ELSEVIER

Journal of Hepatology 49 (2008) 308–312

Journal of Hepatology

www.elsevier.com/locate/jhep

Editorial

PHES: One label, different goods?! ☆

Karin Weissenborn*

Department of Neurology, Hannover Medical School, 30623 Hannover, Germany

See Article, pages 346–353

Psychometric testing has a long tradition in clinical hepatology. It has been used for decades for the assessment and follow-up of hepatic encephalopathy (HE). Initially hand-writing, or the construction of a five-pointed star were common methods to monitor HE. But, since the evaluation of these tests could not be standardized, alternative solutions were sought which combined convenience, practicability, sensitivity, specificity, and low costs in terms of money and time.

A first approximation towards these goals was the introduction of the Number Connection Tests into the assessment of HE [1]. Then comprehensive batteries of psychometric tests were used, and it was shown that HE is characterized by a distinct pattern of cognitive impairment: alterations of attention, visuo-spatial perception and psychomotor function [2–5]. While the verbal IQ was preserved in patients with overt hepatic encephalopathy the performance IQ was found to be decreased even in patients without clinical signs of HE [2,3].

The PHES – the psychometric hepatic encephalopathy score – has its roots in these flourishing times of neuropsychological assessment of HE. The term PHES was coined by Dr. Andres Blei in 2001. He suggested to name the sum score of the PSE-Syndrom-Test [6], a test battery which had been especially developed for diagnosing HE, psychometric hepatic encephalopathy score (PHES) [7]. The basis for the PSE-Syndrom-Test had been laid by the neuropsychologist Wolfgang Hamster† and the gastroenterologist Hans Schomerus† in the early eighties. They aimed to characterize the neuropsychological deficit typical for HE, and to develop a short battery that could be easily used in the clinic to make the diagnosis of minimal HE. Therefore, they presented a battery of more than 20 neuropsychological tests to patients with liver cirrhosis, patients with alcohol-toxic pancreatitis without liver cirrhosis, patients with alcohol-toxic neuropathy without cirrhosis and healthy controls [8,9]. They found deficits in concentration, attention and psychomotor function in the liver patients compared to controls. A discriminant analysis was performed to identify those variables which were capable of differentiating between patients with and without liver cirrhosis and between patients with HE and patients with alcoholic brain atrophy and healthy controls. The Digit Symbol Test, the Benton Test, the Line tracing Test (in particular the number of errors), reaction time to acoustic and visual stimuli, and the tests steadiness, aiming and long sticks from the ‘‘Motorische Leistungsserie’’, a battery aimed to detect alterations of fine motor skills, showed a high discrimination accuracy. For practicability reasons they decided to develop a paper-pencil-test battery that represented the affected domains such as visual-motor coordination, attention shift and motor speed and accuracy. Considering the results of the discriminant analysis the Digit Symbol Test (DST), a paper-pencil version of the Line Tracing Test (LTT) and instead of aiming the Serial Dotting Test (SDT) were included into the battery. Then the Number Connection Tests (NCT) A and B (also referred to as Trailmaking Tests) were added because these two tests were the most frequently used psychometric tests for the diagnosis of minimal HE at that time. The new battery was presented to more than 400 individuals to develop normative data [10].

Associate Editor: J. Bosch

☆ The author who has taken part in the research of this paper declared that she does not have a relationship with the manufacturers of the materials involved either in the past or present and did not receive funding from the manufacturers to carry out her research.

* Tel.: +49 511 532 2339; fax: +49 511 532 3115.
E-mail address: weissenborn.karin@mh-hannover.de

0168-8278/$34.00
doi:10.1016/j.jhep.2008.06.023

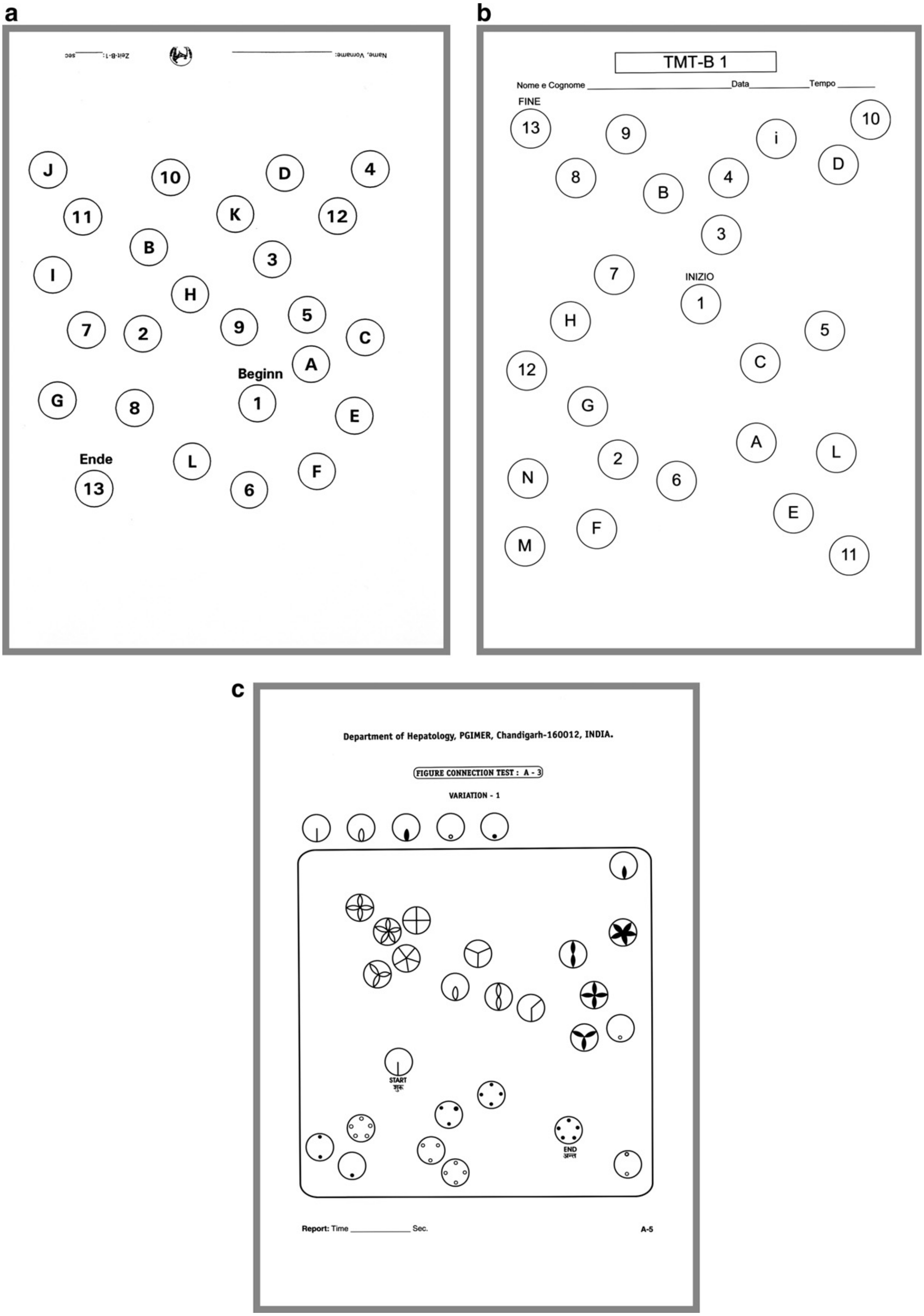

Fig. 1. Examples of the German (a) and Italian (b) versions of the NCT B and of the Figure Connection Test which substitutes the NCT B in the Indian version of the test battery. The Italian version of the NCT B was provided by Dr. Piero Amodio, Clinical and Experimental Medicine; University of Padova. The example of the Figure Connection Test (c) was provided by Dr. Radha K. Dhiman, Additional Professor, Department of Hepatology, Postgraduate Institute of Medical Education & Research, Chandigarh, India.

Most of the tested people were in-patients of the Department of Internal Medicine, part were forced to undergo a neuropsychological examination to keep their driving licence. Unfortunately, the age pattern of this "representative group" was disproportionate. Therefore, in a second step the battery was presented to a group of 120 healthy controls with a balanced pattern of age, education and occupation [11]. Finally, in 1999, the battery was published by SWETS Test Services in German.

Hamster and Schomerus made several attempts to publish their data regarding an optimal test combination for the diagnosis of HE. At the end of the eighties, however, psychometric testing – especially the application of paper-pencil-tests – had become obsolete. More "sophisticated" technical methods were expected to produce a higher sensitivity and a higher specificity than paper-pencil-tests [12–14].

Drs. Hamster and Schomerus would be highly pleased to see their test battery finally being appreciated all over the world. After the standardization in Germany similar test batteries have also been standardized in Spain [15] and, now, in Italy [16] and the United Kingdom [17]. An altered test version has been standardized in India [18].

All test versions used so far are similar with regard to their principal structure. All combine a digit symbol test, a serial dotting test, a line tracing test and number or figure connection tests. Nevertheless, there are significant differences in details that make it difficult to compare the results achieved with the different test versions. Amodio et al. [16], for example, used for their battery presented in this issue of journal the DST, SDT and LTT forms of the PSE-Syndrom-Test and added previously developed forms for the NCT A and B because of differences between the Italian and German alphabet sequences. In addition to the distribution of numbers and letters the dimension of circles and letters in these tests were different from the German forms. The Indian group substituted the NCT B with a test where the subjects had to connect figures belonging to a specific cate-

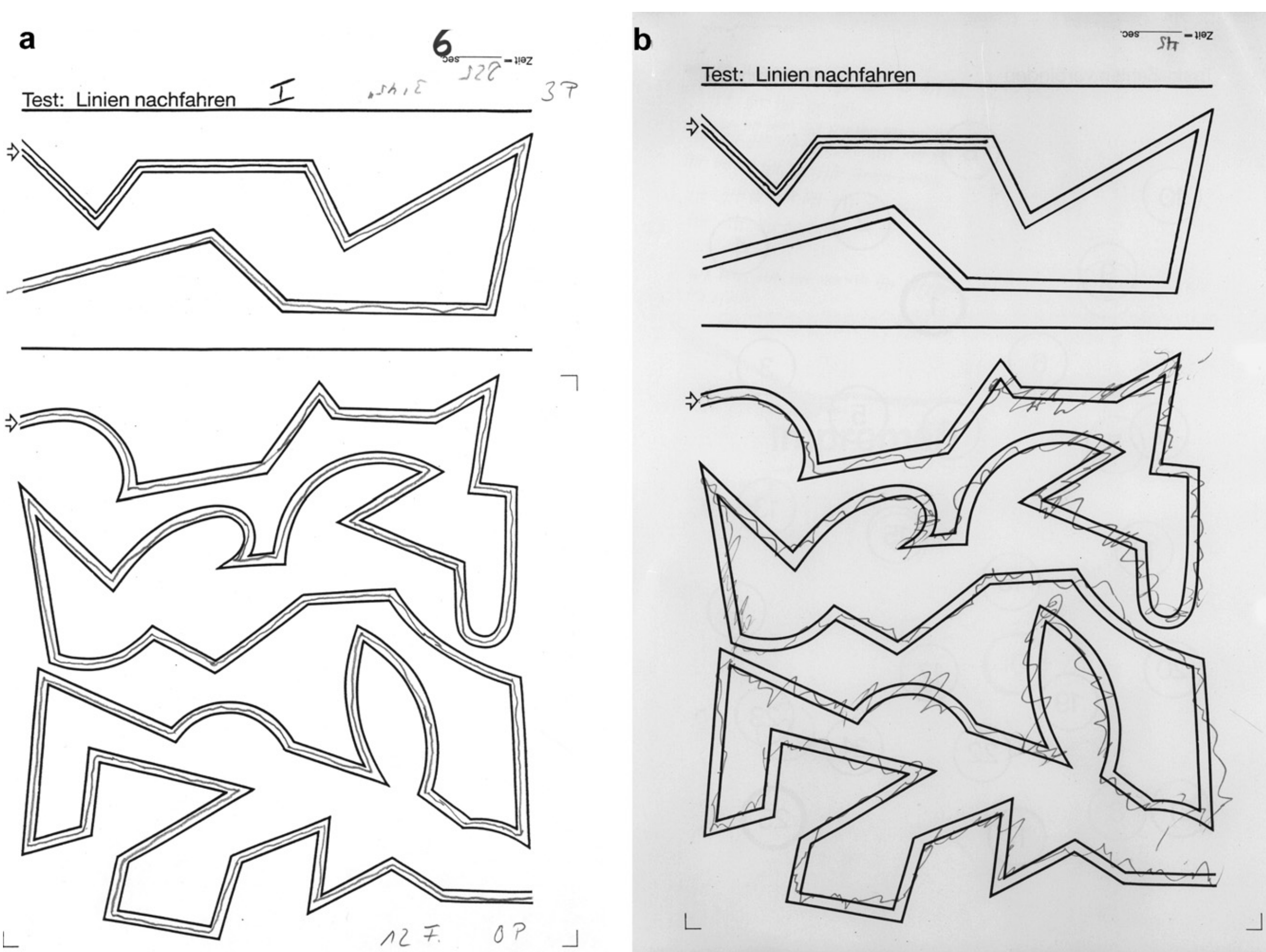

Fig. 2. Apart from cognitive dysfunction hepatic encephalopathy may also result in a reduction of motor speed or motor accuracy or even both. The LTT example on the left (a) shows the result of a patient with extreme bradykinesia but preserved accuracy in fine movements (LTT time: 225 s, number of errors: 12). The example on the right shows the LTT result of a patient without alterations of cognition or motor speed but extensive asterixis (LTT time: 45 s, number of errors: 198).

gory because they have to deal with a high amount of non-alphabetized patients [19]. Both alterations of the test (Fig. 1) have measurable influence on the test performance. The Indian solution even alters the test structure as their substitute for the NCT B assesses other cognitive functions than the NCT B (Table 1).

Besides differences in the structure, there are also marked differences in the evaluation procedure of the tests. All groups compare the subject's sub-test results to norms and score them with +1 to −3 points depending on their position in the +1 to −3 standard deviation range from the mean. Then a sum score is calculated which is referred to as Psychometric Hepatic Encephalopathy Score (PHES). Since all groups use the term PHES for the sum score of their test variation the difference between the batteries tends to be forgotten and PHES results from different study groups are compared without reservations. However, there are several very important differences in the calculation of the sum score. They become most evident considering the scoring of the line tracing test results. Hamster and Schomerus had shown in their discriminant analysis that performance time and number of errors added independently to the discrimination between patients with liver cirrhosis and their control groups with the number of errors being even better than the performance time. Therefore, they included both, performance time and errors into their sum score [8]. Amodio et al. [16] argue that LTT performance time and errors are significantly related. Thus they measure LTT performance by "error-weighted time". Their argumentation is identical to that of the authors of the Spanish and British normal data, who also calculate a combined – though completely differently calculated – score for performance time and errors [15,17]. The argumentation is based on the data of healthy controls. It does not take into account that performance time and errors in the LTT represent different aspects of brain function, which may be altered independently by any pathophysiology (Fig. 2). Therefore, neuropsychologists suggest to assess performance speed and accuracy independently, if possible.

The different handling of the LTT results in differences with regard to the range of the sum score. While the German version provides a range from +6 to −18 the Italian, Spanish and British versions provide a range from +5 to −15. Considering all changes of test structure and evaluation it becomes obvious that the cut-off score between normal and pathological results must be different between the different test versions, and must be determined for each of the different versions, separately. Since the test shall discriminate between normal brain function and hepatic encephalopathy the cut-off score should be determined by comparison of healthy controls and patients with clinically overt HE. Amodio et al. [16] give reasons for their cut-off score of $\leqslant -4$ by comparing their psychometric results with the EEG results as they had not examined patients with clinically overt HE. This approach is comprehensible. However, like all other substitute measures EEG is not able to fully represent HE. Even patients with clinically overt HE may present with normal EEG [13].

The paper by Amodio et al. [16] underscores the principal usefulness of the PSE-Syndrom-Test test battery for diagnosing minimal HE. However, it also highlights facts that have to be considered in psychometric testing principally: the influence of age, education, occupation and socio-cultural background and practice effects (see Table 2).

Table 1
Cognitive and motor functions addressed by the different components of the PHES – delivering test batteries

Sub-test	Assessable functions
Number Connection Test A	Psychomotor speed; visual scanning efficiency, sequencing, attention, concentration
Number Connection Test B	Attention set shifting ability, psychomotor speed, visual scanning efficiency, sequencing, attention, concentration
Figure Connection Test [19]	Visual perception, visual search, visual scanning efficiency, psychomotor speed, attention, concentration, working memory
Digit Symbol Test	Associative learning; graphomotor speed, cognitive processing speed, visual perception, working memory
Serial Dotting Test	Motor speed
Line Tracing Test	Motor speed and accuracy

Table 2
Advantages and limitations of the psychometric test batteries that have been standardized for the assessment of minimal hepatic encephalopathy [6,15–18]

Advantages	Limitations
• High sensitivity • High specificity • High reliability • High validity • Simplicity • Bedside tests • Low cost	• Need for representative norm data for each sample studied considering effects of age, education, occupation, gender and socio-cultural background • Non-applicability of the NCT B in illiterate subjects • Need for the control of practice effects

Importantly, normal data that have been elaborated in a representative sample from one country cannot be used for the evaluation of individual data in other countries without control. Due to alterations in educational levels and daily living activities from one generation to the other, neuropsychologists recommend the periodic re-evaluation of normal data even within a distinct population.

How come, despite of all these drawbacks, the psychometric hepatic encephalopathy score (PHES) based on one or the other variation of the original PSE-Syndrom-Test is increasingly used for diagnosing minimal HE, worldwide? The battery is practical and easy to apply, sensitive and specific, and cheap. Practice effects in the range as in the present study by Amodio et al. [16] have not been observed by other groups [11]. None of the neuro- or psychophysiological methods used competitively has proven to be of greater use for diagnosing mHE, so far. This holds especially true also for the most recently recommended measure – the critical flicker frequency [20]. It is foreseeable that the norm differences between the different European countries will even out within the next decade. It appears worthwhile to directly compare meanwhile the raw data of the different representative samples of different countries, and to see whether the results would be more comparable after the deletion of one or the other sub-test, or after identical handling of the raw data. At best, this combined with a summarisation of all norm data collected to date would result in more widely applicable norms, which, for example, could be used for the assessment of HE in international multi-centre trials.

References

[1] Zeegen R, Drinkwater JE, Dawson AM. Method for measuring cerebral dysfunction in patients with liver disease. Brit Med J 1970;2:633–636.

[2] Gilberstadt SJ, Gilberstadt H, Zieve L, Buegel B, Collier RO, Mc Clain J. Psychomotor performance defects in cirrhotic patients without overt encephalopathy. Arch Intern Med 1908;140:519–521.

[3] Rikkers L, Jenko P, Rudman D, Freides D. Subclinical hepatic encephalopathy: detection, prevalence, and relationship to nitrogen metabolism. Gastroenterology 1978;75:462–469.

[4] Rehnström S, Simert G, Hansson JA, Johnson G, Vang J. Chronic hepatic encephalopathy. A psychometric study. Scand J Gastroenterol 1977;12:305–311.

[5] Tarter RE, Hegedus AM, Van Thiel DH, Schade RR, Gavaler JS, Starzl TE. Nonalcoholic cirrhosis associated with neuropsychological dysfunction in the absence of overt evidence of hepatic encephalopathy. Gastroenterology 1984;86:1421–1427.

[6] Schomerus H, Weissenborn K, Hamster W, Rückert N, Hecker H. PSE-Syndrom-Test. Swets Test Services. Frankfurt: Swets & Zeitlinger B.V.; 1999.

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

[8] Hamster W. Neuropsychologie der latenten portosystemischen Enzephalopathie. PhD thesis. Medizinische Fakultät der Eberhard-Karls-Universität Tübingen; 1982.

[9] Schomerus H, Hamster W. Neuropsychological aspects of portal–systemic encephalopathy. Metab Brain Dis 1998;13:361–377.

[10] Schönleber R. Standardisierung eines psychometrischen Tests zur Erfassung der hepatischen Enzephalopathie. Thesis. Medizinische Fakultät der Eberhard-Karls-Universität Tübingen; 1989.

[11] Ennen J. Diagnosestandardisierung der latenten portosystemischen Enzephalopathie mittels psychometrischer Testverfahren. Thesis. Medizinische Hochschule Hannover; 2000.

[12] Davies MG, Rowan MJ, Feely J. EEG and event related potentials in hepatic encephalopathy. Metab Brain Dis 1991;6:175–186.

[13] Weissenborn K. Neurophysiological methods in the diagnosis of early hepatic encephalopathy. In: Bengtsson F, Jeppsson B, Almdahl T, Vilstrup H, editors. Progress in hepatic encephalopathy and metabolic nitrogen exchange. Boca Raton: CRC Press; 1991. p. 27–39.

[14] Kircheis G, Wettstein M, Timmermann L, Schnitzler A, Häussinger D. Critical flicker frequency for quantification of low-grade hepatic encephalopathy. Hepatology 2002;35:357–366.

[15] Romero Gómez M, Córdoba J, Jover R, del Olmo J, Fernández A, Flavià M, 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.

[16] Amodio P, Campagna F, Olianas S, Iannizzi P, Mapelli D, Penzo M, 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.

[17] Marks ME, Jackson CD, Montagnese S, Jenkins CW, Head M, Morris RW, et al. Derivation of a normative UK database for the psychometric hepatic encephalopathy score (PHES): confounding effect of ethnicity and test scoring. J Hepatol 2008;48:S119.

[18] Thumburu KK, Kurmi R, Dhiman RK, Venkataramarao SH, Prabhakar S, Singh P, et al. Psychometric hepatic encephalopathy score, critical flicker frequency and P300 event-related potential for the diagnosis of minimal hepatic encephalopathy: evidence that psychometric hepatic encephalopathy score is enough. 13th ISHEN, Padua; 2008 [abstract].

[19] Dhiman RK, Saraswat VA, Verma M, Naik SR. Figure connection test: a universal test for assessment of mental state. J Gastroenterol Hepatol 1995;10:14–23.

[20] Romero-Gómez M, Córdoba J, Jover R, del Olmo JA, Ramirez M, Rey R, et al. Value of the critical flicker frequency in patients with minimal hepatic encephalopathy. Hepatology 2007;45:879–885.