

Determination of Turkish Norms of Psychometric Tests For Diagnosing Minimal Hepatic Encephalopathy and Proposal of a Simplified Paper Pencil Test Battery

Burak Özbaş (✉ drburakozbas@hotmail.com)

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi <https://orcid.org/0000-0001-9850-0481>

Onur Keskin

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Hartmut Hecker

Hannover Medical School: Medizinische Hochschule Hannover

İrfan Karahan

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Cansu Özbaş

Gazi University Faculty of Medicine: Gazi Üniversitesi Tıp Fakültesi

Çağdaş Kalkan

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Aysun Kartal

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Fatih Oğuz Önder

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Burcu Kahveci Öncü

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Genco Gençdal

Koç Üniversitesi: Koc Üniversitesi

Murat Akyıldız

Koç Üniversitesi: Koc Üniversitesi

Fulya Günşar

Ege University Faculty of Medicine: Ege Üniversitesi Tıp Fakültesi

Ramazan Idilman

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Karin Weissenborn

Hannover Medical School: Medizinische Hochschule Hannover

Ömer Özütemiz

Ege University Faculty of Medicine: Ege Üniversitesi Tıp Fakültesi

Cihan Yurdaydin

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Keywords: minimal hepatic encephalopathy, portosystemic hepatic encephalopathy score, critical flicker frequency, cirrhosis, Turkish norms

Posted Date: February 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-170557/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Psychometric hepatic encephalopathy score (PHES) needs standardization for countries and is used for minimal hepatic encephalopathy (mHE) diagnosis. Assessing it is cumbersome.

Aims: To standardize PHES for Turkish patients and compare it with German norms; to determine mHE prevalence with two different methods and to assess whether the PHES test can be simplified.

Methods: Healthy volunteers (n=816; 400 male) and cirrhotics (n=124; 58 male) were included. PHES score threshold was set at ≤ -5 points and that of critical flicker frequency (CFF) at < 39 Hz for mHE diagnosis. For comparing German and Turkish norms, datasets were combined. Multiple backward procedure was applied. Receiver operating characteristic (ROC) curves were created for assessing diagnostic capabilities of single subtests of the PHES test.

Results: Turkish norms displayed non-linear dependence on age and education. Rate of mHE in compensated cirrhotics was 29.8% and 27.4% with PHES and CFF tests, respectively. Compatibility of two tests was low (kappa coefficient 0.389); mHE prevalence decreased to 16% when both tests were combined. Turks performed worse vs Germans in the digit symbol (DS) and serial dotting (SD) subtests but generally performed better in the other subtests. In ROC analyzes of subtests, AUROC was 0.974 for DS+SD tests combined.

Conclusions: PHES norms for Turks were developed. Use of 2 methods for diagnosing mHE may be important for research purposes. From a clinical perspective sensitivity with acceptable specificity may suffice. Combined use of DS and SD subtests of the PHES battery may well be suited for use in clinical practice.

Introduction

Hepatic encephalopathy (HE) is a brain dysfunction caused by liver insufficiency and/or portosystemic shunting. It manifests as a wide spectrum of neurological or psychiatric abnormalities ranging from subclinical alterations to coma [1]. Minimal HE (mHE) represents the mildest form of HE which is not recognizable by a standard neurological examination. However, deficits in attention, working memory, psychomotor speed or visuospatial ability are detected by means of psychometric tests. Furthermore, electrophysiological tests such as EEG and evoked potentials or psychophysiological tests such as critical flicker frequency (CFF) may show alterations of brain function in these patients [2, 3].

Despite its concealed or covert clinical characteristics, several reports point to the clinical significance of mHE: it impairs quality of life [4-6]; affects fitness to drive [7-9], predicts overt HE [10, 11] and mortality from liver disease [12, 13]. Further, the working capacity of patients with mHE may be reduced as outlined by early retirement of blue-collar workers with mHE [6].

The prevalence of mHE varies from 20% to 84% in different studies [1]. This high range of variation between studies depends mainly on two factors: the different methods employed for mHE diagnosis and the characteristics of the population studied [14, 15]. Among the latter, variables including cultural background, age and education are known confounding factors affecting test performance. Further, alcohol abuse and other accompanying diseases need to be considered as important confounding factors [15, 16].

Methods used for diagnosing mHE can be grouped into psychometric and neurophysiological tests. The most commonly used psychometric test is the PSE (Portosystemic Encephalopathy)-Syndrome test which yields the psychometric hepatic encephalopathy score (PHES) [17]. Other tests, that have been frequently used are the continuous reaction time test (CRT) [18], inhibition control test (ICT) [19], Stroop [20] and SCAN tests [12].

Neurophysiological measures include the EEG and evoked potentials. Another quite often used method is the critical flicker frequency (CFF) test [21]. Each method has its advantages and disadvantages, and the combination of at least two testing strategies is recommended especially for multi-centered studies, provided that one method is a psychometric test [22].

Of the many methods available the PHES test represents the most often used test for the diagnosis of mHE. The test has been thoroughly validated. Hamster and Schomerus [23], who developed the test battery, had assessed numerous psychometric tests for their use in diagnosing HE, before they finally created the PHES-test battery with the following paper-pencil tests: Digit Symbol Test (DST), Number Connection Test-A (NCT-A), Number Connection Test-B (NCT-B), Serial Dotting Test (SDT), and Line Tracing Test (LTT). This test battery has been considered to be the gold standard diagnostic method for the diagnosis of mHE by experts in the field [24]. One concern with this test battery was the learning effect when repeated. Therefore, four parallel versions of the test have been elaborated and validated for follow-up examinations.

The PHES test battery has been evaluated in several countries which include for example Germany, Italy, Spain, Poland, South Korea and Mexico [17, 25-29]. Population-based norms were established separately for these countries. As a neuro- or psychophysiological test, the CFF test seems attractive since it is easy to perform and may have a pathophysiological link to HE as well. It was reported that in patients with liver cirrhosis retinal glial cells present pathological changes, named 'retinal gliopathy', reminiscent of type II astrocytosis in the central nervous system [30]. CFFs not affected by educational background but changes with age [9, 21, 31, 32].

In the present study, we primarily aim to establish Turkish norms for the PHES test battery, and then determine the prevalence of mHE in patients with compensated cirrhosis by using two different methods for the detection of mHE, namely the PHES test battery and the CFF test as recommended in the recent EASL-AASLD guidelines [33]. In addition, using receiver operating characteristic curve analyses we explored whether the current 5 test assessment for diagnosing mHE can be replaced by a simplified test battery.

Patients And Methods

Selection of Patients and Healthy Subjects

This study was conducted on a total of 816 healthy subjects distributed evenly according to age, sex, and educational background characteristics and 124 patients with a diagnosis of liver cirrhosis at the Gastroenterology Department of Ankara University Medical School from 01/01/2015 to 7/31/2015. Figure 1 shows the flow diagram of the study.

The study consisted of two parts; in the first part normal values of the PHES tests in the Turkish population based on age and educational level was determined, and compared to German norms [17]. In the second part, prevalence of mHE in compensated cirrhotic patients was explored. For the first study, healthy subjects, 18 to 70 years old, were enrolled (Suppl. Table 1). Subjects who had consumed alcohol or had used psychoactive drugs within the last 3 months were excluded. Further exclusion criteria were visual or hearing impairment, neurological diseases and illiteracy. For the cirrhotic patients, besides the listed exclusion criteria for healthy subjects, decompensated cirrhosis, a history of HE, esophageal variceal bleeding, hepatocellular carcinoma and any significant concomitant disease which might have affected the results of the study such as neuropsychiatric disease, respiratory, kidney or heart failure, or any malignant disease were exclusion criteria as well. The study was conducted in line with the Declaration of Helsinki. The study was approved by the University of Ankara Medical School Ethics Committee. All patients gave written informed consent.

Neuropsychometric Tests

The PHES test battery applied to patients and healthy subjects consisted of the DST, NCT-A, NCT-B, SDT and LTT. LTT was assessed separately based on time and the number of errors made.

PHES test score calculation was made according to the study of Weissenborn et al [17]. Thus, for each test the score was depicted as 0 when the test value was between -1 SD to +1 SD; the score was -1 for values between -1 SD to -2 SD; it was -2 for values between -2 SD to -3 SD, -3 for values below -3 SD, and the score was depicted as +1 when the values exceeded +1 SD. Thus, the total score interval ranged between -18 and +6. Scores equal to or lower than -5 were considered indicative of mHE based on the results in the norm population. The cut-off is the same as in the study by Weissenborn et al [17].

Critical flicker frequency

CFF measures the threshold at which fused light impression switched to flickering light impression perceived by the patient when frequency of light pulses from a light source decreased stepwise by 0.1 Hz from 60 Hz downwards. A total of nine measurements were recorded and the mean of those 9 measurements was calculated. CFF <39 Hz was taken to be the threshold for mHE, in line with previous studies [34]. CFF and the PHES test battery were administered at the same session.

Statistical Analysis

Data were evaluated with SPSS 15.0 statistical package program. Descriptive statistics were presented as mean (\pm) standard deviation, or median and percentiles. Statistical methods included the Chi-Square test and Fisher's test, for categorical variables and Student's t and Mann-Whitney U tests for continuous variables. Correlation between continuous variables was evaluated with Spearman's correlation analysis for non-normal distribution and with Pearson's correlation analysis for normal distribution. $P < 0.05$ was considered significant.

For determination of the Turkish norms for the PHES test, a transformation (log for NCT-A, log-log for NCT-B, log-log-log for SDT, square root for DST, log-log for LTT-time, and cube root for LTT-error) was performed for each sub-test as the data presented by healthy individuals did not conform to a normal distribution, assessed with the Kolmogorov-Smirnov test. In the transformed scales, norm limits were defined as the values of age dependent mean and of deviations of $k = -1, +1, +2, +3$ (NCTA, NCTB, LTT Time, SDT, LTT Error) or $k = +1, -1, -2, -3$ (DST) standard deviations from the mean value. Thus, deviances indicating worse performance of a subtest may be classified into three categories, while better performances form a single category. Retransformation for the log-transformation is based on the formula:

$$\exp(a + b \cdot \text{age} \pm k \cdot s).$$

For the iterated logarithm (log-log), the retransformation formula is:

$$\exp \exp(a + b \cdot \text{age} \pm k \cdot s)$$

and for the threefold logarithm (log-log-log):

$$\exp \{ \exp \exp(a + b \cdot \text{age} \pm k \cdot s) \}.$$

For the square root transformation it is:

$$(a + b \cdot \text{age} \pm k \cdot s)^2$$

and

$$(a + b \cdot \text{age} \pm k \cdot s)^3$$

for the cube root transformation.

Here, a denotes the constant (intercept) of the regression line in the transformed scale, b the slope of the line and s the standard deviation of the residuals. k is the integer value to be replaced by the defined number of standard deviations for the different norm limits.

The impact of covariates on the test results was analyzed as follows: for each subtest, age was always included into the regression model, applied to the transformed subtest results as described above. The two further covariates (education and gender) were included and handled in a stepwise procedure: after inclusion of both of them, at each step, the covariate with the highest p-value was excluded from the model as long as this p-value was > 0.05 . Formal Education was handled as a categorical variable with contrast of "H" (for high school) and of "U" (university) vs. "P" (primary school). For each subtest, additional outliers were excluded from the model. The final regression model then contained the age together with the covariates relevant for this subtest. As a consequence, the norm values were recomputed for each subtest taking into account the relevant covariate values of each subject.

The homogeneity of variances among different groups was analyzed with the Levene test to ensure homogeneity. Individuals falling outside of ± 3 standard deviation were excluded from the analysis. Distribution of test data was assessed according to educational background and age with a two-way variance analysis (two-way Anova test).

In analyzing German [17] and Turkish norms, in order to compare the regression characteristics for subtests with equal transformation, the datasets were combined, and a multiple backward procedure was applied allowing for different intercepts and slopes in the Turkish and the German populations. The differences were tested within the framework of the linear model and were eliminated from the model if $p > 0.05$. The standard deviations of the residuals were compared by applying Levene's test of equal variances. The test results are given in the last column of table 1.

Finally, receiver operating characteristic (ROC) curves were created for assessing the diagnostic capabilities of single subtests of the PHES test. Then, ROC curves for combination of two single subtests were established. The area under ROCs (AUROC) of single or double subtests were compared using Hanley-McNeil tests [35]. A p-value of < 0.05 was considered significant.

Results

The study included 816 healthy controls and 124 patients with compensated cirrhosis. Healthy subjects had a mean age of 40.5 ± 14 (mean age \pm SD) and 400 (49%) were male. While 30% (246) of healthy subjects were primary school graduates, 33% (271) were middle or high school graduates and 37% (299) were university and above graduates. The distribution of healthy subjects according to age and educational level is depicted in Suppl. Tables 1 & 2. Patients with compensated cirrhosis had a mean age of 50.6 ± 12.3 and 58 (47%) were male (Suppl. Table 2); more than 50% had only primary school education.

Turkish Norms of the PHES test battery;

Analysis of test results displayed a heterogenous distribution. A normal distribution was achieved after logarithmic transformation.

As shown in figure 2, a non-linear dependence on age by linear regression analysis was found with all paper pencil tests assessed. An influence of formal education is present in all six subtests (with advantages of "H" and "U" vs. "P" as given by the two regression coefficients) (Fig. 3). Gender attributes to the norm values only in the Serial Dotting subtest, with better female results. Based on these data Turkish norms of the PHES test battery were created taking into account age and level of education. In Suppl. Table 3, normal values with standard deviations are presented according to age and education level for every paper pencil test.

The distribution of the PSE sum scores (mean value: -0.121; standard deviation $s = 2.035$) in the Turkish population (after the exclusion of outliers in the subtests) is shown in Suppl. Table 4. The cut-off point for out-of-norm negative (pathological) values was set at ≤ -5 points; 19 of 785 cases (2.4%) fell out of the normal range.

Comparison of German and Turkish Norms

The derivation of the German (Hannover) norm values was based on $N=249$ for the NCTA and NCTB and on $N = 120$ healthy subjects for the other four subtests [17].

As shown in table 1, the transformations used to obtain linear dependency between age and test results and homoscedasticity of the residuals were the same in the Turkish and the German population only for the subtest NCT-A. The discrepancies in the other subtests were mainly due to the finding that in the Turkish population the standard deviations of the residuals were still increasing with age after simple (log- or square-) transformation.

In all five subtests the age-adjusted values differed significantly in the two populations ($p < 0.001$ to $p = 0.006$; see table 1). For NCTB, LTT time and LTT Error, higher subset score values were seen in the Turkish compared to the German population (i.e.: Turkish subjects generally needed less time for the Number Connection Test B and for the Line Tracing Test, and made less errors in the Line Tracing Test test). However, for the DS and SDT tests, lower subset scores were observed for the Turkish compared to the German population (i.e. generally lower numbers were achieved in the Digit Symbol test and more time was needed for the Serial Dotting Test).

Comparison of the PSE sum scores in the German and Turkish healthy populations.

In order to conduct an overall comparison of the German and the Turkish norm population data, the German norms and the Turkish norms were applied to both populations. Thus, a comparison between the populations and between the norm values was possible on the basis of each of the two norm definitions.

On the basis of the Hannover norm definition, the mean Sum Score was significantly lower (that is: worse) in the Turkish population (-1.22 vs. 0.025; unpaired t-test: $p < 0.001$). Also, the standard deviation was significantly higher (2.987 vs. 1.994; $p < 0.001$; Levene test of equal variances, Suppl. Table 5).

If the Turkish norm is applied, no significant differences are found between the populations. In the Turkish population, 19 out-of-norm cases would also be classified as out of norm by the Hannover norm definition, but 142 of 785 subjects classified as normal would be classified as out of norm if the Hannover Scores were applied (Suppl. Fig 1). This is indicating that the Turkish norm limits are generally wider. The difference is significant ($p < 0.001$; test of McNemar). It must be the effect of education which leads to the wider norm limits in the Turkish compared to the German population. It suggests that education standards are more homogenous in Germany and more heterogenous in Turkey. Difference of the norm definitions is mainly apparent for negative Sum Score results (Suppl. Fig 1).

Minimal HE prevalence in patients with compensated cirrhosis

Once PHES test cut-offs within the Turkish population was established, the study proceeded to the second stage i.e. estimation of mHE prevalence in patients with compensated cirrhosis. To this end, two different methods as recommended by the latest EASL-AASLD guidelines were used: the PHES test battery for neuropsychometric assessment, and the CFF test as a neurophysiological method [1].

Patients with cirrhosis were older and less well educated compared to healthy controls (Suppl. Table 2). Within the compensated cirrhotic patient group, the most frequent observed etiology of liver disease was chronic hepatitis B (21.8%), followed by cryptogenic cirrhosis (11.3%), and chronic hepatitis C (11.3%) (Suppl. Table 6). Majority of the so called cryptogenic cirrhotic cases are most likely due to non-alcoholic liver disease. Suppl. Table-7 depicts biochemical and hematological characteristics of the cirrhotic patient group. 80% (100/124) of the patients were Child-Pugh Class A and 20% (24/124) were Class B. Median MELD score was 9 (6-20), median APRI and FIB-4 scores were 0.93 (0.1-5.6) and 3.1 (0.4-8.4), respectively.

mHE prevalence was 29.8% using our newly established Turkish norms for the PHES test battery. With the CFF test, mHE prevalence was 27.4%. The number of patients who were considered to have mHE was 37 according to the PHES test battery and 34 with the CFF test. Of these 37 patients; 17 (46%) were normal according to the CFF test. When patients who were outside the norms with both tests were considered as diagnostic of mHE, patients with mHE would be 16% (20/124). We assessed the compatibility of CFF test with the PHES test battery using the Cohen's kappa coefficient of concordance. Kappa coefficient of concordance was 0.389 ($p < 0.001$). However, this was still a weak concordance due to a low coefficient of concordance.

Assessment of PHES subtests to replace the whole PHES test battery:

We performed receiver operator characteristic (ROC) analyses for PHES subtests for predicting mHE by comparing them with the PHES total score. Of the PHES subtests, the best area under ROC (AUROC) curves were achieved with the DST, NCT-A and SDT (Table 2). The DST performed better than NCT-B, LTT time and LTT error ($p < 0.05$ for all by the Hanley McNeil test, Table 3). In comparison with the PHES test, area under ROC curve (AUROC) for the DST test was 0.925. With a cut-off point of -0.5, the sensitivity was 94.6% and the specificity 79.3% for predicting mHE based on the PHES total score (Cohen's kappa= 0.657). AUROC was 0.974 when DST and SDT were combined and this 2 subtest combination had the best AUROC compared to other binary tests (Table 3). Further, the triple test combination with the highest AUROC (DST+SDT+NCT-A) did not do better than this dual test combination in diagnosing mHE (Table 3). For this pair, at a cut-off of -1.5, sensitivity was 97.3% and specificity 86.2% for diagnosing mHE (Cohen's kappa = 0.769) (Table 2). Thus, out of 37 patients diagnosed with mHE according to the PHES test battery, 36 had been caught with this binary test combination. However, out of 87 patients without a diagnosis of mHE with the PHES test battery, 12 patients would have been misdiagnosed with mHE. For a cut-off of -2.5, sensitivity was 78.4% and the specificity 97.7% (Cohens Kappa = 0.798). When checking other PHES subtests and their combinations, the use of this pair had the closest result to the PHES total score. We performed ROC analysis similarly for the CFF test. The AUROC curve for the CFF test to predict the PHES total score based mHE diagnosis was 0.747. When the cut-off was set at 39.0 hz., the sensitivity and specificity were 54.0% and 83.9%, respectively.

Comparison of demographic, biochemical and non-invasive fibrosis markers in patients with or without mHE:

As shown in Suppl table 8 patients with mHE according to the PHES compared to those without had higher MELD and Child-Pugh scores. They had more advanced Child-Pugh class, and higher FIB-4 scores). Blood ammonia levels and

total bilirubin levels were higher and albumin was lower). CHILD score and class as well as albumin level were also associated with mHE diagnosis with the CFF test, but the most significant effect on CFF had age (Suppl. Table 9).

Discussion

The present study had three aims. The first and main aim was the establishment of the Turkish normal values for the Porto-Systemic Encephalopathy (PSE)-Test that provides the psychometric hepatic encephalopathy score (PHES) and compare it with German norms. The present study has standardized the PHES test battery for Turkish people, according to age and educational background. We have provided as Suppl. material norm values for every test according to age and level of education. The second aim of the study was the assessment of mHE prevalence in cirrhotic patients according to the newly established cut-offs of the PHES test and the CFF test. The third aim was to assess if the PHES test could be replaced with a simpler version. Using receiver operating characteristic curve analyses we were able to provide data to show that two subtests of the PHES test battery, namely the DST and SDT can be used for a rapid screening for mHE.

Our data confirm that the PHES test battery score varied depending on age and educational background of subjects. Test performance decreased with advanced age. These findings are in line with the data reported for Germany, Italy, Spain, Poland, South Korea and Mexico [17, 25-29]. In addition, test performance also decreased with lower educational background in line with data from Poland, Spain, Italy, South Korea and Mexico but in variance with data obtained from Germany in 1999 [36] (Suppl. Table 10). The results obtained from healthy volunteers were compared in detail with German data and there was a significant difference between the groups for all sub-tests ($p < 0.001$ - $p = 0.006$) [17]. Healthy Turks compared to healthy Germans needed less time for the NCT B and for the LTT, and made less errors in the LTT. However, Turks compared to Germans achieved lower numbers in the DST and they needed more time to complete the SDT. NCT A performance was similar between healthy Germans and Turks. Looking at the PSE Sum Scores PHES, the mean Sum Scores were significantly lower (i.e. worse) in the Turkish population. This is interesting as out of the 6 parameters assessed in the PHES test battery, in three (NCT-B, LTT time and LTT error) Turks did actually perform better than Germans and in only two (DST and SDT) did Germans perform better than Turks. It may suggest that the differential power of the DST and SDT in the overall PHES test battery exceeds the other tests. This is supported by our finding that DST and SDT combined displayed a very high AUROC (0.974) for predicting mHE when compared to the PHES total score. When the DST was used alone it displayed again a quite high AUROC of 0.925 for predicting mHE. These data suggest that using DST and SDT instead of the whole battery can be considered, as one of the perceived disadvantages of the PHES test battery is the time needed to complete the full battery of tests (30 minutes on average). A simpler psychometric test would certainly be an important adjunct for the assessment of mHE. Several attempts were made in this regard [37, 38]. Our observation related to the DST and SDT is actually in line with the study by Riggio et al who used a triple test battery consisting of the DST, SDT and LT tests and suggested that this triple test was as 'good' as the PHES test [37]. This triple test reached an AUROC of 0.981 in the current study and our data suggest that the binary test serves the purpose of diagnosing mHE good enough. It must be considered, however, that depending on the cut-off used for a sum score given by these two subtests of the PHES battery, there is considerable number of false negative or false positive results, respectively, compared to the results of the whole battery (cut-off < -1 : false negative: 2.7%; false positive: 13.8%; cut-off: < -3 : false negative: 21.6% , false positive: 2.3%). The use of a cut-off with a very low false negativity rate such as the -1,5 cut-off of the DST and SDT combined test can increase the use of these tests in routine clinical practice and could have important clinical implications. This could be important in clinical research and would serve well for use in a busy Hepatology clinic.

For the assessment of mHE the cutoff value for the PHES was set at ≤ -5 in line with studies conducted in Germany, Spain, Poland, South Korea and Mexico [17, 26-29]. The only exception is Amodio's study from Italy in which the cutoff value was set at ≤ -4 [25]. The mean overall score of the cirrhotic patient group was -3.54 ± 3.98 in our cohort. This score was similar to German (-3.5 ± 4.7), but lower than in Spanish (-1.4 ± 3.4) and Italian (-1.5 ± 3.7) studies. In a study from Italy the proportion of patients with only primary school education was 35% compared to the 57% observed in our study [25]. Age and education should not affect the test results in our patients since the norms are adjusted for age and education. However, it has been shown before that patients with less educational training and patients of older age are more frequently diagnosed with mHE even using adjusted norms. This finding has been interpreted as result of a lower cognitive reserve in these subjects.

MHE prevalence of cirrhotic patients was 29.8% according to the PHES test battery. This rate changes between 20% and 84% in different studies [1]. The results from Spain, South Korea and Portugal are similar to our study (30.7%, 25.6% and 34%, respectively) [26, 28, 39]. The differences may be due to the variability of the studied patient cohorts. Supporting this, a rather high mHE prevalence according to PHES test (48%) was reported by Dhiman et al in a patient cohort that included Child-Pugh class C patients [32].

We also used the CFF test for diagnosing mHE in cirrhotic patients. The CFF test has been reported to be able to reflect both retinal functioning and cerebral cortex functions. CFF assessment appears to be a reliable, simple, convenient test and test performance is reported not to be affected by variables such as age or educational background [30]. The latter is supported by results of the current study which has used a large cohort of healthy subjects with different educational background, and this property of the CFF test provides an important advantage for its use for mHE diagnosis. However, age significantly affected CFF test results in the current study in line with several other studies [21, 31, 32]. Nevertheless, we used the fixed cut-off (39 Hz) suggested by Kircheis et al to evaluate the CFF results to be able to compare our results with previous studies [21].

The present study found the mHE rate to be 27% with the CFF test. Similarly mHE was 21% in the study by Dhiman et al [32]. However, a study conducted in Spain reported a much higher rate (42%) which is likely due to a sicker patient cohort which included Child-Pugh class C patients (about 18%) as well [30]. In addition to that, in the Spanish study about 42% of the patients had alcoholic liver cirrhosis, and alcoholism is known to affect CFF test results [21, 30].

An important question is the concordance between PHES test battery and CFF tests. The Cohen's kappa coefficient of concordance was 0.389 consistent with a significant albeit poor concordance. Several reasons may account for this such as a lower sensitivity of the CFF test or focus on a different characteristic of HE than the PHES test. Thus, any characteristic considered to indicate HE with one method might not be captured by the other method. Out of 37 patients diagnosed as having mHE according to the PHES test battery, 17 patients were not detected by the CFF test. This difference is high and suggests that the CFF test fails to detect mHE with sufficiently high precision as has been recently reported in a meta-analysis [40].

In conclusion, the present study provides Turkish norms for the PHES test battery. The PHES test results varied, in line with previous data according to age and educational level. The EASL-AASLD guidelines recommend the use of two different methods for detection of mHE. This was done in the current study. Abnormal results were found in 29.8% using the PHES test, in 27.4% using the CFF test but in only 16% with both methods. The use of 2 methods for diagnosing mHE increases specificity and may be important for research purposes. From a clinical perspective however, it makes sense to diagnose mHE early, treat it early as mHE is not associated with a good prognosis. Through its treatment we may increase patients' quality of life, prevent early loss of workforce, as well as likely accidents and deaths. In this context for the diagnosis of mHE, for the practicing clinician sensitivity matters more

than specificity. The use of a simple paper pencil test can serve that purpose and based on the data presented here we think that the combined use of the DST and SDT is well suited for this.

Declarations

Conflict of interest: The authors declare no competing interests.

Declaration of funding interests: The Critical Flicker Frequency device was kindly provided by ASSOS Pharmaceuticals throughout the study. We did not receive any financial support for conducting the study.

Author contributions: CY, FG and KW contributed to the study design. BO, OK, IK, CO, CK, AC, CO, FOO, BKÖ contributed to acquisition of data, BO, HH, FG, KW and CY contributed to data analysis, interpretation of data, drafting of the manuscript, BO, OK, HH, GG, MA, FG, OO, KW, CY contributed to revising of the manuscript. All authors have seen and approved the last version of the manuscript.

References

1. Vilstrup H, Amodio P, Bajaj J, et al. *Hepatic encephalopathy in chronic liver disease: 2014 Practice Guideline by the American Association for the Study of Liver Diseases and the European Association for the Study of the Liver*. Hepatology, 2014. **60**(2): p. 715-35.
2. Ridola L, Cardinale V, Riggio O. *The burden of minimal hepatic encephalopathy: from diagnosis to therapeutic strategies*. Ann Gastroenterol, 2018. **31**(2): p. 151-164.
3. Morgan MY, Amodio P, Cook NA, et al. *Qualifying and quantifying minimal hepatic encephalopathy*. Metab Brain Dis, 2016. **31**(6): p. 1217-1229.
4. Groeneweg M, Quero JC, De Bruijn I, et al. *Subclinical hepatic encephalopathy impairs daily functioning*. Hepatology, 1998. **28**(1): p. 45-9.
5. Agrawal S, Umapathy S, Dhiman RK. *Minimal hepatic encephalopathy impairs quality of life*. J Clin Exp Hepatol, 2015. **5**(Suppl 1): p. S42-8.
6. Schomerus H, Hamster W. *Quality of life in cirrhotics with minimal hepatic encephalopathy*. Metab Brain Dis, 2001. **16**(1-2): p. 37-41.
7. Wein C, Koch H, Popp B, Oehler G, Schauder P. *Minimal hepatic encephalopathy impairs fitness to drive*. Hepatology, 2004. **39**(3): p. 739-45.
8. Bajaj JS, Ananthakrishnan AN, McGinley EL, Hoffman RG, Brasel KJ. *Deleterious effect of cirrhosis on outcomes after motor vehicle crashes using the nationwide inpatient sample*. Am J Gastroenterol, 2008. **103**(7): p. 1674-81.
9. Kircheis G, Knoche A, Hilger N, et al. *Hepatic encephalopathy and fitness to drive*. Gastroenterology, 2009. **137**(5): p. 1706-15 e1-9.
10. Das A, Dhiman RK, Saraswat VA, Verma M, Naik SR. *Prevalence and natural history of subclinical hepatic encephalopathy in cirrhosis*. J Gastroenterol Hepatol, 2001. **16**(5): p. 531-5.
11. Romero-Gomez M, Boza F, Garcia-Valdecasas MS, Garcia E, Aguilar-Reina J. *Subclinical hepatic encephalopathy predicts the development of overt hepatic encephalopathy*. Am J Gastroenterol, 2001. **96**(9): p. 2718-23.
12. Amodio P, Del Piccolo F, Marchetti P, et al. *Clinical features and survival of cirrhotic patients with subclinical cognitive alterations detected by the number connection test and computerized psychometric tests*. Hepatology, 1999. **29**(6): p. 1662-7.

13. 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**(6): p. 1483-9.
14. Quero JC, Hartmann IJ, Meulstee J, Hop WC, Schalm SW. *The diagnosis of subclinical hepatic encephalopathy in patients with cirrhosis using neuropsychological tests and automated electroencephalogram analysis*. Hepatology, 1996. **24**(3): p. 556-60.
15. Groeneweg M, Moerland W, Quero JC, Hop WC, Krabbe PF, Schalm SW. *Screening of subclinical hepatic encephalopathy*. J Hepatol, 2000. **32**(5): p. 748-53.
16. Lee Y, Kim C, Suk KT, et al. *Differences in cognitive function between patients with viral and alcoholic compensated liver cirrhosis*. Metab Brain Dis, 2016. **31**(2): p. 369-76.
17. Weissenborn K, Ennen JC, Schomerus H, Ruckert N, Hecker H. *Neuropsychological characterization of hepatic encephalopathy*. J Hepatol, 2001. **34**(5): p. 768-73.
18. Lauridsen MM, Thiele M, Kimer N, Vilstrup H. *The continuous reaction times method for diagnosing, grading, and monitoring minimal/covert hepatic encephalopathy*. Metab Brain Dis, 2013. **28**(2): p. 231-4.
19. Bajaj JS, Hafeezullah M, Franco J, et al. *Inhibitory control test for the diagnosis of minimal hepatic encephalopathy*. Gastroenterology, 2008. **135**(5): p. 1591-1600 e1.
20. Bajaj JS, Thacker LR, Heuman D, et al. *The Stroop smartphone application is a short and valid method to screen for minimal hepatic encephalopathy*. Hepatology, 2013. **58**(3): p. 1122-32.
21. Kircheis G, Wettstein M, Timmermann L, Schnitzler A, Haussinger D. *Critical flicker frequency for quantification of low-grade hepatic encephalopathy*. Hepatology, 2002. **35**(2): p. 357-66.
22. Bajaj JS, Cordoba J, Mullen KD, et al. *Review article: the design of clinical trials in hepatic encephalopathy—an International Society for Hepatic Encephalopathy and Nitrogen Metabolism (ISHEN) consensus statement*. Aliment Pharmacol Ther, 2011. **33**(7): p. 739-47.
23. Schomerus H, Hamster W. *Neuropsychological aspects of portal-systemic encephalopathy*. Metab Brain Dis, 1998. **13**(4): p. 361-77.
24. Ferenci P, Lockwood A, Mullen K, Tarter R, Weissenborn K, Blei AT. *Hepatic encephalopathy—definition, nomenclature, diagnosis, and quantification: final report of the working party at the 11th World Congresses of Gastroenterology, Vienna, 1998*. Hepatology, 2002. **35**(3): p. 716-21.
25. Amodio P, Campagna F, Olianias 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**(3): p. 346-53.
26. 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**(7): p. 246-9.
27. Wunsch E, Koziarska D, Kotarska K, Nowacki P, Milkiewicz P. *Normalization of the psychometric hepatic encephalopathy score in Polish population. A prospective, quantified electroencephalography study*. Liver Int, 2013. **33**(9): p. 1332-40.
28. Seo YS, Yim SY, Jung JY, et al. *Psychometric hepatic encephalopathy score for the detection of minimal hepatic encephalopathy in Korean patients with liver cirrhosis*. J Gastroenterol Hepatol, 2012. **27**(11): p. 1695-704.
29. Duarte-Rojo A, Estradas J, Hernandez-Ramos R, Ponce-de-Leon S, Cordoba J, Torre A. *Validation of the psychometric hepatic encephalopathy score (PHES) for identifying patients with minimal hepatic encephalopathy*. Dig Dis Sci, 2011. **56**(10): p. 3014-23.
30. Romero-Gomez M, Cordoba J, Jover R, et al. *Value of the critical flicker frequency in patients with minimal hepatic encephalopathy*. Hepatology, 2007. **45**(4): p. 879-85.

31. Goldbecker A, Weissenborn K, Hamidi Shahrezaei G, et al. *Comparison of the most favoured methods for the diagnosis of hepatic encephalopathy in liver transplantation candidates*. Gut, 2013. **62**(10): p. 1497-504.
32. Dhiman RK, Kurmi R, Thumburu KK, et al. *Diagnosis and prognostic significance of minimal hepatic encephalopathy in patients with cirrhosis of liver*. Dig Dis Sci, 2010. **55**(8): p. 2381-90.
33. American Association for the Study of Liver, D. and L. European Association for the Study of the, *Hepatic encephalopathy in chronic liver disease: 2014 practice guideline by the European Association for the Study of the Liver and the American Association for the Study of Liver Diseases*. J Hepatol, 2014. **61**(3): p. 642-59.
34. Kircheis G, Hilger N, Haussinger D. *Value of critical flicker frequency and psychometric hepatic encephalopathy score in diagnosis of low-grade hepatic encephalopathy*. Gastroenterology, 2014. **146**(4): p. 961-9.
35. Hanley JA, McNeil BJ. *A method of comparing the areas under receiver operating characteristic curves derived from the same cases*. Radiology, 1983. **148**(3): p. 839-43.
36. Schomerus H, Weissendorn K, Hamster W, RuĖckert N, Hecker H. *PSE-Syndrom-Test Swets Test Services*. Frankfurt: Swets & Zeitlinger, 1999.
37. Riggio O, Ridola L, Pasquale C, et al. *A simplified psychometric evaluation for the diagnosis of minimal hepatic encephalopathy*. Clin Gastroenterol Hepatol, 2011. **9**(7): p. 613-6 e1.
38. Campagna F, Montagnese S, Ridola L, et al. *The animal naming test: An easy tool for the assessment of hepatic encephalopathy*. Hepatology, 2017. **66**(1): p. 198-208.
39. Barbosa M, Marinho C, Mota P, Cotter J. *Minimal Hepatic Encephalopathy: The Reality Beyond Our Eyes*. Acta Med Port, 2015. **28**(4): p. 480-5.
40. Torlot FJ, McPhail MJ, Taylor-Robinson SD. *Meta-analysis: The diagnostic accuracy of critical flicker frequency in minimal hepatic encephalopathy*. Aliment Pharmacol Ther, 2013. **37**(5): p. 527-36.

Tables

Table 1. Characteristics of the regression lines in the transformed scale; comparison between the Turkish and Hannover data

Subtest		Transformation / Regression parameter	Data Source		Comparison Hannover vs. Turkish Data (P-Value) ¹⁾	
Turkish Data	PSE Manual		Hannover (N = 249)	Hannover (N = 120)	Turkish (N = 816)	
Transformation / Parameter values						
NCTA	ZV A	Transformation:	log		Log	
Number connection test (A)	Connecting Numbers (A) (Time)	a (constant):	2.7157		2.7807	n.s.
		b (regression slope):	0.0167		0.01442	n.s.
		s (standard deviation):	Sqrt(0.0768) = 0.2771		sqrt (0.11268) = 0.33568	0.004
		Outlier:	none		Nr. 133, 386, 391	
NCTB	ZV B	Transformation:	log		log-log	
Number connection test (B)	Connecting numbers (B) (Time)	a (constant):	3.5503		1.28512	0.006
		b (regression slope):	0.0183		0.0034026	
		s (standard deviation):	Sqrt(0.0816)= 0.2856		sqrt (0.008161)= 0.090338	
		Outlier:	none		Nr. 1, 15, 124, 625.	
LTT time	LNZ	Transformation:		log	log-log	
Line Tracing test (time)	Line tracing (Time)	a (constant):		3.7984	1.27285	0.001
		b (regression slope):		0.0067	0.002396	
		s (standard deviation):		Sqrt(0.0859) = 0.2931	sqrt (0.00601) = 0.07752	
		Outlier:		none	Nr 102, 485, 599, 663	
LTT Error	LNF	Transformation		SquareRoot	Cube Root	
Line Tracing test (errors)	Line Tracing (Error)	a (constant):		4.9617	2.2333	< 0.001
		b (regression slope):		0.0283	0.0031	
		s (standard deviation):		Sqrt(3.8811)	sqrt (0.74994) = 0.8660	

			=1.9701		
		outlier	none	Nr. 102,560	
DST	ZS	Transformation:	log	SquareRoot	
Digit Symbol test	Counting Symbol Test (Number)	a (constant):	4.3403	7.9854	< 0.001
		b (regression slope):	-0.0101	- 0.03674	
		s (standard deviation):	Sqrt(0.0356) =0.1887	sqrt(0.96138)= 0.98050	
		Outlier:	none	Nr. 479, 551, 569	
SDOT	KP	Transformation:	log	log-log-log	
Serial Dotting test	Dotting circles (Time)	a (constant)	3.5224	0.23604	< 0.001
		b (regression slope):	0.0047	0.001275	
		s (standard deviation):	Sqrt(0.0261) =0.1615	s = sqrt(0.0019954) =0.04467	
		Outlier:	none	Nr. 133, 244, 264, 740, 754	

Table 2. Receiver operator characteristic (ROC) analyses of PHES subtests for predicting minimal hepatic encephalopathy (mHE) through comparison to the PHES total score.

PHES subtest	AUROC	Cut-off	Sensitivity	MHE (n)	Specificity	PPV	NPV
NCT-A	0.894	-0.5	33/37 (89.2)	52	68/87 (78.2)	33/52 (63.5)	68/72 (94.4)
NCT-B	0.818	-1.5	23/37 (62.2)	32	78/87 (89.7)	23/32 (71.9)	78/92 (84.8)
LTT-Time	0.777	-0.5	22/37 (59.5)	33	76/87 (87.4)	22/33 (66.7)	79/91 (83.5)
LTT-Error	0.664	-0.5	19/37 (51.4)	40	66/87 (75.9)	19/40 (47.5)	66/84 (78.6)
SDT	0.876	-0.5	34/37 (91.9)	59	62/87 (71.3)	34/59 (57.6)	62/65 (95.4)
DST	0.925	-0.5	35/37 (94.6)	53	69/87 (79.3)	35/53 (66)	69/71 (97.2)
DST + NCT-A	0.951	-1.5	31/37 (83.8)	41	77/87 (88.5)	31/41 (75.6)	77/83 (92.8)
DST + NCT-B	0.920	-1.5	32/37 (86.5)	50	69/87 (79.3)	32/50 (64)	69/74 (93.2)
DST + SDT	0.974	-1.5	36/37 (97.3)	48	75/87 (86.2)	36/48 (75)	75/76 (98.7)
DST + SDT+ NCT-A	0.989	-2.5	35/37 (94.6)	39	83/87 (95.4)	35/39 (89.7)	83/85 (97.6)
DST + SDT+ NCT-B	0.988	-2.5	36/37 (97.3)	47	76/87 (87.4)	36/47 (76.6)	76/77 (98.7)
DST + SDT + LTT-Time	0.981	-2.5	34/37 (91.9)	38	83/87 (95.4)	34/38 (89.5)	83/86 (96.5)

Abbreviations: AUROC: area under ROC curve, PHES: Psychometric hepatic encephalopathy score, NCT: Number Connection Test, LTT: Line Tracing Test, SDT: Serial Dotting Test, DST: Digit Symbol Test

Table 3. Comparison of receiver operator characteristics (ROC) curves of subtests of the PHES test battery for diagnosing mHE in relation to the PHES test sum score

PHES subtest 1	AUROC	PHES subtest 2	AUROC	p value
DST	0.925	NCT A	0.894	0.467
DST	0.925	NCT B	0.818	0.035
DST	0.925	SDT	0.876	0.259
DST	0.925	LTT time	0.777	0.006
DST	0.925	LTT error	0.664	<0.0001
DST + SDT	0.974	DST + NCT A	0.951	0.269
DST + SDT	0.974	DST + NCT B	0.920	0.059
DST + SDT	0.974	DST + SDT + NCT A	0.988	0.332
DST + SDT	0.974	DST	0.925	0.085
DST + SDT	0.974	SDT	0.876	0.007
DST + SDT	0.974	NCT A	0.894	0.023
DST + SDT	0.974	NCT B	0.818	0.0005
DST + SDT	0.974	LTT time	0.777	<0.0001
DST + SDT	0.974	LTT error	0.664	<0.0001

Abbreviations: AUROC: area under ROC curve, PHES: Psychometric hepatic encephalopathy score, NCT: Number Connection Test, LTT: Line Tracing Test, SDT: Serial Dotting Test, DST: Digit Symbol Test

Figures

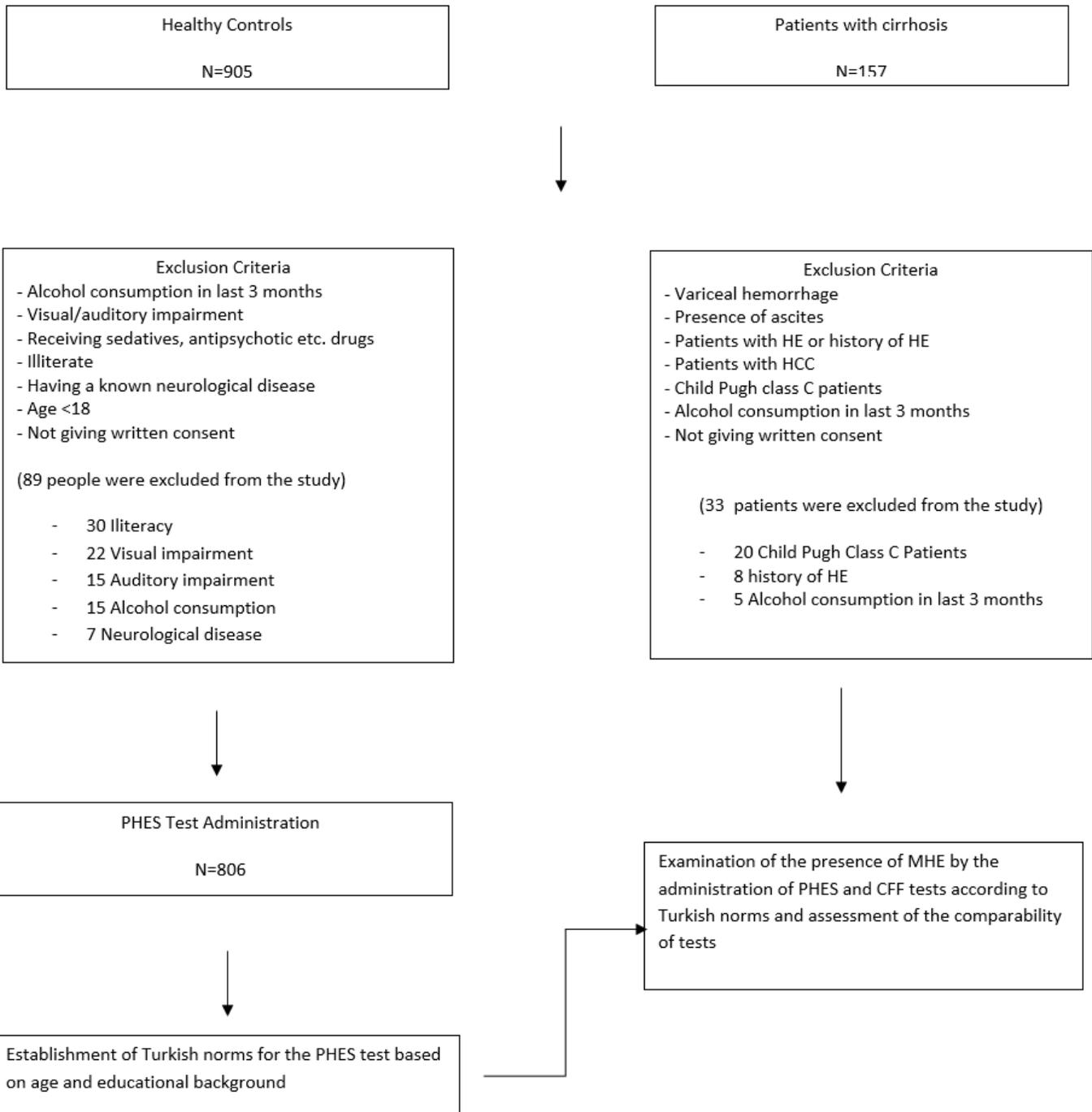


Figure 1

Flow diagram of individuals who participated in the establishment of Turkish norms for the psychometric hepatic encephalopathy score (HE = Hepatic Encephalopathy, HCC= Hepatocellular Carcinoma, PHEs = psychometric hepatic encephalopathy score)

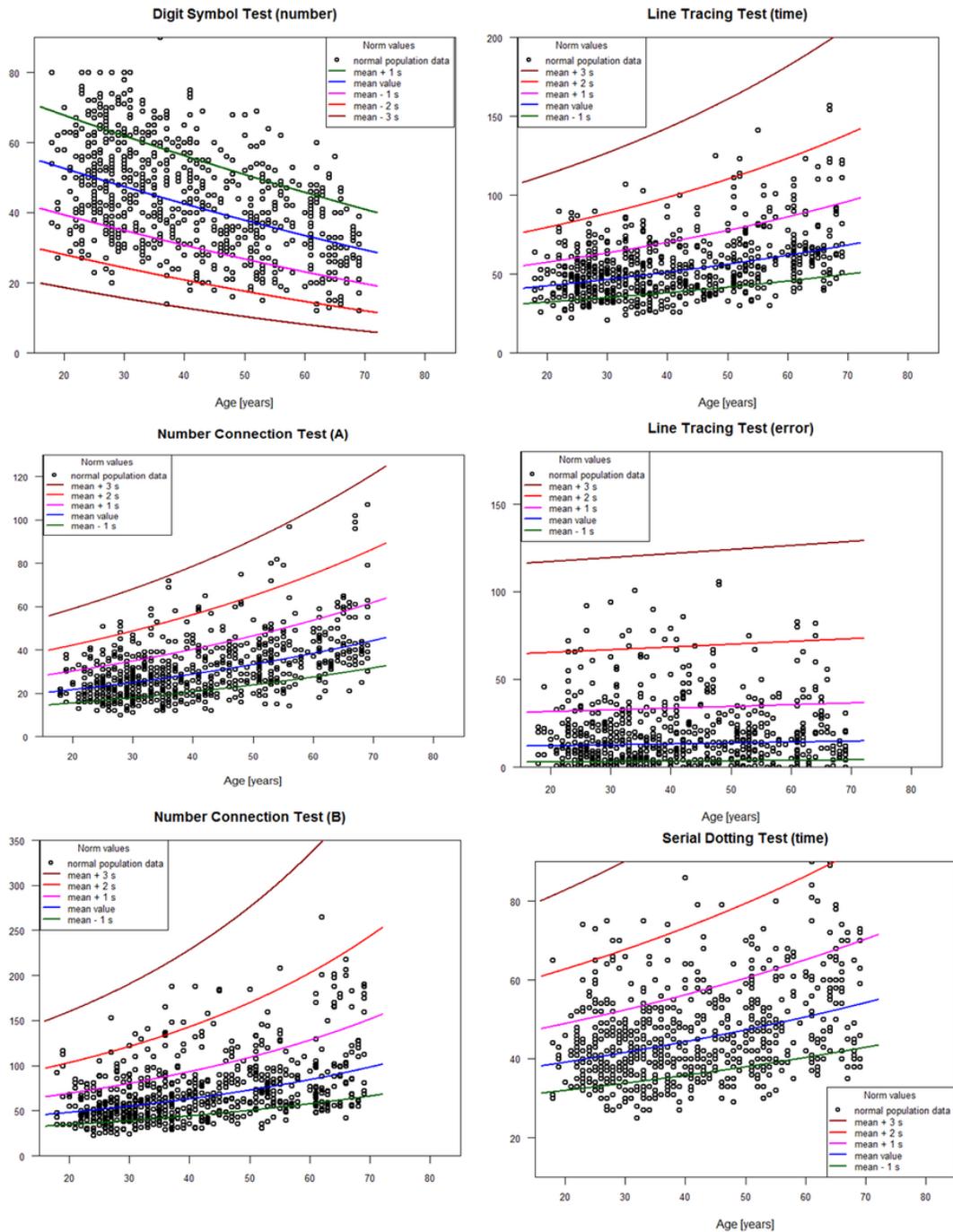


Figure 2

Norms obtained for the various subtests of the portosystemic hepatic encephalopathy score based on performance of 816 adult Turkish healthy subjects. Mean with standard deviation values are given according to age for the Number Connection test A (2A), Number Connection test B (2B), Line Tracing test (time) (2C), Line Tracing test (error) (2D), Digit Symbol test (2E) and Serial Dotting test (2F)

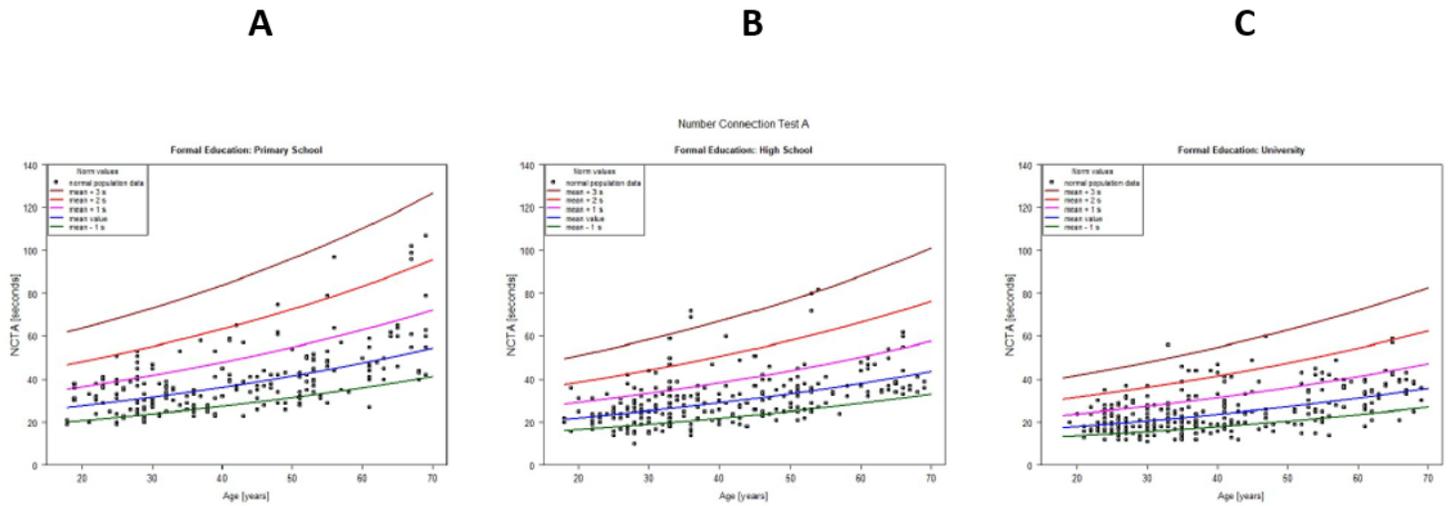


Figure 3

Effect of education on test performance exemplified here for the Number Connection test-A. Test performance improves at every age level with higher education: data is shown for primary school (A), high school (B) and university education (C)

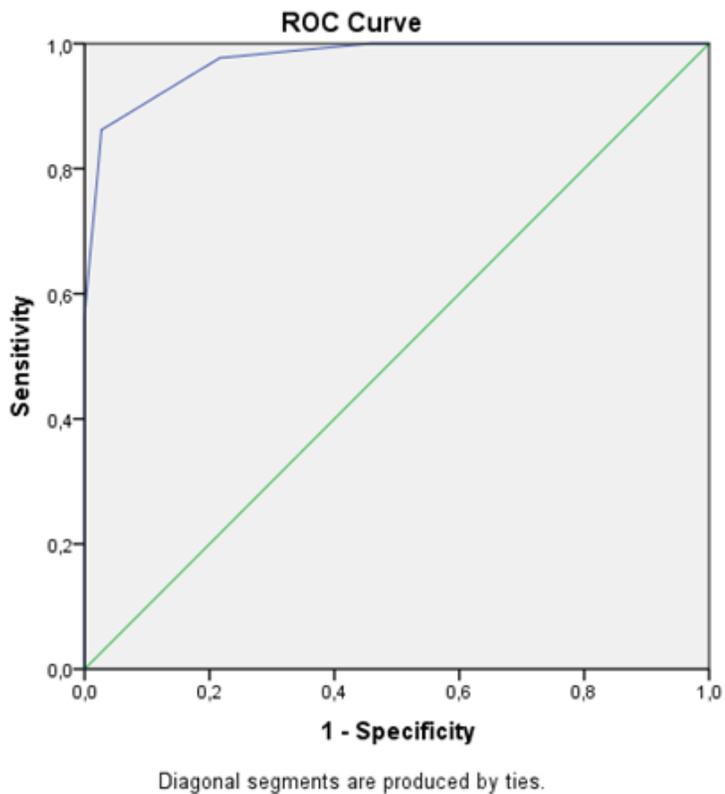


Figure 4

Performance of the digit symbol and serial dotting tests according to the portosystemic hepatic encephalopathy score using receiver operator characteristic curve analysis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure1.docx](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable10.docx](#)
- [SupplementaryTable2.docx](#)
- [SupplementaryTable3AF.docx](#)
- [SupplementaryTable4.docx](#)
- [SupplementaryTable5.docx](#)
- [SupplementaryTable6.docx](#)
- [SupplementaryTable7.docx](#)
- [SupplementaryTable8.docx](#)
- [SupplementaryTable9.docx](#)