Optimizing the Performance of the Psychometric Hepatic Encephalopathy Score (PHES) for the Diagnosis of Hepatic Encephalopathy

Clive D Jackson,[1] Michael Marks,[2] Marsha Y Morgan[2]

[1]Department of Neurophysiology, The Royal Free Hospital, The Royal Free London NHS Foundation Trust, London, [2]UCL Institute for Liver & Digestive Health, Division of Medicine, Royal Free Campus, University College London, London, UK

Background and Aims:

PHES is the most established test for the diagnosis of hepatic encephalopathy (HE). The availability of normative data to adjust for potential confounding variable is a necessity. However, there are other variables which may confound the performance of the PHES test and might result in the misclassification of patients, including: (i) the confounding variables used for test correction; (ii) possible interactions between the corrections variables; (ii) the scoring of the line tracing test (LTT) and (iii) the final summation of the component scores. The aim of this study was to produce a series of models to determine the influence of these variables on the performance of the PHES test.

Methods:

PHES testing was undertaken in 524 permanent UK residents aged > 18 years, who spoke good English, and were generally healthy. A UK PHES model was constructed as follows: (i) a reciprocal model for the LTT time/error relationship was devised; (ii) univariate/multivariate analyses identified significant correction factors for the individual test scores, without applying interaction terms; and (iii) each corrected test score was transformed to a Z-Score and then summed. The final PHES score was expressed as a Z-Score; values below -2.0 were considered abnormal. The effects of changing the basic PHES model by adding in interaction terms, changing the LTT scoring system and the final summation method were examined by applying the models to PHES data collected in 226 patients with cirrhosis.

Results:

Seventy–six (33.6%) of the 226 patients had abnormal UK PHES scores. Variation in the scoring of the LTT test (German, Spanish and Italian versions) resulted in a minor reduction in the number of abnormal results (73 [32%]). Add-in and subtraction models for the independent correction factors – age, sex, ethnicity, alcohol consumption, years/place of education –were associated with changes of -11% to +9.8% and +2.4 to -8.8% respectively, in the proportion of abnormal result; ethnicity and age had the greatest effects. Adding interaction terms for the correction factors marginally increased the number of abnormal results to 84 (37.2%). Changing the final score summation from Z-scores to integerized scores significantly increased the number of abnormal tests to 100 (44.2%) p=0.026.

Conclusion:

The major factor affecting the proportion of abnormal PHES results is the method used for final summation of the scores. Integerization of the scores does not allow for the exact application of thresholds and its use should perhaps be reconsidered.