ELSEVIER

Journal of Hepatology 47 (2007) 642–650

Journal of Hepatology

www.elsevier.com/locate/jhep

Assessment of low-grade hepatic encephalopathy: A critical analysis☆

Gerald Kircheis[1,*], Wolfgang E. Fleig[2], Roman Görtelmeyer[3,4], Susanne Grafe[3], Dieter Häussinger[1]

[1]*Department of Gastroenterology, Hepatology and Infectiology, Heinrich-Heine-University Düsseldorf, Moorenstrasse 5, 40225 Düsseldorf, Germany*
[2]*Department of Internal Medicine I, Martin-Luther-University Halle-Wittenberg, Germany*
[3]*Merz Pharmaceuticals GmbH, Frankfurt, Germany*
[4]*Department of Psychology, University of Mannheim, Germany*

***Background/Aims*: The value of paper–pencil tests and West-Haven-criteria for assessment of low-grade hepatic encephalopathy under conditions of a randomized, double-blind, placebo-controlled, clinical trial was evaluated in a cohort of 217 cirrhotics.**

***Methods*: Patients were graded at least twice clinically for severity of hepatic encephalopathy and tested concomitantly with a recommended psychometric test battery.**

***Results*: Re-evaluation of the study documentation showed that at study entry 33% and during the study even 50% of the patients were wrongly allocated to minimal or overt hepatic encephalopathy. Despite the participating physicians' training, 31% of the number-connection-tests-A, 20% of the number-connection-tests-B and 28% of the line-tracing-test were in retrospect considered invalid by an independent psychologist. Neither the Portosystemic-Encephalopathy-Syndrome (PSE) test nor the Psychometric-Hepatic-Encephalopathy-Sum (PHES)-score reliably picked up clinical improvement in the individual patient. Although these test scores could statistically differentiate between patients with minimal and overt hepatic encephalopathy, the clinical classification of individual patients into one of the groups will have a high rate of error. The PHES-Score was less balanced than the score derived from the PSE-Syndrome-Test.**

***Conclusions*: Inaccuracies in conducting paper–pencil tests together with the subjectivity and incorrectness of clinical HE-grading question the usefulness of West-Haven-criteria and paper–pencil tests including related scores for quantification of low-grade HE at least in multicenter approaches.**

***Keywords*: Cirrhosis; Minimal hepatic encephalopathy; Paper–pencil tests; Psychometry**

1. Introduction

Depending on the severity of clinically apparent neuropsychiatric symptoms, overt hepatic encephalopathy (oHE) is usually classified into four stages according to the West-Haven-criteria [1–3]. However, a significant proportion of cirrhotic patients, which appear normal on clinical examination, may exhibit neuropsychological deficits, when subjected to selected neuropsychological test batteries [4–13]. This condition was termed minimal hepatic encephalopathy (mHE), which by definition cannot be picked up by West-Haven-criteria [4–13]. In order to diagnose mHE, it is important to use several validated and reliable tests which pick up a broad range of neuropsychiatric abnormalities [13–20]. However, the value of these tests is limited with regard to retest-reliability and is affected

Received 17 January 2007; received in revised form 24 May 2007; accepted 25 May 2007; available online 23 July 2007
Associate Editor: J. Bosch

☆ RG and SG are employees of Merz Pharmaceuticals GmbH, Frankfurt, Germany. GK and DH belong to a group of patent holders for a portable bed-side device for critical-flicker-frequency analysis.

* Corresponding author. Tel.: +49 (0) 211 811 6210; fax: +49 (0) 211 811 8752.
E-mail address: Kircheis@med.uni-duesseldorf.de (G. Kircheis).

0168-8278/$32.00
doi:10.1016/j.jhep.2007.05.019

by factors such as age, training, and education of the patient [10,18,19,21–31]. Psychometric test batteries such as the Portosystemic-Encephalopathy-Syndrome-Test (PSE-Syndrome-Test;32,33) or the Psychometric-Hepatic-Encephalopathy-Sumscore (PHES-Score; 18,19, 31) have been proposed as potential standards for neuropsychological characterization of mHE [10,11]. However these test batteries and scores have been employed in only two European populations, and their validity for mHE diagnosis was never proven in comparison to a standard diagnostic procedure [17–19,32–34]. Their value in clinical multicenter trials remained to be established. Therefore we used the PSE-Syndrome-Test for the first time as an endpoint for efficacy-analysis in a clinical trial on mHE/oHE [35]. Additionally, data analysis was directed to the accuracy of HE-severity assessment using West-Haven-criteria and the precision in conducting the paper–pencil tests. The PSE-Syndrome-Test uses the same individual paper–pencil tests as does the sumscore of the PHES-test and both test systems differ in their published normative values only [18,31] with impact for the sumscores [10,11,18,19,32,33]. Since in the meantime the recently validated PHES-Score was proposed as a standard diagnostic procedure for mHE-grading [18,19] the analysis further focused on the performance, the value and the balancedness of both sumscores and their treatment sensitivity.

2. Patients and methods

2.1. Study cohort and design

The data were obtained in a cohort of 217 cirrhotic patients with low-grade HE who participated in a multicenter, randomized, double-blind, placebo-controlled, clinical trial, which was completed in 1998 [35]. The primary objective of the trial was the efficacy of an oral active compound compared to placebo in cirrhotic patients with mHE or oHE [2,3] with regard to performance time in NCT-B. One further objective was the efficacy regarding performance in the sumscore of the PSE-Syndrome-Test, which was used for the first time in a clinical trial. For these purposes the study protocol had been reviewed and approved by German and Austrian responsible Ethics Committees. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki [36].

Only patients with stable mHE or oHE without identifiable precipitating factors were included. Exclusion criteria were severe HE (grades III and IV), decompensated cirrhosis (CHILD–PUGH >2), acute alcoholic hepatitis or obvious alcohol abuse at the beginning of the trial, a history of gastrointestinal or variceal bleeding during the previous 2 weeks, and uncorrectable electrolyte disturbances. Other exclusion criteria as well as additional medical and nutritional aspects of this trial were identical to those used previously [37].

Patients were randomly assigned to the treatment groups after giving informed consent and fulfilling inclusion and exclusion criteria. The treatment-phase covered a period of 6 weeks, followed by a 2 week washout period. Control time points for determination of HE-severity according to West-Haven-criteria and psychometric testing (paper–pencil test battery) were at day 0 (baseline), and days 14, 42, and 56.

2.2. Psychometric tests and sumscore of the PSE-Syndrome-Test (paper–pencil test battery)

The paper–pencil test battery consisted of the number-connection-test-A (NCT-A) and -B (NCT-B), digit-symbol-test (DST), line-tracing-test (LTT – time and errors), serial-dotting-test (SDOT) and had been validated for use in patients with HE [32,33]. The four available validated parallel forms of each test, the NCT-A, NCT-B, DST, SDOT, and LTT, were applied to the individual patients in a randomized way as described previously [5,12,19,27,31]. All physicians participating in the study were trained in conduct and evaluation of the psychometric tests. Test results within ±1SD from the age-adjusted mean were scored with 0 points, those between −1 and −2SD were scored −1, those between −2 and −3SD beyond the mean were scored −2 points and those worse than −3SD were scored with −3 points. Results better than means + 1SD were scored +1 point. The individual test scores of the PSE-Syndrome-Test were summarized to a sumscore of the PSE-Syndrome-Test, allowing score results to range from +6 to −18 points (see Supplementary Table 1). All investigators got an instruction folder which explained the exact use of the PSE-Syndrome-Test and its normatives, as well as the West-Haven-criteria (see Supplementary Table 2).

2.3. Definition of mHE and oHE

Assessment of oHE was performed according to the West-Haven-criteria, assessment of mHE according to the psychometric test results [2,3,31–33,37]. Test results beyond one standard deviation from the age-adjusted mean standard value of the PSE-Syndrome-Test were considered abnormal [32,33]. Cirrhotic patients without evidence of clinical abnormalities, but in whom neuropsychological abnormalities from the age-adjusted mean value of >1SD were found in at least two psychometric tests (one of these had to be the NCT-B), were classified as mHE patients. Thus, patients with only one pathological paper–pencil test were classified as HE 0.

2.4. Re-evaluation of data documented in the case report files

An erroneous classification of HE 0, mHE or HE I by the investigators was observed during data management procedures. In order to guarantee a uniform HE-grading, a re-evaluation of the HE-graduations by the independent institution WIMED GmbH (Munich, Germany) was arranged. This re-evaluation was based on individual data from the case report files and the psychometric test documentations according to the definitions in the study protocol.

The psychometric tests in many cases were not performed in accordance with the instruction-folder and were therefore also re-evaluated by an independent psychologist, who was unaware of the assignment to treatment groups. If two or more major violations from the test instructions were present, the individual test was considered invalid (see also Table 2).

2.5. Data management and statistical methods

Double data entry in a SAS-database (SAS Inst. Inc., Cary, NY, USA) was carried out. SPSS, version 11.0 for Windows (SPSS Inc. Headquarters, Chicago, USA), was used for the present analysis. The data from day 0 (baseline) as well as data from the final visit day 42 were used for the main statistical analysis. Missing data were replaced by last value carried forward (LOCF) principle. Regression and variance analyses were performed to investigate the influence of HE-grading on various test results. In order to test for internal consistency of the sumscores of PSE-Syndrome-Test and PHES-Score the Cronbach-alpha-Index was calculated. Categorical data were analyzed by means of frequency methods. The HE-graduation from the investigators and the re-evaluated graduation were cross-tabulated. As a measure of agreement the Kappa coefficient was calculated from these tables. Sensitivity and specificity analyses were performed for the single