




ORIGINAL ARTICLE OPEN ACCESS

Minimal Hepatic Encephalopathy in Cirrhotic Patients: A New Simple and Fast Digital Screening Method

Henrike Dobbermann¹ | Raffael Schlüter¹ | Yevgeniy Lyubchenko² | Johanna Beder¹ | Sven Danneberg¹ | Katharina Mitzlaff¹ | Iris Engelbart¹ | Jakob Jessberger¹ | Felix Braun³ | Thomas Becker³ | Christian Labenz^{4,5}  | Monique Janneck² | Denys Matthies² | Friedhelm Sayk¹  | Jens U. Marquardt¹ 

¹Department of Medicine I, University Medical Center Schleswig-Holstein, Lübeck, Germany | ²Department of Electrical Engineering and Computer Science, Technische Hochschule Lübeck, University of Applied Sciences, Lübeck, Germany | ³Department of General, Visceral-, Thoracic-, Transplantation- and Pediatric Surgery, University Medical Centre Schleswig-Holstein (UKSH), Kiel, Germany | ⁴Department of Internal Medicine I, University Medical Center of the Johannes Gutenberg-University, Mainz, Germany | ⁵Cirrhosis Center Mainz (CCM), University Medical Center of the Johannes Gutenberg-University, Mainz, Germany

Correspondence: Jens U. Marquardt (jens.marquardt@uksh.de)

Received: 5 October 2024 | **Revised:** 18 December 2024 | **Accepted:** 31 December 2024

Funding: Publication fees funded by Fa. Merz Therapeutics GmbH, Frankfurt, Germany.

Keywords: app | CHILD | cirrhosis | cognitive deterioration | diagnosis | digital | MELD | psychometric hepatic encephalopathy score | smartphone | Stroop test

ABSTRACT

Background: Minimal hepatic encephalopathy is a common and prognostically severe complication of cirrhosis with a significant impact on the quality of life. Detailed diagnostic work-up of minimal hepatic encephalopathy is time-consuming and difficult to integrate into daily clinical routine.

Objective: We aimed to develop a new, simple, and easy-applicable smartphone-based self-assessment method for screening of minimal hepatic encephalopathy.

Methods: 92 patients with cirrhosis and 20 healthy controls were recruited to perform 3 different short digital tests on smartphones (Tip test (TT), number connection test (dNCT) and modified Stroop test (ST)). Results were correlated with the Psychometric Hepatic Encephalopathy Score (PHES) as the presumed gold standard for minimal hepatic encephalopathy. The impact of age, gender, education, CHILD and MELD scores was further investigated.

Results: All 3 digital tests showed good correlation with PHES (TT $r = -0.76$, dNCT $r = -0.58$, and ST $r = -0.65$; all $p < 0.001$). Digital tests were performed significantly faster (TT median 41s (IQR 36–51s); dNCT median 21s (IQR 8–16s); ST median 76s (IQR 55–99s) than PHES (median 322s; IQR 261–434s); There were significant differences between age groups and different levels of education ($p < 0.05$). AUC for TT was 0.835 (95% confidence interval 0.747–0.922, $p < 0.001$) and highest among all digital tests.

Conclusion: All 3 digital tests proved to be suitable for screening of minimal hepatic encephalopathy. TT showed the highest correlation with reference PHES and was not affected by language skills or color blindness, and, thus, might represent a new and fast method for minimal hepatic encephalopathy detection. Intra-individually adjusted smartphone-based thresholds might further eliminate the influence of age, gender, educational level or training, to refine early app-based alerts in case of cognitive deterioration in cirrhotic patients.

Henrike Dobbermann and Raffael Schlüter are the first authors shared first author.

Friedhelm Sayk and Jens U. Marquardt shared last author.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *United European Gastroenterology Journal* published by Wiley Periodicals LLC on behalf of United European Gastroenterology.

Summary

- Established knowledge on the subject
 - Minimal hepatic encephalopathy (mHE) is cumbersome to diagnose and often neglected in clinical routine.
 - Established psychometric assessments like Psychometric Hepatic Encephalopathy Score (PHES) paper-and-pencil test are time-consuming, resource-demanding and not suitable for self-assessment.
 - Integration of new, smartphone-based digital methods promises autonomous self-conducted surveillance with potentially earlier mHE detection.
- Significant and new findings of this study
 - In the present head-to-head comparison, three novel smartphone-based digital tests proved to be rapid and reliable approaches to detect mHE as correlated with concurrent reference PHES.
 - Tip Test showed the best performance. Its pragmatic ease of use, speed, and discriminatory capacity could make it valuable in clinical settings inside and outside the hospital.
 - Potential learning effects at regularly repeated self-screening and constitutional individual factors warrant adjustment. Predetermined individual baseline and threshold levels based on a smartphone-app could generate a continuous degradation alert in the event that mHE develops or deteriorates.

1 | Introduction

Hepatic encephalopathy (HE) is a common and severe complication of cirrhosis, which can significantly affect patients' prognosis [1, 2]. This also applies to minimal hepatic encephalopathy (mHE), the subclinical form of HE [3–5]. mHE is present in 30%–70% of cirrhosis patients and can significantly impair memory, attention, and reaction time, thereby severely affecting the quality of life [6]. Individuals with mHE typically exhibit prolonged reaction times, which can compromise activities like driving [7, 8]. Furthermore, mHE significantly increases the risk of progression to overt HE (oHE) leading to hospitalizations and higher socioeconomic costs [9–11]. Diagnosis of mHE presents challenges in clinical practice. The most validated method is the internationally recognized Psychometric Hepatic Encephalopathy Score (PHES), which includes five manual paper-and-pencil tests. To perform the test battery usually takes 10–20 min, with an additional 3–10 min for evaluation. Since PHES requires clinical staff and is not suitable for self-conduct by patients, comprehensive testing of mHE is often neglected in daily clinical routine due to limited resources [12]. This hampers the early detection and treatment of HE [13]. Recently, digital diagnostic tools, such as the critical flicker frequency (CFF) and animal naming test (ANT), have been developed and validated for faster and simpler HE diagnosis [2, 5, 14].

New smartphone-based methods such as the EncephalApp Stroop Test offer significant advantages compared to conventional methods. This app uses the Stroop effect, a cognitive phenomenon related to inhibitory control [15] where patients have to discriminate between a denominated color and the apparent font

color (e.g., the word “blue” displayed in red letters). This multilingual application has been internationally validated as a reliable stand-alone tool for regular follow-up testing [16–18]. Its smartphone compatibility makes it a pragmatic option for detecting mHE in clinical practice, with the potential for wider accessibility in the future. However, the app has some limitations. Results can be severely affected by factors such as age or education level, and it is not suitable for patients who are illiterate or have red-green color blindness. Additionally, the test can still be time-consuming and mentally demanding, which might make it frustrating for patients and limit its use as a self-performed screening tool at home.

Thus, developing a novel, easy-applicable and fast digital test approach independent of language skills and color vision is highly desirable. Our aim was to develop a practical tool for patient-based repeated mHE screening at home. As a first step we here report the accuracy of novel digital tests compared to the standard PHES.

2 | Methods

2.1 | Study Cohort

Ninety-two patients with cirrhosis, who were familiar with using a smartphone, were recruited from the hepatology division of our medical department. The diagnosis of ~~liver~~ cirrhosis was established by conventional definition through a combination of imaging techniques (CT, MRI, sonography including transient elastography), laboratory findings (thrombocytopenia and diminished liver synthesis parameters), clinical signs and endoscopic indicators (presence of esophageal varices and signs of portal hypertensive gastropathy) or liver biopsy. To stage the severity of liver cirrhosis, Pugh score and Model of end-stage liver disease (MELD) were calculated. We excluded patients who either did not provide consent, showed cognitive impairments indicative of a manifest neurological disorder or overt HE (≥ 2 , West Haven classification), or had taken sedative medications (e.g. benzodiazepines or narcotics) or alcohol on the day of testing. Recruitment was performed during regular outpatient visits. A subset of the study group received inpatient care at our university hospital due to cirrhosis-related complications other than overt HE. 20 healthy control subjects without any history of liver disease were recruited from the municipal normal population.

2.2 | Testing for Hepatic Encephalopathy

All patients underwent a clinical examination by a hepatologist to rule out overt hepatic encephalopathy, acute intoxication, or acute withdrawal symptoms prior to inclusion. This was complemented by laboratory tests, to control for confounding metabolic disturbances. Subsequently, patients and controls performed the PSE test battery, developed by Weissenborn et al. (version 2019) to obtain the individual PHES [12, 19]. Test results were compared with PSE-inherent reference values tailored to age, gender, and years of school education. A score lower than -4 is considered pathological, indicating mHE [12].

Following PHES assessment, participants underwent three smartphone-based tests: the Tip Test (TT), digital Number Connection Test (dNCT), and the modified Stroop test (Stroop test).

1. *Tip Test*: On the smartphone display, a grid pattern of 70 tiles was presented. Next to 69 gray tiles, one tile was distinctively colored in blue. Once the participant tapped this blue tile, its position in the grid changed. Participants were requested to tap this alternating blue tile as fast as possible, the whole task consisted of 40 taps, and the time needed for the whole run was recorded.
2. *dNCT*: Inspired by the Number Connection Test A from the PHES, the dNCT was adapted for the smartphone display. Due to the smaller screen size, it showed only 9 numbers scattered randomly across the screen. Participants had to connect these numbers in ascending order by “swiping” their finger, thus also including a motoric component akin to the line drawing test of PHES. The time needed to complete the task and finger movements were recorded.
3. *Modified Stroop test*: The test described hereafter examines inhibitory control similar to the EncephalApp by Bajaj et al. In an attempt to create a faster, but still cognitively challenging test, we chose to limit the total number of iterations (30 items compared to at least 140 items for “On” and “Off” state of EncephalApp) and, to retain complexity, included five different colors instead of the original three. Patients had to identify the font color of a word and select the corresponding color out of one of five buttons displayed at the screen's bottom (white font on a blue background) (See Figure S1). However, the word itself denominated a different color (e.g. the word “green” displayed in blue font, thus, the “blue” button should be selected). Both the time required and the number of incorrect selections was recorded.

Screenshots of the three digital tests described above are depicted in Figure S1.

2.3 | Statistical Data Analysis

The results from the smartphone-based tests for cirrhosis patients and controls were correlated with the corresponding individual total PHES results. Additionally, the results were correlated with the subject's age, educational level, as well as MELD and Child Score. Moreover, the results of digital tests were sub-analyzed for effects of gender, age, and educational level.

To determine the normal distribution of metric data from the smartphone-based tests, the Shapiro-Wilk test was performed. In cases where data were not normally distributed, Spearman's rank correlation method was applied for correlation analyses and data were displayed as the median and inter-quartile range (IQR). Statistical significance of results across various groups was analyzed using the-U test. All tests were two-tailed, and a p -value ≤ 0.05 was considered to be statistically significant.

To examine the diagnostic accuracy of the digital TT, dNCT and Stroop test, a ROC (Receiver Operating Characteristic) curve analysis was conducted. Cut-off values were determined by calculating Youden's Index (J) as well as the point closest top left (C) by using the formulas listed below and choosing the highest and lowest values as a suitable cut-off value.

$$J = \text{sensitivity} + \text{specificity} - 1$$

$$C = \sqrt{(1 - \text{sensitivity})^2 + (1 - \text{specificity})^2}$$

All statistical evaluations were performed using SPSS software, Version 28.0 (SPSS IBM Corp., New York, USA).

2.4 | Ethics

This study was approved by the ethics committee of the University of Lübeck (file number 2022–556). Written informed consent was obtained from all participants prior to testing.

3 | Results

3.1 | Baseline Characteristics

A total of 100 patients with cirrhosis were screened for eligibility. Eight patients had to be excluded due to missing smartphone experience ($n = 5$), intake of sedative medication ($n = 2$) or underlying neurological conditions ($n = 1$) as predefined. Finally, 92 patients with cirrhosis were enrolled comprising 35 females (38%) and 57 males (62%) with an average age of 60 ± 9 years (range: 38–87 years). 67 patients were recruited from the outpatient department and 25 were inpatients. The majority of the patients were classified as compensated cirrhosis Child stage A ($n = 55$; 59.8%). The flow chart of eligibility and enrollment for mHE testing is displayed in Figure 1. As a control cohort, 20 healthy individuals (9 males) with a mean age 56 ± 16 years (range: 26–76) were additionally recruited. 20 patients (21.7%) consumed alcohol on a regular basis, 9 (9.8) within the last week, but not within 24 h prior to mHE testing, and two patients used other drugs (cannabis $n = 1$, benzodiazepine $n = 1$), occasionally, but not within the last 48 h. Baseline characteristics of patients and controls are detailed in Table 1.

3.2 | Test Results of PHES and Smartphone-Based Tests

In order to document present individual cerebral performance, all participants underwent the standardized and validated PSE-test to obtain the PHES, which comprises five subtests and is the conventionally accepted gold standard for mHE detection [19]. One patient declined to perform the “line tracing” subtest and was excluded from further analysis.

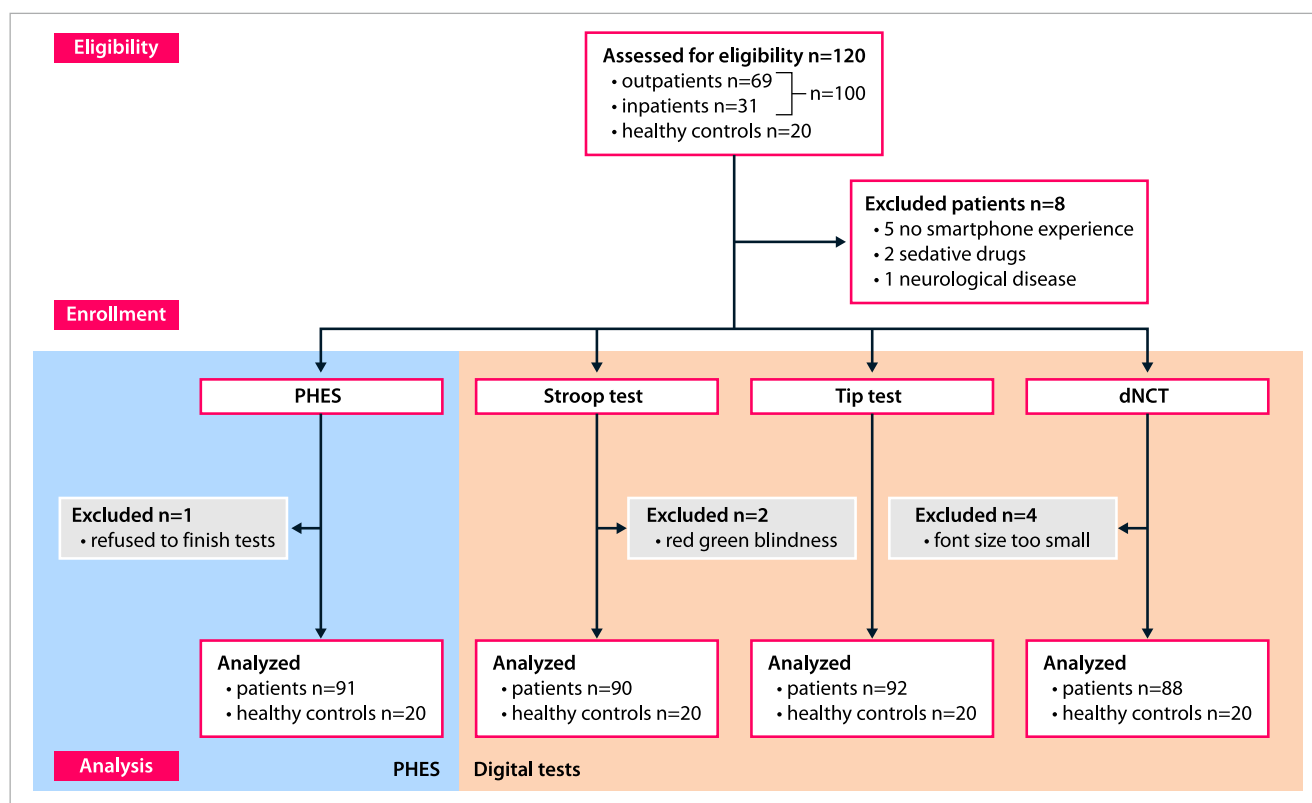


FIGURE 1 | Flow diagram of eligible and enrolled or excluded participants (cirrhotic patients and healthy controls) for PHES (left side, blue) and digital mHE-tests (right side, orange). dNCT: digital number connecting test; PHES: psychometric hepatic encephalopathy score.

According to the PHES, mHE was confirmed in 60 of the 92 patients (65%) and was ruled-out in all healthy controls as expected. To complete the PHES assessment (combined durations of the five subtests, excluding the time taken for explanations and evaluation of test results), patients required a median time of 322 s (IQR 261–434 s) and healthy individuals about 314 s (median, IQR 302–361 s).

Subsequently, participants performed the three digital tests. All 92 participants of the patient cohort completed the TT. Two of the 92 subjects were excluded from the Stroop test due to red-green color blindness. Moreover, the results of the first four participants for the dNCT had to be excluded, because the initially chosen font size of 30 pt proved too small for appropriate smartphone screen visibility. After adjusting the font size to 50 pt, the test was well applicable to the remaining 88 participants. Missing cases were excluded from the respective correlation analyses (Figure 1). All healthy control subjects completed the digital tests.

To perform the digital TT, cirrhotic patients required a median time of 41 s (IQR 36–51 s), while healthy controls needed about 31 s (median, IQR 28–31 s). The dNCT patients required a median of 21 s (IQR 8–16 s) and healthy individuals 7 s (median, IQR 5–9 s). In the patient cohort, the median of time required to complete the modified Stroop test was 76 s (IQR 55–99 s) and for control subjects 46 s (IQR 40–49 s). The respective boxplots of results are shown in Figure 2.

3.3 | Differences Between Patients With and Without mHE

As expected, there were marked differences in digital test results between the patient subgroup diagnosed with mHE by PHES and those without PHES (all $p < 0.001$); results are detailed in Table 2 and depicted in Figure 3a. Patients with mHE had a rate of about three errors (median, IQR 1–8). In contrast, patients without mHE only made about one mistake (median, IQR 0–2.5). Notably, since each error has to be corrected instantaneously by the subject, higher error rate results in a longer duration to complete the test. Therefore, test duration is the primary read-out combining the speed of performance and error rate.

3.4 | Correlation Between PHES and Digital Test Results

The digital test results were thoroughly evaluated by within-subject comparison with the composite total PHES. In cirrhotic patients, all digital tests demonstrated statistically significant ($p < 0.001$) correlations with the PHES. Notably, in this cohort, the TT showed the strongest correlation ($r = -0.76$), followed by the Stroop test ($r = -0.67$) and the dNCT ($r = -0.59$). In healthy controls, in contrast, no significant correlations were observed between the cumulative PHES and the digital tests (Figure 4 a and b).

TABLE 1 | Baseline characteristics of the total cohort of patients with cirrhosis and subgroups with and without minimal hepatic encephalopathy (mHE) according to PHES or of healthy controls.

Variable	Total number of patients <i>n</i> = 92	Without mHE <i>n</i> = 31 (34%)	mHE <i>n</i> = 60 (66%)	Healthy controls
Male <i>n</i> (%)	57 (62)	17 (55)	40 (67)	9 (45)
Female <i>n</i> (%)	35 (38)	14 (45)	20 (33)	11 (55)
Age in years, mean \pm SD	60 \pm 9	57 \pm 9	62 \pm 10	56 \pm 16
Education in years, mean \pm SD	9.8 \pm 1.4	10.2 \pm 1.6	9.8 \pm 1.3	12.2 \pm 1.4
Outpatient number <i>n</i> (%)	67 (73)	27 (87)	39 (65)	
Inpatient number <i>n</i> (%)	25 (27)	4 (13)	21 (35)	
Etiology of cirrhosis <i>n</i> (%)				
Alcohol	57 (62)	22 (71.0)	35 (58.3)	
Alcohol + hepatitis B/C	6 (6.6)	1 (3.2)	5 (8.3)	
Metabolic	13 (14.1)	2 (6.5)	11 (18.4)	
Viral hepatitis (B + C)	3 (3.3)	1 (3.2)	1 (1.7)	
Autoimmune (PBC/AIH)	6 (6.5)	2 (6.5)	4 (6.7)	
Other	7 (7.6)	3 (9.7)	4 (6.7)	
Child pugh score, mean \pm SD	6 \pm 1.6	6 \pm 1	7 \pm 2	a
A <i>n</i> (%)	55 (59.8)	27 (87)	27 (45)	
B <i>n</i> (%)	34 (37)	4 (13)	30 (50)	
C <i>n</i> (%)	3 (3.3)	0 (0)	3 (5.1)	
MELD score, mean \pm SD	11 \pm 5	9 \pm 4.6	12 \pm 5.1	a
MELD Na score, mean \pm SD	13 \pm 6	10 \pm 4.8	14 \pm 1.7	a
Laboratory values mean \pm SD				
Bilirubin μ mol/L	23.3 \pm 31.7	23.5 \pm 46.4	23.4 \pm 21.2	
Creatinine μ mol/L	103.7 \pm 71.3	81.7 \pm 25.7	116 \pm 84	
Albumin mg/L	36.5 \pm 7.3	41.1 \pm 6.2	34 \pm 6.5	
INR	1.2 \pm 0.3	1.2 \pm 0.4	1.3 \pm 0.3	
Thrombocytes/nL	159.9 \pm 84.8	160.4 \pm 78	157.4 \pm 87.7	
Na mmol/L	137.2 \pm 4.8	138.9 \pm 2.9	137 \pm 5.3	
Extent of ascites <i>n</i> (%)				
None	63 (68.5)	27 (87)	35 (58.3)	
Mild	13 (14.1)	3 (9.7)	10 (16.7)	
Moderate to severe	16 (17.4)	1 (3.2)	15 (25)	
Prior decompensations <i>n</i> (%)				
Ascites	59 (64.1)	19 (61.3)	40 (66.7)	
SBP	14 (15.2)	1 (3.2)	13 (21.7)	
Eos. varices	55 (59.8)	20 (64.5)	35 (58.3)	
Variceal bleeding	14 (15.2)	3 (9.7)	11 (18.3)	
HRS	12 (13)	3 (9.7)	9 (15)	
oHE	9 (9.8)	1 (3.2)	8 (13.3)	
cHE	11 (12)	3 (9.7)	8 (13.3)	
Alcohol consumption in the last 3 months <i>n</i> (%)	20 (21.7)	6 (19.4)	13 (21.7)	
Alcohol consumption in the last 7 days <i>n</i> (%)	9 (9.8)	4 (12.9)	5 (8.3)	
Consumption of > 24 g (men) or 12 g (women) alcohol per day <i>n</i> (%)	4 (4.5)	2 (6.5)	2 (3.3)	

(Continues)

TABLE 1 | (Continued)

Variable	Total number of patients <i>n</i> = 92	Without mHE <i>n</i> = 31 (34%)	mHE <i>n</i> = 60 (66%)	Healthy controls
Reason for hospitalization <i>n</i> (% of inpatients)				
Hydrope decomp.	11 (44)	1 (25)	10 (47.6)	
Icteric decomp.	1 (4)	1 (25)	0 (0)	
SBP	1 (4)	0 (0)	1 (4.8)	
Varic. ligation	2 (8)	1 (25)	1 (4.8)	
Other infections	4(16)	0 (0)	4 (19.1)	
Anemia	2 (8)	1 (25)	1 (4.8)	
Pulmonary embolism	1 (4)	0 (0)	1 (4.8)	
Unspecific	3 (12)	0 (0)	3 (14.3)	

Note: Unspecific: contains stomach pain, exsiccosis.

Abbreviations: decomp., decompensation; Eos., esophageal; SBP, spontaneous bacterial peritonitis; varic., variceal.

^a*p* < 0.001 for comparison of patient subgroups with versus without mHE.

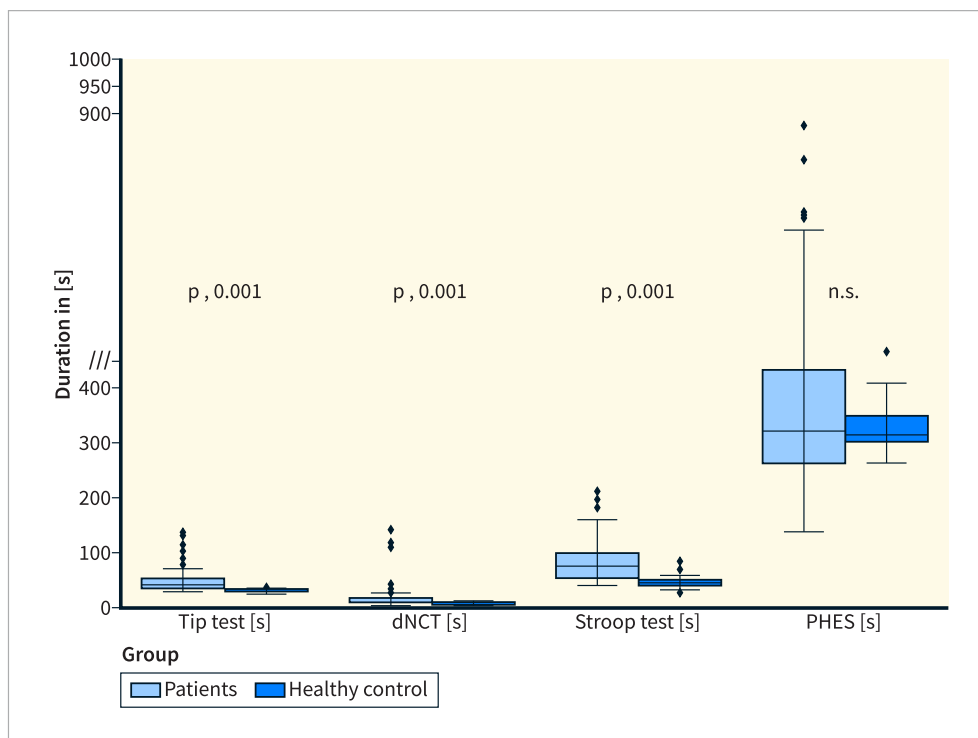


FIGURE 2 | Box-plots of time (seconds) required to perform each of the three smart-phone based mHE tests in comparison to the time needed for PHES testing (time required for explanation and evaluation of PHES not included) in cirrhotic patients versus healthy controls.

3.5 | Correlation Between Digital Test Results and Clinical Characteristics

When correlating MELD and Child Score with the PHES and digital test results, positive coefficients with high statistical significance were found for the PHES ($p < 0.001$) and for the TT ($p < 0.001$), but not for the dNCT or Stroop Test. Age (all $p < 0.001$) and educational level (TT $p = 0.017$, dNCT $p = 0.003$, Stroop test $p = 0.002$), in contrast, showed a significant correlation with the results of all digital tests ($p < 0.05$). Both are well-acknowledged influencing factors for the PHES, and, therefore, standard evaluation of the PHES is already adjusted to

age and educational level. Consistently, the correlation of the PHES with age was only weak in our patient cohort ($r = -0.259$; $p = 0.013$) and there was no correlation with years of school education ($r = 0.153$; $p = 0.149$) (Figure 5).

In our cohort of healthy control subjects, there was a positive correlation between age and the results of each digital test (TT $p = 0.004$, dNCT $p < 0.001$, Stroop test $p = 0.002$). However, educational level showed no significant correlation since there was only one individual with less than 10 years of schooling. As expected, results of the age- and education-adjusted PHES did not correlate with these parameters.

TABLE 2 | Test duration in seconds (median; IQR) of the three digital tests in subgroups of cirrhotic patients according to presence of mHE and dependence on age or level of education.

Cirrhotic patients	Without mHE			With mHE			Education		
	(<i>n</i> = 60)	(<i>n</i> = 32)	<i>p</i>	≤ 60 years (<i>n</i> = 46)	> 60 years (<i>n</i> = 46)	<i>p</i>	< 10 years (<i>n</i> = 42)	≥ 10 years (<i>n</i> = 48)	<i>p</i>
Tip Test	37 (33–39)	48 (39–57)	< 0.001	39 (33–48)	47 (37–55)	0.005	45 (37–52)	39 (33–49)	0.033
dNCT	10 (7–12)	14 (9–18)	< 0.001	9 (7–15)	14 (10–17)	< 0.001	15 (11–18)	10 (7–14)	< 0.001
Modified stroop	56 (48–76)	82 (62–124)	< 0.001	64 (50–81)	85 (66–119)	0.002	95 (63–126)	66 (50–81)	0.002

3.6 | Detailed Analysis of the Impact of Age and Educational Level on Digital Test Results

In order to further characterize the potential impact of age as an influential factor on the results of the smartphone-based tests, the patient cohort was divided into two subgroups based on the 50th percentile, that is individuals of ≤ 60 years and > 60 years (each group *n* = 46). For all three tests, the median time to perform the digital tests was significantly shorter in the younger compared to the older subgroup (TT *p* = 0.005; dNCT *p* < 0.001; Stroop test *p* = 0.002). Results are reported in detail in Table 2 and depicted in Figure 3b.

Overall, healthy controls performed the smartphone-based tests faster than our patient cohort (all *p* < 0.001) (Figure 2). Likewise, to the cirrhotic patient cohort, those healthy individuals aged less than 60 years (*n* = 10) were significantly faster than healthy individuals aged 60 years or older (*n* = 10) (TT *p* = 0.005; dNCT *p* < 0.001; Stroop test *p* = 0.002).

To explore the potential influence of educational level on the outcomes of the smartphone-based tests, participants were subgrouped based on the duration of their school education: individuals with 6–9 years were compared to those with 10 or more years of formal education. We chose this categorization in analogy to the subgroups of the PHES. Subjects with longer education completed all three digital tests significantly faster compared with those with less education (TT *p* = 0.033; dNCT *p* < 0.001; Stroop test *p* = 0.002) (Table 2). As expected, there was no statistically significant difference between the two patient subgroups based on their educational levels in PHES (−7.8 vs. −6.7 points; *p* = 0.328). For the cohort of healthy controls, we could not well divide into two subgroups because there was only one individual with an educational level of < 10 years.

Furthermore, we found no significant gender-related differences in the digital test results between males and females.

3.7 | Sensitivity and Specificity

Additionally, the diagnostic accuracy of the smartphone-based digital tests was examined in relation to the PHES using Receiver Operating Characteristic (ROC) curve analysis. There was an AUC of 0.832 (95% confidence interval (CI) 0.742–

0.921, *p* < 0.001) for the TT, an AUC of 0.738 (95% CI 0.632–0.845, *p* = 0.001) for the dNCT, and an AUC of 0.762 (95% CI 0.680–0.878, *p* < 0.001) for the Stroop test (Figure 6). For the TT, using Youden's Index, we calculated an optimal cut-off value of 44.5s resulting in a sensitivity of 63% and a specificity of 93%. In our clinical setting, we felt that a higher sensitivity at the cost of specificity was warranted and thus calculated the point closest to the top left, resulting in a cut-off value of 39 s with a corresponding sensitivity of 75% and specificity of 76%.

4 | Discussion

Diagnostic work-up of subclinical mHE in cirrhotic patients is highly relevant but time consuming and staff intensive, and thus difficult to integrate into clinical routine when assessed by standardized psychometric paper-and-pencil tests. Digital methods promise autarkical self-assessment with potentially earlier mHE detection. Hence, smartphone-based mHE surveillance might save time and staff resources. In this study, we present three simple smartphone-based digital mHE tests all showing good correlations with the PHES, which is the internationally accepted diagnostic standard for mHE detection.

On a closer view, the Tip test (TT) demonstrated the highest correlation coefficients with the PHES. In contrast to the PHES and Stroop test, it is independent of native language, literacy level and color vision, allowing widespread applicability. Furthermore, it is independent of red-green color-blindness, which further enhances its universal use. The TT in particular received favorable feedback from patients (data not shown). Our modified Stroop test demonstrated short testing times with still acceptable correlations to PHES. Recently, Bajaj et al. have also addressed the issues of test duration and ease of use in their improved “Quick Stroop” version, featuring only the easier “Off” State of their test [20]. Our third test, the dNCT, showed the weakest correlations and AUC. This mirrors the low sensitivity of NCT A and B as stand-alone tests as described by Weissenborn et al. in their original characterization of PHES [12]. Therefore, the dNCT does not appear promising as a future stand-alone testing method.

In our healthy control subjects, an excellent performance at digital tests was in concordance with normal PHES-results. Since these controls showed no signs of cognitive impairment,

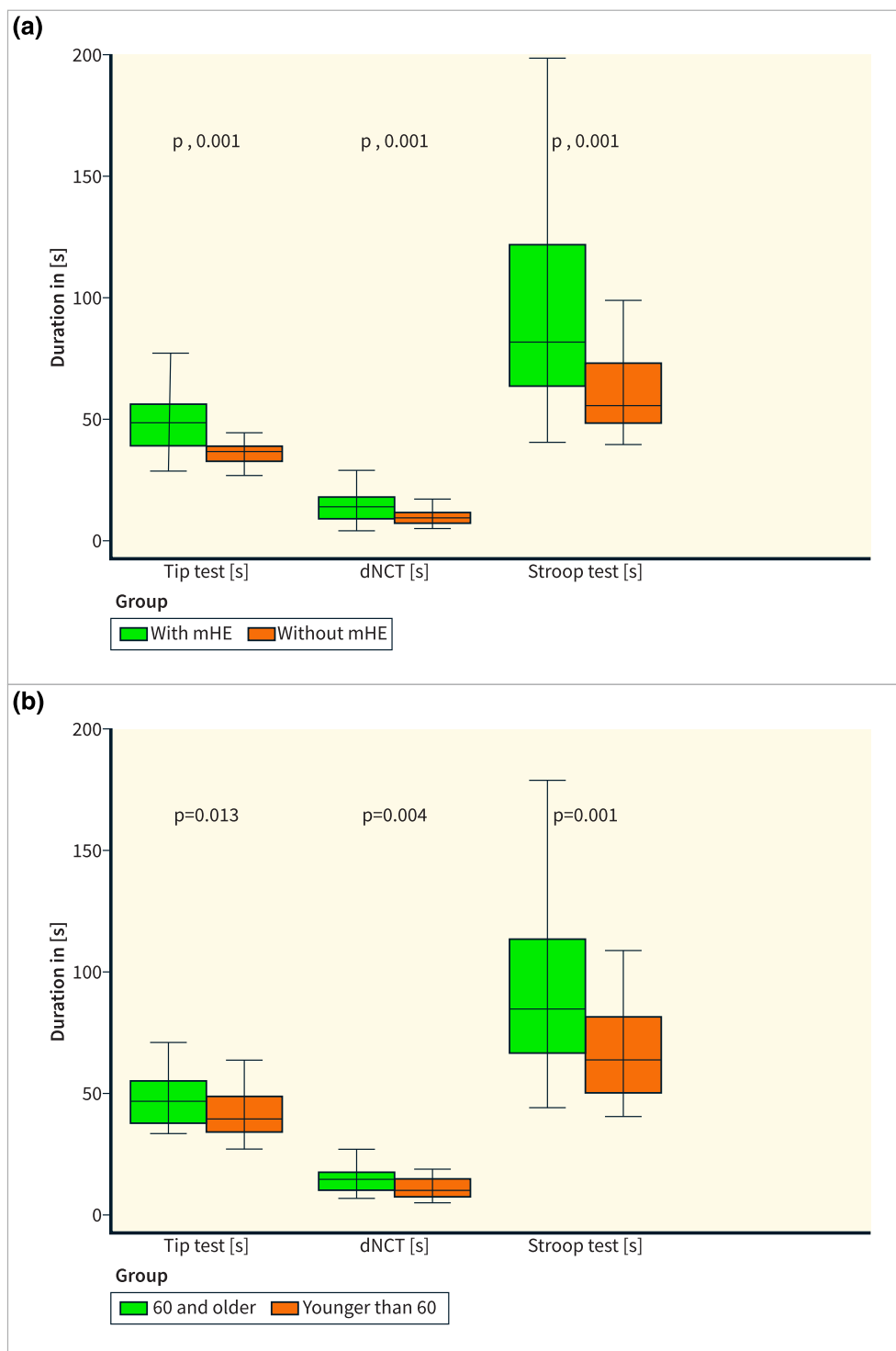


FIGURE 3 | Box-plots comparing time required to perform digital Tip test, Stroop test, and dNCT (a) in cirrhotic patients with versus without mHE and (b) for the patient cohort dichotomized at the 50th percentile into age group < 60 years versus ≥ 60 years.

there was little variation in test results in both PHES and digital testing. Due to this rather uniform performance without substantial scattering, we found no correlation between digital tests and PHES, as expected.

In patients with cirrhosis, digital tests discriminated well between patients with and without mHE and correlated significantly with

MELD and Child–Pugh scores. These findings are plausible as severe cognitive impairment is well-described in patients with advanced liver disease [2, 21]. In cirrhotic patients, multiple co-morbid factors might additionally influence the development of mHE. In our cohort, for example, chronic alcohol consumption was found to be the leading etiology of cirrhosis. Underlying neurotoxic damage might interfere with smartphone-based mHE

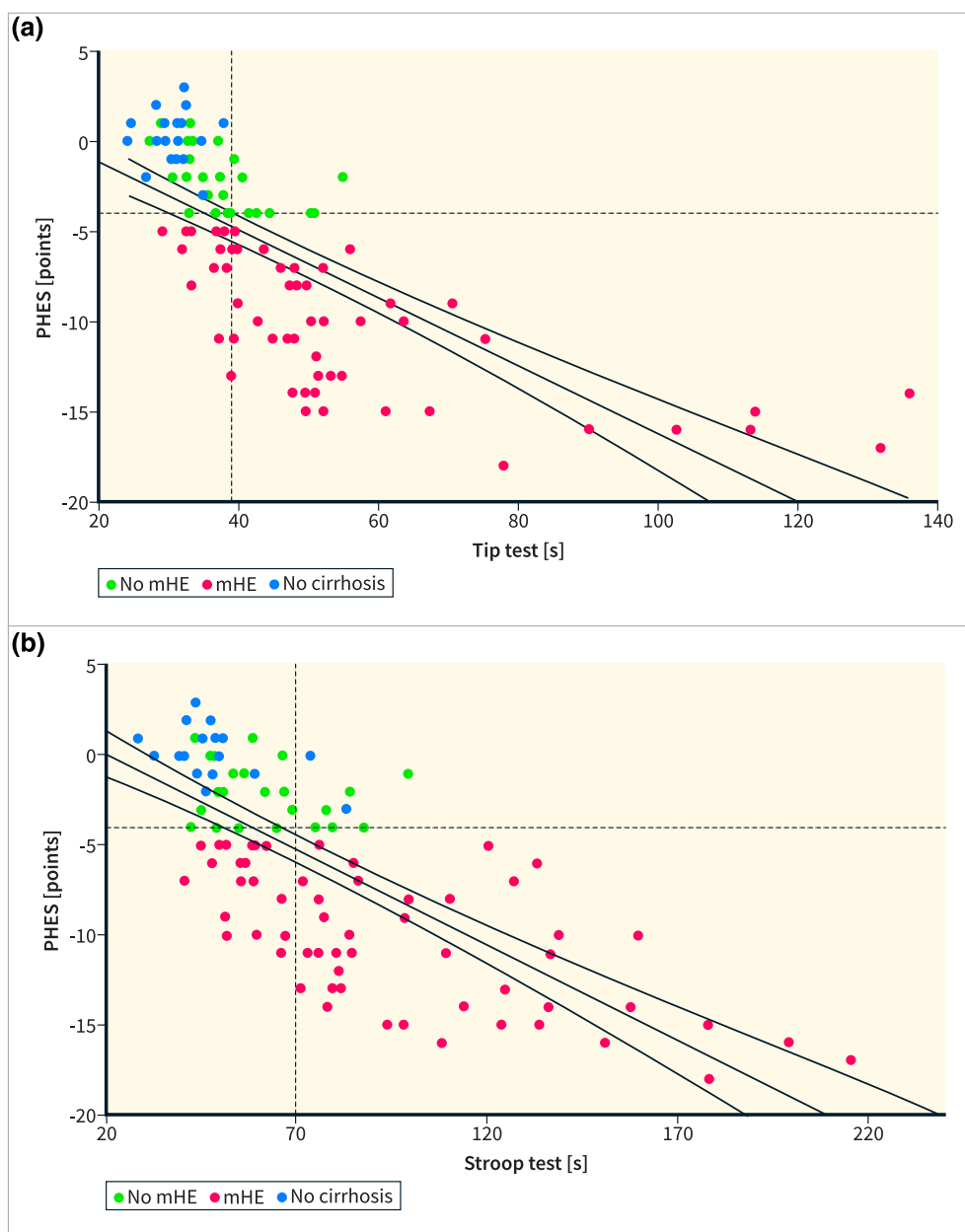


FIGURE 4 | Scatter plot and linear regression (CI 95%) of the (a) digital Tip Test (TT) and (b) Stroop Test with reference to PHES with a PHES of < -4 defining mHE (depicted by the horizontal dotted line) Green dots depict cirrhotic patients without mHE, red dots cirrhotic patients with mHE and blue dots healthy controls without mHE. In accordance with the point closest to the top left at ROC analysis (Figure 6), a cut-off duration of about 39 s for the TT (depicted by the vertical dotted line) yielded a sensitivity of 75% with specificity of about 76%, and of about 70 s for the modified Stroop Test (sensitivity 70%, specificity 72%).

self-assessment, but should likewise confound the PHES and therefore not affect our correlation analysis [21].

Apart from clinical confounders and medication, constitutional factors such as ethnicity, age, gender, or education might influence mHE assessment. The PHES already acknowledges this by adjusted reference values [19]. In digital testing, age emerged as an important determinant in both patients and controls. This was most evident in all digital tests ($p < 0.05$) when comparing patients aged ≤ 60 with those > 60 years, representing the 50th percentile of our study population, emphasizing age as a substantial determinant. It's worth noting that in addition to a

confounding effect, advanced age can elevate the risk of mHE, leading to a higher prevalence in older patients [6]. To further explore age-related effects, future studies should aim at deriving age-adjusted thresholds for better digital mHE discrimination. Still, the fact that the present non-age-adjusted version of our digital tests did correlate well with the age-corrected PHES obviously argues in favor of its strong value independent of age.

Educational level exhibited a significant correlation with the digital test results in our patient cohort. Notably, for the TT, this aspect had only weak statistical significance. In our healthy controls, we did not observe any significant correlation, which

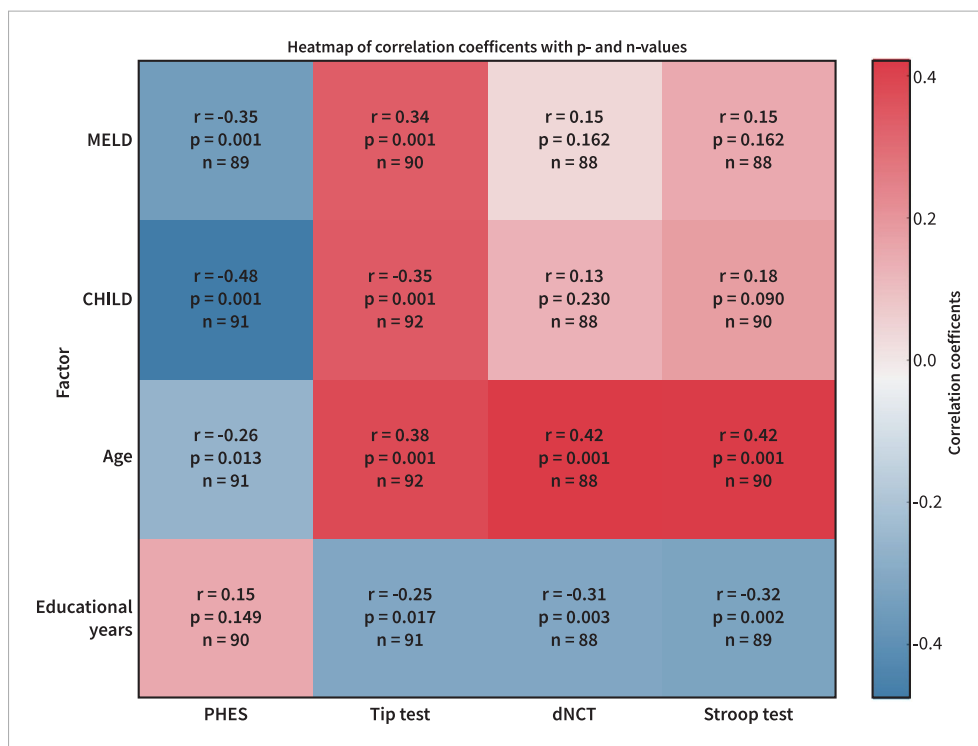


FIGURE 5 | Correlation coefficients (r) for the PHES and the 3 digital tests in cirrhotic patients with regard to scores of disease severity (MELD, Child Pugh), age and years of school education. Blue: negatively correlated, red: positively correlated, intensity of color: strength of correlation.

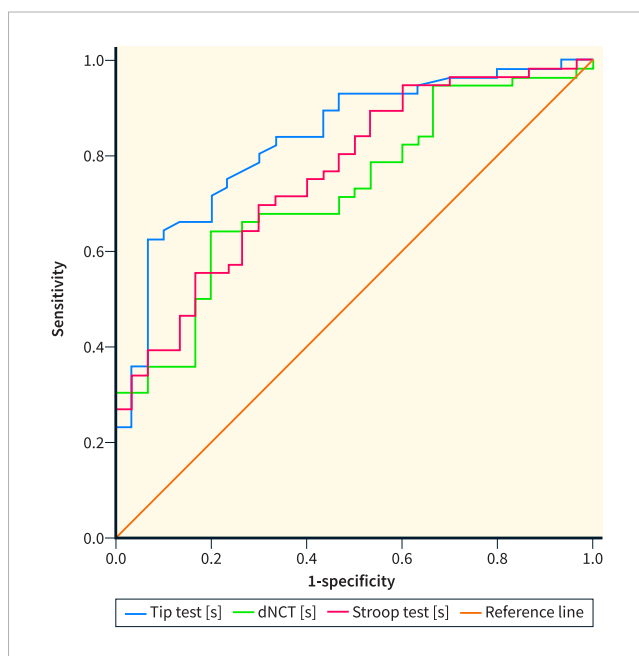


FIGURE 6 | Receiver operating characteristic (ROC) curves for the digital mHE tests (Tip test: AUC 0.832, dNCT: 0.738 and Stroop test: AUC 0.779) with reference to the PHES as diagnostic standard.

might be attributed to the homogeneity of educational levels in this cohort. The PHES, already adjusted for education, showed no such disparities [19]. In addition, evaluation of the PHES is gender-specific. For the digital tests, our patient cohort did not show any gender-related differences. However, the gender imbalance in our cohort of patients with mHE, with male

participants outnumbering females ($n = 40$ vs. 20), might limit statistical accuracy and warrant approval in larger cohorts.

4.1 | Perspective

The present work reports the head-to-head comparison of smartphone-based digital tests with concurrent mHE assessment by PHES as the well-acknowledged gold standard. At this level, the digital tests proved to be a rapid and reliable patient-based approach to detect mHE. Their ease of use, speed, and discriminatory capacity could make them particularly valuable in clinical settings inside and outside the hospital. In particular, the TT emerges as an attractive and pragmatic, language and color-neutral screening tool. Potential learning effects of regularly repeated self-screening as well as constitutional individual factors warrant consideration. To address these caveats, it would be conceivable to determine individual baseline and threshold levels based on a smartphone-app. Subsequently, this app could generate a continuous degradation alert in case mHE develops or deteriorates. This should prompt consultation at the hepatology outpatient clinic at an early stage, to check for the need of interventions and avoid hospitalization. While the present pilot-trial aimed at characterizing the novel digital test strategies for mHE detection as compared to the PHES, our findings warrant confirmation and further adjustment in larger longitudinal trials. As next steps these should aim (1) at defining age-adjusted reference margins for the digital tests, (2) at investigating their suitability for regular self-surveillance in everyday life and (3) at validating their diagnostic efficacy and clinical prognostic value. Further studies should implement individualized adaptations, to improve independent and regular self-

surveillance, thereby facilitating early and patient-refined mHE management with improved prognostic outcome.

Acknowledgments

Thanks to Marc Strugar from Department of Electrical Engineering and Computer Science, Technische Hochschule Lübeck, University of Applied Sciences for his computational help with digital test programming.

Conflicts of Interest

J.U.M. and F.S. received honoraria from Merz and abbvie, which were not related to the current study. No conflicts of interest, financial or otherwise, are declared by the other authors.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. Montagnese, P. E. Rautou, M. Romero-Gómez, et al., "EASL Clinical Practice Guidelines on the Management of Hepatic Encephalopathy," *Journal of Hepatology* 77, no. 3 (2022): 807–824, <https://doi.org/10.1016/j.jhep.2022.06.001>.
2. H. Vilstrup, P. Amodio, J. Bajaj, 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: Vilstrup et al," *Hepatology* 60, no. 2 (2014): 715–735, <https://doi.org/10.1002/hep.27210>.
3. I. J. C. Hartmann, M. Groeneweg, J. C. Quero, et al., "The Prognostic Significance of Subclinical Hepatic Encephalopathy," *American Journal of Gastroenterology* 95, no. 8 (2000): 2029–2034, <https://doi.org/10.1111/j.1572-0241.2000.02265.x>.
4. J. Cordoba, M. Ventura-Cots, M. Simón-Talero, et al., "Characteristics, Risk Factors, and Mortality of Cirrhotic Patients Hospitalized for Hepatic Encephalopathy With and Without Acute-On-Chronic Liver Failure (ACLF)," *Journal of Hepatology* 60, no. 2 (2014): 275–281, <https://doi.org/10.1016/j.jhep.2013.10.004>.
5. M. Luo and W. K. Cao, "Advances in Psychometric Tests for Screening Minimal Hepatic Encephalopathy: From Paper-And-Pencil to Computer-Aided Assessment," *Turkish Journal of Gastroenterology* 30, no. 5 (2019): 398–407, <https://doi.org/10.5152/tjg.2019.18226>.
6. V. Felipo, "Hepatic Encephalopathy: Effects of Liver Failure on Brain Function," *Nature Reviews Neuroscience* 14, no. 12 (2013): 851–858, <https://doi.org/10.1038/nrn3587>.
7. G. Kircheis, A. Knoche, N. Hilger, et al., "Hepatic Encephalopathy and Fitness to Drive," *Gastroenterology* 137, no. 5 (2009): 1706–1715.e9, <https://doi.org/10.1053/j.gastro.2009.08.003>.
8. E. B. Tapper, M. Romero-Gómez, and J. S. Bajaj, "Hepatic Encephalopathy and Traffic Accidents: Vigilance Is Needed," *Journal of Hepatology* 70, no. 4 (2019 Apr): 590–592, <https://doi.org/10.1016/j.jhep.2019.01.017>.
9. V. V. Mittal, B. C. Sharma, P. Sharma, and S. K. Sarin, "A Randomized Controlled Trial Comparing Lactulose, Probiotics, and L-Ornithine L-Aspartate in Treatment of Minimal Hepatic Encephalopathy," *European Journal of Gastroenterology and Hepatology* 23, no. 8 (2011): 725–732, <https://doi.org/10.1097/meg.0b013e32834696f5>.
10. M. M. Lerch, M. Rathmayer, B. Siegmund, et al., "Die Grenzen des G-DRG-Systems bei der Abbildung von Komplexität in der Universitätsmedizin," *Zeitschrift für Gastroenterologie* 58, no. 08 (2020): 747–753, <https://doi.org/10.1055/a-1219-8245>.

11. S. J. Gairing, C. Mangini, L. Zarantonello, et al., "Minimal Hepatic Encephalopathy Is Associated With a Higher Risk of Overt Hepatic Encephalopathy and Poorer Survival," *Journal of Internal Medicine* 295, no. 3 (2024): 331–345, <https://doi.org/10.1111/joim.13747>.
12. K. Weissenborn, J. C. Ennen, H. Schomerus, N. Rückert, and H. Hecker, "Neuropsychological Characterization of Hepatic Encephalopathy," *Journal of Hepatology* 34, no. 5 (2001): 768–773, [https://doi.org/10.1016/s0168-8278\(01\)00026-5](https://doi.org/10.1016/s0168-8278(01)00026-5).
13. J. S. Bajaj, A. Etemadian, M. Hafeezullah, and K. Saeian, "Testing for Minimal Hepatic Encephalopathy in the United States: An AASLD Survey," *Hepatology* 45, no. 3 (2007): 833–834, <https://doi.org/10.1002/hep.21515>.
14. C. Labenz, L. Beul, G. Toenges, et al., "Validation of the Simplified Animal Naming Test as Primary Screening Tool for the Diagnosis of Covert Hepatic Encephalopathy," *European Journal of Internal Medicine* 60 (2019): 96–100, <https://doi.org/10.1016/j.ejim.2018.08.008>.
15. P. Amodio, J. C. Quero, F. Del Piccolo, A. Gatta, and S. W. Schalm, "Diagnostic Tools for the Detection of Subclinical Hepatic Encephalopathy: Comparison of Standard and Computerized Psychometric Tests With Spectral-EEG," *Metabolic Brain Disease* 11, no. 4 (1996): 315–327, <https://doi.org/10.1007/bf02029493>.
16. J. S. Bajaj, L. R. Thacker, D. M. Heuman, et al., "The Stroop Smartphone Application Is a Short and Valid Method to Screen for Minimal Hepatic Encephalopathy," *Hepatology* 58, no. 3 (2013): 1122–1132, <https://doi.org/10.1002/hep.26309>.
17. J. S. Bajaj, D. M. Heuman, R. K. Sterling, et al., "Validation of EncephalApp, Smartphone-Based Stroop Test, for the Diagnosis of Covert Hepatic Encephalopathy," *Clinical Gastroenterology and Hepatology* 13, no. 10 (2015): 1828–1835.e1, <https://doi.org/10.1016/j.cgh.2014.05.011>.
18. S. Allampati, A. Duarte-Rojo, L. R. Thacker, et al., "Diagnosis of Minimal Hepatic Encephalopathy Using Stroop EncephalApp: A Multicenter US-Based, Norm-Based Study," *American Journal of Gastroenterology* 111, no. 1 (2016): 78–86, <https://doi.org/10.1038/ajg.2015.377>.
19. H. Schomerus, K. Weissenborn, H. Hecker, W. Hamster, and N. Rückert, *PSE Syndrom Test Psychodiagnostisches Verfahren zur quantitativen Erfassung der (minimalen) portosystemischen Enzephalopathie (PSE)* 3 (Medizinische Hochschule Hannover, 2021): überarbeitete Auflage.
20. C. Acharya, J. Shaw, N. Duong, et al., "QuickStroop, a Shortened Version of EncephalApp, Detects Covert Hepatic Encephalopathy With Similar Accuracy Within One Minute," *Clinical Gastroenterology and Hepatology* 21, no. 1 (2023): 136–142, <https://doi.org/10.1016/j.cgh.2021.12.047>.
21. C. F. Rose, P. Amodio, J. S. Bajaj, et al., "Hepatic Encephalopathy: Novel Insights into Classification, Pathophysiology and Therapy," *Journal of Hepatology* 73, no. 6 (2020): 1526–1547, <https://doi.org/10.1016/j.jhep.2020.07.013>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.