,22] Although we did not find any education dependence, age was a significant factor in determining the results. Therefore, age-based cutoffs were developed. Interestingly, because most patients were older than 45 years of age, the ROC inflection point cutoff (>190 seconds) was similar to that determined when 2 SDs were added to the >45-year-old controls' OffTime+OnTime result. Our patients had a high standard of education, which may have mitigated the potential influence of education on these results. Further validation is needed in those with fewer years of formal education.

Going beyond simple cognitive testing into potential real-world actions such as driving or driving simulation brings into focus how these results can potentially impact functioning in the real world.[23] We found that the times and, to a lesser extent, runs needed for completion were correlated with navigation skills and crashes, but not speeding. As shown in prior studies, crashes reflect a breakdown in reaction time, visuomotor coordination, and psychomotor speed, whereas successful navigation requires these same cognitive skills, along with intact working memory.[8] Speeding instances represent a failure of divided attention, ie, inability to concentrate simultaneously on driving and paying attention to the speed limit and the speedometer. The app is not specifically geared to test this domain. Although simulation does not have the vestibular component of real driving, it can ensure similar driving situations without medicolegal issues for research testing. Therefore, the app specifically was correlated to basic domains involved in driving, reaction time, working memory, and psychomotor speed, which highlights its face validity.

The results also indicate that the app results, especially those centered on psychomotor speed, track the underlying patient status. The results significantly worsened after TIPS placement and improved after sodium increase. Prior studies have shown that cognitive ability worsens after TIPS placement, which was confirmed by EncephalApp.[24] Trials have also demonstrated quality of life enhancement with hyponatremia correction, which our results extended further into the realm of EncephalApp performance.[25,26] These 2 specific situations are important because they represent patients in end-stage liver disease where the cognitive impairment burden is at its extreme compared with that in early-stage cirrhotic patients. Although the cross-sectional portion of this study demonstrated the ability of EncephalApp to differentiate between patients relatively early in their disease course, these results show that it can also change with underlying patient status change in the end stages. This range of functioning increases the applicability of this app across SONIC and increases the breadth of its application.

It is intriguing that the consistent differentiators between unaffected and affected groups are the indices of psychomotor speed and reaction time (OffTime, OnTime)

and not measures of cognitive flexibility (OnTime-OffTime) or those of accuracy (number of runs required).[20] All paper-pencil tests were also correlated with the time results individually. Interestingly, this correlation also extended to psychomotor speed–independent measures, NCT-B–NCT-A and OnTime-OffTime, indicating that several aspects of paper-pencil tests and EncephalApp were impaired in the same direction in affected groups. The task stops when mistakes are made; therefore, each mistake or requirement to start a new run increases the insight of the subject into their poor response and could potentially improve subsequent performance. When the ROC analysis was performed, however, the measures of psychomotor speed overshadowed those of cognitive flexibility and accuracy in differentiating groups. Therefore, even though all domains tested by EncephalApp were more impaired in affected groups, the final common pathway of most responses, ie, the psychomotor speed, remained the dominant factor in differentiating between subjects.[20]

Because of the average age of the subjects involved, the iPad was considered to be an easier approach; however, because the size of the stimuli presented would now increase several fold, we systematically tested the results. We did not find a significant change in the EncephalApp outcomes regardless of the device, which is encouraging and could be due to the built-in training runs that precede each administration.

Our results replicate prior studies of the paper Stroop tests in cirrhosis and also our study with the old App in an entirely new set of subjects and furthers it with study of validation and age-based cutoffs.[3–6] The relatively high educational background of our subjects may limit its generalizability. Our age range was limited for norms because of the age of the majority of cirrhotic patients in our clinic. An additional potential limitation of using this app is that psychomotor speed was the ultimate differentiator, which can be affected by several disorders unrelated to CHE. Therefore, like most cognitive tests, this test is not specific for CHE. From a logistic standpoint, this app is only available for Apple devices, which could also limit its use. The app has been translated into several other languages (although only the English version was used for this study), is available for free download from iTunes, and detailed Webcasts of administration and interpretation instructions are available at www.chronicliverdisease.org. Validation in other languages and cultures is also needed.

We conclude that the EncephalApp has good face validity, test-retest reliability, and external validity for the diagnosis of CHE in patients with cirrhosis. Further trials to validate this are needed and are underway in other centers.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at http://dx.doi.org/10.1016/j.cgh.2014.05.011.

References

1. Kappus MR, Bajaj JS. Covert hepatic encephalopathy: not as minimal as you might think. Clin Gastroenterol Hepatol 2012;10:1208–1219.
2. Bajaj JS, Etemadian A, Hafeezullah M, et al. Testing for minimal hepatic encephalopathy in the United States: an AASLD survey. Hepatology 2007;45:833–834.
3. Amodio P, Schiff S, Del Piccolo F, et al. Attention dysfunction in cirrhotic patients: an inquiry on the role of executive control, attention orienting and focusing. Metab Brain Dis 2005;20:115–127.
4. Zhang LJ, Yang G, Yin J, et al. Neural mechanism of cognitive control impairment in patients with hepatic cirrhosis: a functional magnetic resonance imaging study. Acta Radiol 2007;48:577–587.
5. Felipo V, Ordono JF, Urios A, et al. Patients with minimal hepatic encephalopathy show impaired mismatch negativity correlating with reduced performance in attention tests. Hepatology 2012;55:530–539.
6. Bajaj JS, Thacker LR, Heuman DM, et al. The Stroop smartphone application is a short and valid method to screen for minimal hepatic encephalopathy. Hepatology 2013;58:1122–1132.
7. Ferenci P, Lockwood A, Mullen K, et al. 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.
8. Bajaj JS, Hafeezullah M, Hoffmann RG, et al. Navigation skill impairment: another dimension of the driving difficulties in minimal hepatic encephalopathy. Hepatology 2008;47:596–604.
9. Bajaj JS, Wade JB, Sanyal AJ. Spectrum of neurocognitive impairment in cirrhosis: implications for the assessment of hepatic encephalopathy. Hepatology 2009;50:2014–2021.
10. Weissenborn K, Ennen JC, Schomerus H, et al. Neuropsychological characterization of hepatic encephalopathy. J Hepatol 2001;34:768–773.
11. Ortiz M, Jacas C, Cordoba J. Minimal hepatic encephalopathy: diagnosis, clinical significance and recommendations. J Hepatol 2005;42(Suppl):S45–S53.
12. Nabi E, Thacker LR, Wade JB, et al. Diagnosis of covert hepatic encephalopathy without specialized tests. Clin Gastroenterol Hepatol 2013 Dec 19. Epub ahead of print.
13. Bajaj JS, Riggio O, Allampati S, et al. Cognitive dysfunction is associated with poor socioeconomic status in patients with cirrhosis: an international multicenter study. Clin Gastroenterol Hepatol 2013;11:1511–1516.
14. 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;45:549–559.
15. Bustamante J, Rimola A, Ventura PJ, et al. Prognostic significance of hepatic encephalopathy in patients with cirrhosis. J Hepatol 1999;30:890–895.
16. Mullen KD, Prakash RK. Management of covert hepatic encephalopathy. Clin Liver Dis 2012;16:91–93.
17. Bajaj JS, Pinkerton SD, Sanyal AJ, et al. Diagnosis and treatment of minimal hepatic encephalopathy to prevent motor

vehicle accidents: a cost-effectiveness analysis. Hepatology 2012;55:1164–1171.

18. Schiff S, D'Avanzo C, Cona G, et al. Insight into the relationship between brain/behavioral speed and variability in patients with minimal hepatic encephalopathy. Clin Neurophysiol 2014; 125:287–297.
19. Weissenborn K, Heidenreich S, Ennen J, et al. Attention deficits in minimal hepatic encephalopathy. Metab Brain Dis 2001;16:13–19.
20. Lauridsen MM, Thiele M, Kimer N, et al. The continuous reaction times method for diagnosing, grading, and monitoring minimal/covert hepatic encephalopathy. Metab Brain Dis 2013; 28:231–234.
21. Bajaj JS, Schubert CM, Heuman DM, et al. Persistence of cognitive impairment after resolution of overt hepatic encephalopathy. Gastroenterology 2010;138:2332–2340.
22. Riggio O, Ridola L, Pasquale C, et al. Evidence of persistent cognitive impairment after resolution of overt hepatic encephalopathy. Clin Gastroenterol Hepatol 2010;9:181–183.
23. Kircheis G, Knoche A, Hilger N, et al. Hepatic encephalopathy and fitness to drive. Gastroenterology 2009;137:1706–1715.
24. Riggio O, Nardelli S, Moscucci F, et al. Hepatic encephalopathy after transjugular intrahepatic portosystemic shunt. Clin Liver Dis 2012;16:133–146.
25. Cardenas A, Gines P, Marotta P, et al. Tolvaptan, an oral vasopressin antagonist, in the treatment of hyponatremia in cirrhosis. J Hepatol 2012;56:571–578.
26. Ahluwalia V, Wade JB, Thacker L, et al. Differential impact of hyponatremia and hepatic encephalopathy on health-related quality of life and brain metabolite abnormalities in cirrhosis. J Hepatol 2013;59:467–473.

Reprint requests
Address requests for reprints to: Jasmohan S. Bajaj, MD, Division of Gastroenterology, Hepatology and Nutrition, Virginia Commonwealth University and McGuire VA Medical Center, 1201 Broad Rock Boulevard, Richmond, Virginia 23249. e-mail: jsbajaj@vcu.edu; fax: (804) 675-5816.

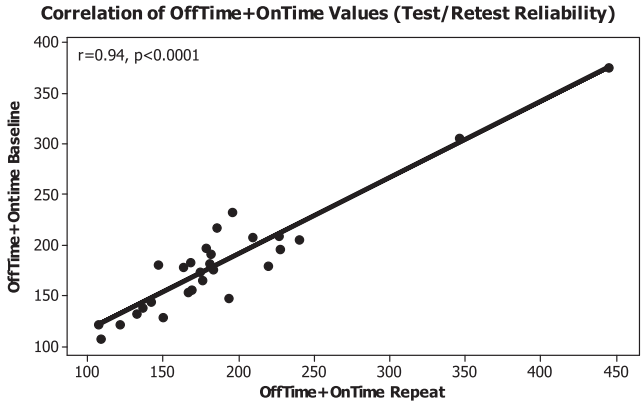
Conflicts of interest
The authors disclose no conflicts.

Funding
Supported by grant number R01AA020203 from NIAAA, R01DK087913 from NIDDK, and UL1TR00058 from the National Center for Advancing Translational Research and McGuire Research Institute. None of the grantors were involved in protocol design and implementation, data collection, analysis, or interpretation of the study results. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supplementary Table 1. Test-Retest Reliability

	Baseline (n = 30)	Repeat testing (n = 30)
MELD score	10.6 ± 4.5	10.5 ± 4.7
Standard tests		
NCT-A	38 ± 19	42 ± 34
NCT-B	101 ± 67	104 ± 101.5
Digit symbol	58 ± 19	60 ± 20
Block design	34 ± 14	36 ± 14
EncephalApp		
OffTime	81 ± 24	85 ± 31
OnTime	101 ± 32	104 ± 39
No. of runs for Off state, median (range)	5 (5–10)	5 (5–12)
No. of runs for On state, median (range)	5 (5–14)	6 (5–13)
OffTime+OnTime	182 ± 56	188 ± 69
OnTime–OffTime	20 ± 11	19 ± 13

NOTE. None of the comparisons were significantly different.

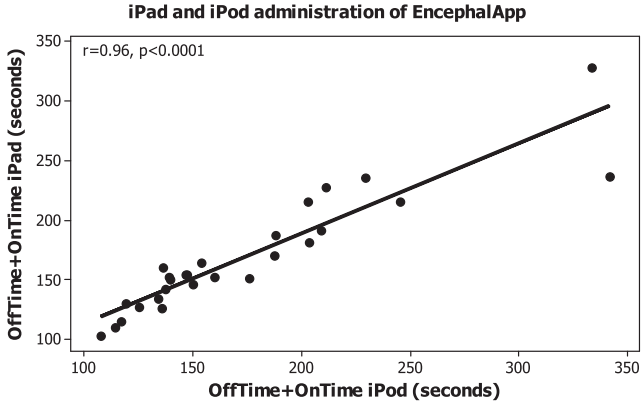


Supplementary Figure 1. Correlation of OffTime+OnTime results on repeat testing.

Supplementary Table 2. iPad Compared With iPod/iPhone

	iPad (n = 27)	iPod (n = 27)
OffTime	80 ± 24	79 ± 26
OnTime	90 ± 26	94 ± 35
No. of runs for Off state, median (range)	5 (5–12)	5 (5–8)
No. of runs for On state, median (range)	5.5 (5–11)	6 (5–14)
OffTime+OnTime	169 ± 49	173 ± 24

NOTE. None of the comparisons were significantly different.



Supplementary Figure 2. Correlation of iPod and iPad OffTime+OnTime results of EncephalApp.