Published in final edited form as:

Metab Brain Dis. 2019 February ; 34(1): 289–295. doi:10.1007/s11011-018-0350-z.

Diagnosis of covert hepatic encephalopathy: a multi-center study testing the utility of single versus combined testing

Andres Duarte-Rojo, MD, DSc[1], **Sanath Allampati, MD**[2], **Leroy R. Thacker, PhD**[3], **Christopher R. Flud, BS**[1], **Kavish R. Patidar, MD**[4], **Melanie B. White, RN**[4], **Jagpal S. Klair, MD**[1], **Douglas M. Heuman, MD**[4], **James B. Wade, PhD**[5], **Edith A. Gavis, RN**[4], and **Jasmohan S. Bajaj, MD, MS**[4,6]

[1]Division of Gastroenterology and Hepatology, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA.

[2]Internal Medicine and Gastroenterology, Cleveland Clinic, Cleveland, Ohio, USA.

[3]Family and Community Health Nursing and Biostatistics, Richmond, Virginia, USA

[4]Gastroenterology, Hepatology and Nutrition, Richmond, Virginia, USA

[5]Psychiatry, Virginia Commonwealth University, Richmond, Virginia, USA

[6]McGuire VA Medical Center, Richmond, Virginia, USA

Abstract

Background & Aims: Covert hepatic encephalopathy (CHE) affects cognition in a multidimensional fashion. Current guidelines recommend performing Psychometric Hepatic Encephalopathy Score (PHES) and a second test to diagnose CHE for multi-center trials. We aimed to determine if a two-test combination strategy improved CHE diagnosis agreement, and accuracy to predict overt hepatic encephalopathy (OHE), compared to single testing.

Methods: Cirrhotic outpatients without baseline OHE performed PHES, Inhibitory Control Test (ICT), and Stroop EncephAlapp (StE) at three centers. Patients were followed for OHE development. Areas under the receiver operation characteristic curve (AUROC) were calculated.

Results: We included 437 patients (399 with follow-up data). CHE prevalence varied with testing strategy: PHES+ICT 18%, ICT+StE 25%, PHES+StE 29%, ICT 35%, PHES 37%, and StE 54%. Combination with best test agreement was PHES+StE (k=0.34). Sixty patients (15%) developed OHE. Although CHE by StE showed the highest sensitivity to predict OHE, PHES and PHES+StE were more accurate at the expense of a lower sensitivity (55%, AUROC: 0.587; 36%, AUROC: 0.629; and 29%, AUROC: 0.623; respectively). PHES+ICT was the most specific (85%) but all strategies including ICT showed sensitivities in the 33–45% range. CHE diagnosis by PHES

Correspondence: Jasmohan S Bajaj, MD, MS, Division of Gastroenterology, Hepatology and Nutrition, Virginia Commonwealth University and McGuire VA Medical Center, 1201 Broad Rock Boulevard, Richmond, Virginia 23249, USA Telephone: (804) 675-5802, Fax: (804) 675-5816, jasmohan.bajaj@vcuhealth.org.

Disclosures: No potential conflict of interest.

Presentations: Portions of this manuscript were presented at The International Liver Congress 2015, European Association for the Study of the Liver; and The Liver Meeting 2016, American Association for the Study of Liver Diseases.

(HR=1.79, p=0.04), StE (HR=1.69, p=0.04), and PHES+StE (HR=1.72, p=0.04), were significant OHE predictors even when adjusted for prior OHE and MELD.

Conclusions: Our results demonstrate that combined testing decreases CHE prevalence without improving the accuracy of OHE prediction. Testing with PHES or StE alone, or a PHES+StE combination, is equivalent to diagnose CHE and predict OHE development in a multi-center setting.

Keywords

neuropsychological test; neurophysiological test; overt hepatic encephalopathy; PHES; inhibitory control test; stroop encephalapp

INTRODUCTION

Hepatic encephalopathy (HE) is a prevalent and disabling complication of liver failure and portal hypertension.[1] The clinical form of HE, overt HE (OHE) is particularly relevant to providers as it is associated with higher mortality. On a recent study the severest forms of OHE were associated with increased in-hospital and 30-day mortality, irrespective of the degree of liver failure or the presence of extrahepatic organ failure.[2] Moreover, OHE has been associated with hospital re-admissions, job termination, and it greatly affects the quality of life of caregivers, therefore creating a significant burden to health systems and society overall.[1, 3]

HE is now seen as a continuum from subtle cognitive abnormalities only diagnosed by means of specialized tests, also known as covert HE (CHE), to the clinically obvious OHE where cognition and arousal are affected to various degrees.[1] CHE is relevant to patients and its caregivers given its association with suboptimal work performance,[4] impaired learning capacity,[5] falls,[6] and motor vehicle accidents.[7, 8] Importantly, CHE is recognized as a factor for future development of OHE,[9, 10] and therefore it allows the early recognition of the preclinical phase heralding the advanced phases (i.e. OHE) associated with robust clinical outcomes.

The main difficulty in establishing risk stratification and opportune management to patients with CHE is the lack of a standardized diagnosis, therefore resulting in studies including a heterogeneous group of patients.[11, 12] Given that CHE is mostly diagnosed on the basis of a comparison of the cognitive capabilities to a control population, diagnostic testing and endpoints in clinical trials need to be individualized to a specific population, which may not be translatable across different sets of patients. Following the AASLD/EASL HE practice guideline recommendations,[13–16] our group recently published the US norms for three widely available psychometric tests used for the diagnosis of CHE: the psychometric hepatic encephalopathy score (PHES), inhibitory control test (ICT), and Stroop EncephalApp (StE).[10]

In an attempt to improve the accuracy of CHE diagnosis, current HE practice guideline has recommended using a consensus strategy for multi-center studies, where two psychometric and/or neurophysiological test(s), one of them being PHES, need to agree on the finding of

CHE.[17] However, there are a few concerns with such two-test CHE strategy: different tests evaluate different cognitive testing capabilities and therefore disagreement[18] would not necessarily rule out CHE diagnosis, also it is not known what combination of tests is the most appropriate. Finally, there is a gap in knowledge as to whether a combination of tests would improve stability of the CHE diagnosis and help in prediction of OHE development over a single test.

The aims of this study were to determine if a combination of two psychometric tests showed improved agreement among three different centers testing for CHE, and if a specific combination would better prognosticate the development of future OHE, when compared to single testing.

PATIENTS AND METHODS

Recruitment of patients has been previously described in detail.[10] In brief, subsequent patients with cirrhosis prospectively completed PHES, ICT, and StE at three participating centers (Virginia Commonwealth University and McGuire VA Medical Center [VA], Cleveland Clinic [OH], and University of Arkansas for Medical Sciences [AR]) between December 2012 and June 2014. Patients exhibiting clinical manifestations of OHE at the time of testing, having history of alcohol/illicit drug abuse in the 3 months prior to testing, on psychoactive medications (*e.g.* benzodiazepines and antipsychotics), prior transjugular intra-hepatic portosystemic shunting (TIPS) or those with red-green color blindness were excluded from study. The West Haven criteria were used to rule out OHE and all participants had an MMSE >25 at the time of inclusion. After 6 months from baseline CHE evaluation, each participant's medical record was reviewed to document any hospital admission for OHE. The protocols were approved by the IRBs of all the participating centers.

Statistical Analysis

We used previously validated norms to diagnose CHE by means of PHES, ICT, and StE.[10] CHE by PHES was defined as an impairment of at least −4 standard deviations when compared with expected results derived from controls, whereas in the case of both ICT and StE, CHE was determined by a score <0 after expected results were subtracted from observed results. In all instances, expected results were derived from formulas adjusted by age, gender, and education.

The outcome was development of OHE necessitating hospital admission. Incidental OHE was investigated in patients with or without CHE, the latter being defined by each competing psychometric test (*i.e.* PHES, ICT and StE) or any possible two-test combination. Standard statistics, including t-test, ANOVA, and chi-square, were used for group comparisons, as appropriate. In order to determine agreement between tests Cohen's kappa coefficient was calculated. The accuracy of each CHE testing strategy was assessed by calculating the corresponding area under the receiver operating characteristic curve (AUROC), and by Cox proportional hazards regression as part of a separate time-to-event analysis. AUROC were compared with Hanley-McNeil test. Analyses were performed using SAS (SAS Institute, Cary, NC) and Stata 12 (StataCorp, College Station, TX).

RESULTS

Table 1 shows the baseline characteristics of our cohort. Although there were no differences in age, patients included from Virginia were more frequently male, had a higher MELD score, and had a higher proportion with prior OHE at inclusion. Hepatitis C infection, as a cause of cirrhosis, was more frequently seen in patients from Arkansas. Use of anti-ammonia medications in patients with prior OHE also varied among centers.

Prevalence of CHE per testing strategy is shown in Table 2. StE was the test classifying the largest proportion of patients as having CHE (54%), whereas the combination of PHES + ICT was the one finding the lowest prevalence of CHE (18%). There was wide variation in the prevalence of CHE for each test or its combinations across the three participating centers (between centers variation as low as 5% for StE, and as high as 24% for ICT), what was not statistically different when comparing performance of single versus combination testing. As shown in Table 3, PHES and StE were the two tests showing the stronger agreement when classifying a patient as having CHE, either when considered isolated or in a consensus strategy and. Their combination seemed the most consistent when the underlying liver disease severity was taken into consideration.

Follow up data was available for 399 patients (mean follow up of 11±5 months); the rest were lost to follow-up. Sixty patients (15%) developed OHE during this period. Six patients either died or required a liver transplant, of which two had already developed OHE and were counted in the sixty patients above. Prevalence of CHE in this cohort did not differ from the one observed in all 437 patients. When considering single testing, StE was the most sensitive psychometric test, whereas PHES was the most specific (Table 4). Of the available two-test combinations, PHES + StE showed the best sensitivity *vs.* specificity tradeoff. However, when comparing the AUROC for each testing strategy, no statistically significant differences were observed across single or consensus strategies, in spite of the numerically higher AUROCs noted for PHES and PHES + StE. A sensitivity analysis considering CHE as the presence of at least one abnormal test out of each possible pair combination did not improve accuracy (Supplementary Table 1).

On a final time-to-event analysis adjusted by MELD and prior episode of OHE, we noted that PHES and StE, whether alone or in combination, were all acceptable strategies to predict incidental OHE - initial or recurrent episode (Table 5). Although PHES + ICT was statistically significant on this multivariable analysis, it lacked sensitivity to diagnose CHE and for OHE prediction.

DISCUSSION

Our results demonstrate that in a multi-center study specific test combinations yield a lower prevalence of CHE, without decreasing the variability in CHE prevalence among centers. This decrease in prevalence with consensus testing is a consequence of the poor to moderate agreement observed between individual tests. In this regard, PHES and StE was the test combination showing the best agreement. Remarkably, available test combinations were not statistically superior to individual testing to predict subsequent OHE, although the use of

either PHES alone or PHES + StE showed improved performance. On a time-to-event multivariable analysis, PHES, StE, and the combination of PHES + StE were all reasonable options to test for CHE and predict future OHE by showing similar performance.

While the clinical and psychosocial importance of CHE has been demonstrated by several studies, its diagnostic methods, especially across multiple sites, have not been standardized. The specific type of tests within each class and the diagnostic thresholds have remained an institutional choice and a matter of debate. Current HE practice guidelines by AASLD/EASL have attempted to systematize CHE diagnosis by proposing the use of PHES (paper-pencil psychometric test) plus a computerized psychometric test (such as ICT, StE or continuous reaction time [CRT]) or a neurophysiological test (such as critical flicker frequency [CFF] or electroencephalogram [EEG]) for multi-center experiences. Patients showing abnormal results on both the chosen tests (*e.g.* PHES + ICT or PHES + CFF) should be provided with a CHE diagnosis. However, being that CHE affects the brain in several dimensions with great heterogeneity in its manifestations, and since each of the proposed tests assesses a different cognitive or physiological skill set, it is not surprising to find the relatively poor concordance between any pair of chosen tests. Therefore, the consensus strategy, while intending to improve the specificity of CHE diagnosis, would suffer from lack of agreement between tests that might not concur under normal circumstances. Also, prospective data from multi-center studies showing how the consensus strategy would benefit CHE diagnosis has been lacking.

In the present study comparing the use of two tests, we found that single testing was able to identify the largest proportion of patients with CHE - whether using PHES, ICT, or StE. There was no specific pattern for paper-pencil, computer-based, or smartphone-based strategy identifying the largest prevalence of CHE across the three centers, suggesting that technology acceptance likely did not affect our results. Unexpectedly, we did not find a consistent attenuation in the CHE prevalence gradient across centers with the use of two tests, something we were somehow expecting, as this has been one of the reasons for the AASLD/EASL HE practice guideline providing such recommendation for multi-center studies. In other words, the differences in CHE prevalence were not reduced by the use of two versus one psychometric test in this multi-center study.

The present study extends prior reports that have cited a lack of agreement between psychometric tests into a multi-center realm within the United States.[19] Amodio *et al.* found a very poor correlation when raw results from PHES and ICT lures were compared to each other, and MHE diagnosis agreement with each test on 67% (κ=0.34, p<0.001) of studied cases in Italy.[18] With a different design, Gupta *et al.* found an agreement of 88% (κ=0.72, p<0.001) between PHES and ICT in an Indian cohort.[20] Of note, these two studies used different definitions for MHE both for PHES and ICT: cut-off value of ≤-4 (norm-based) and >24 weighted lures (a proposed ICT metric) in Italy, and ≤−5 (norm-based) and ≥14 lures in India, respectively. In our dataset, using all norm-based thresholds, agreement was found to be numerically lower at 64% (κ=0.22, p<0.001). Although there is no published precedent on the agreement between StE and other psychometric or neurophysiological tests, in our dataset, agreement with PHES (66%, κ=0.34, p<0.001) or ICT (60%, κ=0.22, p<0.001) was not better than between PHES and ICT, although the

higher kappa coefficient for the former comparison reflects improved certainty of the result. With these differences in mind, it is clear that at the cross-sectional level, the combinations derived from PHES, ICT, and StE provide operationally different information such that finding agreement in CHE diagnosis is not a very frequent event. Actually our results suggested that a consensus strategy would result in revising CHE diagnosis in up to 50% of patients initially identified as CHE on the basis of single testing. Despite the limitations of the relatively lower CHE diagnosis rate, the PHES+StE combination seemed the most consistent between centers when the underlying liver disease severity was taken into consideration.

Longitudinal validation of the preclinical phase within the spectrum of a disease is of the utmost importance. Ultimately, if the CHE diagnostic test results are not able to predict occurrence of OHE and other relevant clinical outcomes (*i.e.* motor vehicle accidents, hospitalization, or death), then early identification of the disease may lose its relevance from a clinical standpoint. This is a well-known concept in other subspecialties (metabolic syndrome in endocrinology or prehypertension in cardiology) allowing proper preventive or early intervention strategies, but most hepatology practices remain unconvinced. Multiple studies have shed prospective evidence on the usefulness of diagnosing CHE or MHE to predict future development of OHE, thus prospectively validating PHES, ICT, CRT, EEG, and CFF.[9, 19–23] However, there is a scarcity of evidence regarding the advantages of using a consensus over single-test strategy for the diagnosis of CHE and prediction of clinical outcomes. In a recent single-center Danish study including 129 patients with cirrhosis, PHES identified MHE in 34%, CRT in 53%, whereas consensus diagnosis was established only in 23%. There was no prognostic usefulness in consensus diagnosis for the prediction of OHE (26% *vs.* 23% for either PHES or CRT), although there was a gain in performance to predict mortality (40% *vs.* 21%, p=0.02).[23] Another single-center study from the UK evaluating 106 patients with cirrhosis by means of PHES, EEG, and CFF also found a drop in CHE prevalence when utilizing a consensus strategy (60% with PHES *vs.* 9% for PHES plus either EEG or CFF abnormal), replicating the lack of gain in prediction of OHE (17% for two out of three tests *vs.* 18% for PHES and 0% for EEG)[22] Our results, derived from a larger multi-center sample and including a test (StE) that has not been studied previously in combination with PHES, are in agreement with the lack of benefit of a consensus strategy over single testing in the prediction of future OHE (as per AUROC), in spite of a gain in specificity across the three test combinations. Moreover, we further evaluated incidental OHE by means of a time-to-event analysis, where single testing with PHES and StE remained as independent predictors, along with prior OHE, and no gain was noted as a result of combining PHES with StE.

Our study has a few limitations since we did include patients who had already developed OHE as a subset; however, this was adjusted for in the multi-variable analysis. We did not include any neurophysiological test (i.e. critical flicker frequency or electroencephalogram), which could have changed the interpretation of the data. Finally, we only followed our patients for less than 1 year post-enrollment and it is reasonable to expect further events should this period be extended. However, we did observe a large number of outcomes with adequate power, except for calculations including ICT.

In summary, our study showed that the use of a consensus strategy to diagnose CHE, as proposed by current practice guidelines, did not stabilize performance of PHES, ICT, and StE, and it decreased CHE prevalence without a significant gain in OHE prognostication. Without discouraging the use of a consensus strategy in future multi-center clinical trials, particularly until other research groups replicate our results in a multi-center setting, we believe that single testing with PHES and StE are sufficient to diagnose CHE in clinical practice, with the latter yielding more false positives (*i.e.* CHE not predicting OHE). However, if a consensus strategy is planned, the combination of these two tests is recommended, what could be used in a sequential manner: screening with StE followed by confirmatory PHES.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant Support: ADR receives partial support from the University of Arkansas for Medical Sciences College of Medicine Clinician Scientist Program. This work was also partly supported by NIH RO1DK089713 and VA Merit Review CX1076 to JSB.

Abbreviations:

AASLD:	American Association for the Study of Liver Disease
ANOVA:	analysis of variance
AR:	Arkansas
AUROC:	area under the receiving operating characteristic
CFF:	critical flicker frequency
CRT:	continuous reaction time
CHE:	convert hepatic encephalopathy
EASL:	European Association for the Study of the Liver
EEG:	electroencephalogram
HE:	hepatic encephalopathy
HCV:	hepatitis C virus
HR:	hazards ratio
ICT:	inhibitory control test
IRB:	Institutions Review Board
OH:	Ohio

OHE:	overt hepatic encephalopathy
MELD:	model for end-stage liver disease
MMSE:	mini-mental state examination
NASH:	non-alcoholic liver disease
PHES:	Psychometric Hepatic Encephalopathy Score
StE:	Stroop EncephalApp
US:	United States
VA:	Virginia

REFERENCES

1. Patidar KR, Bajaj JS. Covert and Overt Hepatic Encephalopathy: Diagnosis and Management. Clin Gastroenterol Hepatol 2015;13:2048–2061. [PubMed: 26164219]
2. Bajaj JS, O'Leary JG, Tandon P, et al. Hepatic Encephalopathy Is Associated With Mortality in Patients With Cirrhosis Independent of Other Extrahepatic Organ Failures. Clin Gastroenterol Hepatol 2016.
3. Volk ML, Tocco RS, Bazick J, et al. Hospital readmissions among patients with decompensated cirrhosis. Am J Gastroenterol 2012;107:247–252. [PubMed: 21931378]
4. Weissenborn K The Clinical Relevance of Minimal Hepatic Encephalopathy--A Critical Look. Dig Dis 2015;33:555–561. [PubMed: 26159273]
5. Ortiz M, Cordoba J, Jacas C, et al. Neuropsychological abnormalities in cirrhosis include learning impairment. J Hepatol 2006;44:104–110. [PubMed: 16169117]
6. Soriano G, Roman E, Cordoba J, et al. Cognitive dysfunction in cirrhosis is associated with falls: a prospective study. Hepatology 2012;55:1922–1930. [PubMed: 22213000]
7. 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;50:1175–1183. [PubMed: 19670416]
8. Wein C, Koch H, Popp B, et al. Minimal hepatic encephalopathy impairs fitness to drive. Hepatology 2004;39:739–745. [PubMed: 14999692]
9. Riggio O, Amodio P, Farcomeni A, et al. A Model for Predicting Development of Overt Hepatic Encephalopathy in Patients With Cirrhosis. Clin Gastroenterol Hepatol 2015;13:1346–1352. [PubMed: 25572976]
10. 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. [PubMed: 26644276]
11. 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. [PubMed: 17326150]
12. Rai R, Ahuja CK, Agrawal S, et al. Reversal of Low-Grade Cerebral Edema After Lactulose/Rifaximin Therapy in Patients with Cirrhosis and Minimal Hepatic Encephalopathy. Clin Transl Gastroenterol 2015;6:e111. [PubMed: 26378384]
13. Duarte-Rojo A, Estradas J, Hernandez-Ramos R, et al. Validation of the psychometric hepatic encephalopathy score (PHES) for identifying patients with minimal hepatic encephalopathy. Dig Dis Sci 2011;56:3014–3023. [PubMed: 21461913]
14. Amodio P, Campagna F, Olianas S, et al. Detection of minimal hepatic encephalopathy: normalization and optimization of the Psychometric Hepatic Encephalopathy Score. A neuropsychological and quantified EEG study. J Hepatol 2008;49:346–353. [PubMed: 18602716]

15. Weissenborn K, Ennen JC, Schomerus H, et al. Neuropsychological characterization of hepatic encephalopathy. J Hepatol 2001;34:768–773. [PubMed: 11434627]
16. Romero Gomez M, Cordoba J, Jover R, et al. [Normality tables in the Spanish population for psychometric tests used in the diagnosis of minimal hepatic encephalopathy]. Med Clin (Barc) 2006;127:246–249. [PubMed: 16942726]
17. 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;60:715–735. [PubMed: 25042402]
18. Amodio P, Ridola L, Schiff S, et al. Improving the inhibitory control task to detect minimal hepatic encephalopathy. Gastroenterology 2010;139:510–518, 518 e511–512. [PubMed: 20470775]
19. Montagnese S, Balistreri E, Schiff S, et al. Covert hepatic encephalopathy: agreement and predictive validity of different indices. World J Gastroenterol 2014;20:15756–15762. [PubMed: 25400460]
20. Gupta D, Ingle M, Shah K, et al. Prospective comparative study of inhibitory control test and psychometric hepatic encephalopathy score for diagnosis and prognosis of minimal hepatic encephalopathy in cirrhotic patients in the Indian subcontinent. J Dig Dis 2015;16:400–407. [PubMed: 25858627]
21. 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;149:1483–1489. [PubMed: 26299413]
22. Thomsen KL, Macnaughtan J, Tritto G, et al. Clinical and Pathophysiological Characteristics of Cirrhotic Patients with Grade 1 and Minimal Hepatic Encephalopathy. PLoS One 2016;11:e0146076. [PubMed: 26745876]
23. Lauridsen MM, Schaffalitzky de Muckadell OB, Vilstrup H. Minimal hepatic encephalopathy characterized by parallel use of the continuous reaction time and portosystemic encephalopathy tests. Metab Brain Dis 2015;30:1187–1192. [PubMed: 26016624]

Key Points Box:

- When testing for covert hepatic encephalopathy (HE) in multi-center studies, it is recommended to use consensus diagnosis based on two psychometric tests.
- Prevalence of covert HE shows wide variation depending on testing strategy.
- Use of two-test consensus strategy does not outperform single testing when prediction of future overt HE is used as the endpoint.
- In this multi-center study, single testing with Stroop EncephalApp or PHES was the most efficient way to identify and prognosticate covert HE.

Table 1.

Baseline characteristics of patients

	All centers (n=437)	Virginia (n=230)	Ohio (n=107)	Arkansas (n=100)	*p**
Age	57 ± 8	57 ± 7	58 ± 11	56 ± 7	0.4
Sex (male)	64%	74%	59%	46%	<0.001
Etiology					
HCV	54%	50%	30%	87%	<0.001
Alcohol	15%	19%	20%	2%	
NASH	19%	23%	26%	4%	
Other	12%	9%	24%	7%	
MELD	11 ± 5	13 ± 6	10 ± 4	10 ± 3	<0.0001
Prior OHE	36%	41%	30%	30%	0.05
OHE Treatment					
Lactulose	69%	80%	26%	59%	
Rifaximin	7%	5%	11%	9%	<0.001
Both	24%	15%	63%	32%	

*Refers to comparison across patients included in Virginia, Ohio, and Arkansas. HCV: hepatitis C virus, MELD: model for end-stage liver disease, NASH: non-alcoholic steatohepatitis, OHE: overt hepatic encephalopathy

Table 2.

Prevalence of CHE according to testing strategy

	All Centers (n=437)	Virginia (n=230)	Ohio (n=107)	Arkansas (n=100)	*p**
PHES	37%	45%	30%	25%	<0.001
ICT	35%	25%	39%	49%	<0.001
StE	54%	56%	51%	55%	0.7
PHES + ICT	18%	18%	14%	20%	0.4
PHES + StE	29%	37%	25%	15%	<0.001
ICT + StE	25%	20%	25%	34%	0.03

*Refers to comparison across patients included in Virginia, Ohio, and Arkansas. CHE: covert hepatic encephalopathy, ICT: inhibitory control test, PHES: psychometric hepatic encephalopathy score, StE: Stroop EncephalApp

Table 3.

Agreement in CHE diagnosis among centers by abnormal test(s) results

	PHES	ICT	StE	PHES+StE	PHES+ICT	ICT+StE
PHES		0.22	0.34	0.83	0.54	0.27
ICT			0.22	0.22	0.57	0.77
StE				0.50	0.19	0.42
PHES+StE					0.54	0.39
PHES+ICT						0.62
ICT+StE						

Result is expressed as kappa coefficient. Grayed cells represent the correlation between each pair of tests and its components. ICT: inhibitory control test, PHES: psychometric hepatic encephalopathy score, StE: Stroop EncephalApp

Table 4.

Accuracy of CHE diagnosis as a predictor of OHE according to the use of single or two-test strategy.

Test	CHE Prevalence (n=399)	Incident OHE	AUROC (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)
PHES	36%	24% (35/145)	0.629 (0.562–0.697)	58 (45–71)	68 (62–73)
ICT	35%	20% (27/138)	0.561 (0.493–0.630)	45 (32–58)	67 (62–72)
StE	55%	19% (42/220)	0.587 (0.523–0.652)	70 (57–81)	48 (42–53)
PHES+ICT	18%	29% (20/70)	0.593 (0.530–0.656)	33 (22–47)	85 (81–89)
PHES+StE	29%	26% (30/115)	0.623 (0.557–0.693)	50 (37–63)	75 (70–80)
ICT + StE	25%	21% (21/99)	0.560 (0.495–0.624)	35 (23–48)	77 (72–81)

All possible comparisons between AUROC were non-significant.

95% CI: confidence interval, AUROC: area under the receiving operating characteristic curve, CHE: covert hepatic encephalopathy, ICT: inhibitory control test, OHE: overt hepatic encephalopathy, PHES: psychometric hepatic encephalopathy score, StE: Stroop EncephalApp.

Table 5.

Risk for incidental OHE in patients with baseline CHE according to testing strategy.

Single-test strategy						
	PHES HR (95%CI)	***p***	**ICT HR (95%CI)**	***p***	**StE HR (95%CI)**	***p***
CHE	1.79 (1.02, 3.13)	0.04	1.54 (0.90, 2.63)	0.11	1.69 (1.01–3.07)	0.04
Prior OHE	2.06 (1.17, 3.63)	0.01	2.11 (1.18, 3.75)	0.01	2.04 (1.16, 3.60)	0.01
MELD	1.03 (0.99, 1.08)	0.05	1.05 (1.01, 1.09)	0.01	1.04 (1.00, 1.08)	0.14
Two-test strategy						
	PHES + ICT HR (95%CI)	***p***	**PHES + StE HR (95%CI)**	***p***	**ICT + StE HR (95%CI)**	***p***
CHE	2.02 (1.13, 3.59)	0.01	1.72 (1.02, 3.03)	0.04	1.48 (0.85, 2.59)	0.16
Prior OHE	2.07 (1.16, 3.69)	0.01	2.04 (1.16, 3.61)	0.01	2.11 (1.18, 3.76)	0.01
MELD	1.04 (1.00, 1.08)	0.05	1.03 (0.99, 1.08)	0.15	1.05 (1.01, 1.09)	0.01

CHE: covert hepatic encephalopathy, HR (95%CI): hazards ratio and confidence interval, ICT: inhibitory control test, MELD: model for end-stage liver disease, OHE: overt hepatic encephalopathy, PHES: psychometric hepatic encephalopathy score, StE: Stroop EncephalApp