Use of the Stroop EncephalApp for covert hepatic encephalopathy screening in cirrhotic patients in Southern Brazil

Paulo André Bispo MACHADO JÚNIOR1, Rafaela Deneka ZILIOTTO1, Ana Paula Vieira Neves FERREIRA1, Thiago Mesquita BUSON1, Cynthia Müller do COUTO2 and Alcindo PISSAIA JUNIOR1,2

ABSTRACT – Background – Covert hepatic encephalopathy (CHE) is difficult to detect due to the lack of easily applicable screening tools. The Stroop EncephalApp is a smartphone application already validated for CHE screening. However, its applicability to the Brazilian population is not known. Objective – To estimate the prevalence of CHE and evaluate the use of Stroop EncephalApp in a cirrhotic population in Brazil. Methods – In this cross-sectional study, we evaluated 99 patients previously diagnosed with liver cirrhosis in a Private Hospital in Curitiba/PR. Patients were initially submitted to the mini mental state examination (MMSE) to exclude individuals with dementia. After, the Psychometric Hepatic Encephalopathy Score (PHES) test was performed and lastly, the Stroop EncephalApp test. Results were adjusted for age, sex and education levels to evaluate the accuracy of the app on detecting the disease, comparing its results with the gold standard method (PHES). Patients with one or more of the following were excluded: dementia, inadequate MMSE score, illiteracy, color blindness, history of drugs/alcohol abuse within the past 3 months and previous or actual episodes of encephalopathy. The statistical analysis was performed by SPSS 2.0 and the significance adopted by 5%. Results – We included 82 individuals in the final analysis. Among these patients, 29 were diagnosed with CHE by the PHES test (35.36% prevalence) and 28 of those obtained equal diagnosis by the Stroop EncephalApp (96.6% sensitivity). A total of 53 patients obtained negative results for CHE by PHES, while the Stroop test classified 27 of them as having the disease. In the multivariate analysis, high levels of education were associated with better performance during the tests. No significant relationship was observed between age and sex with the probability of diagnosing CHE through the PHES test. Conclusion – The Stroop EncephalApp is a viable tool with good sensitivity for CHE screening, but its specificity in the studied population is low.

INTRODUCTION

Hepatic encephalopathy (HE) is one of the most common complications of advanced liver disease, present in almost 40% of these individuals10. HE clinically presents through sleep disorders, lethargy, psychomotor agitation and coma, being graduated according to the severity of the signs and symptoms16. The mildest expression of HE is covert hepatic encephalopathy (CHE), which cannot be graduated due to the subtlety of its clinical presentation1,3.

CHE has important implications on quality of life, affecting routine situations such as working memory and driving skills16. Therefore, it is fundamental to obtain a definite diagnosis in order to prevent its evolution and reduce social and economic impacts. Clinical examination is not sufficient to identify the disease due to the lack of obvious symptoms; thus, complementary tests are necessary5,6,8.

The most commonly used test is the Psychometric Hepatic Encephalopathy Score (PHES), a set of five tests capable of evaluating attention, perception, visuospatial construction, psychomotor and accuracy speed7. Despite being considered as the gold-standard method for the diagnosis of CHE, this test has a prolonged time of application and it is not available for universal use for free8,9. Due to this difficulty, more accessible tools have been developed to facilitate the clinical diagnosis of CHE.

Bajaj et al. developed the Stroop EncephalApp in 2013, a smartphone application that evaluates mental flexibility and psychomotor speed10, and it is already studied for CHE screening in the United States and China11,13,14. Despite this, its applicability to the Brazilian population is not known. The aim of our study was to estimate the prevalence of CHE and evaluate the accuracy of Stroop App in a cirrhotic population in Brazil.

METHODS

Population and study design

This cross-sectional study evaluated 99 cirrhotic patients, aged 18-80 years, from the Hepatology Service of Hospital Nossa Senhora das Graças Hospital (Curitiba/PR) between 2017 and 2019. All patients filled and signed consent forms. This study was approved by local ethics committee (process 2.237.426) in accordance with the Brazilian resolution 466/2012.

1 Declared conflict of interest of all authors: none
2 Disclosure of funding: no funding received

1 Pontifícia Universidade Católica do Paraná (PUCPR), Escola de Medicina, Curitiba, PR, Brasil. 2 Hospital Nossa Senhora das Graças (HNSG), Serviço de Hepatologia, Curitiba, PR, Brasil. Corresponding author: Alcindo Pissaia Junior. E-mail: alcindojr@yahoo.com
Patients who presented any of the following conditions were excluded from the study: history of previous encephalopathy, inability to obtain a minimal Mini-mental score (≥22 for individuals with years of schooling between 1 to 11; ≥27 for patients with more than 11 years of schooling), illiteracy, color blindness, history of drug or alcohol abuse over the past 3 months, psychological illness (analyzed through medical history and current medications), inability to understand and/or sign the consent forms and those who spontaneously did not agree to participate. Patients without exclusion criteria were then submitted to three neurocognitive tests in sequence: Mini-mental score, PHES and Stroop EncephalApp. The study protocol is presented in FIGURE 1.

The Stroop EncephalApp is a smartphone application (Play in its Portuguese version. The test consists of two steps: the on and off state. Two runs were performed before the timer starts to allow patients to practice and evaluate the comprehension about how to correctly answer the test.

At the off state (FIGURE 2), patients were presented with several hashtags (#####) on either blue, red or green colors. The objective was to answer, on the shortest amount of time, the corresponding color option shown at the bottom of the screen. Color positions options were randomly changed by the application. Each run contained 10 steps; if a mistake was made during any step, such as choosing the wrong color, the phase was automatically interrupted and reinitiated. All patients had to have 10 steps correctly answered. If the patient answered incorrectly in any of the states more than 20 times the app automatically closed down.

After finishing the test, the application provided the following parameters: total amount of time for five correct runs in the off state (Offtime), number of runs needed to complete five correct runs in the off state, total amount of time for five correct runs in the on state (Ontime) and number of runs needed to complete five correct runs in the on state. The processing of the psychomotor speed analysis consisted on adding Ontime + Offtime. Patient final result was obtained after the examiner accessed the Stroop App website and completed the forms with patients age, gender and education years. Then, the website generated a result indicating the presence or absence of CHE, based on the time taken to complete the test compared to the time expected for it, according to the subjects age, gender and educational level.

Psychometric hepatic encephalopathy score (PHES)

Considered as the gold-standard method for CHE diagnosis, the test consists in five different subtests directed to evaluate the cognitive dysfunction caused by cirrhosis. The first subtest is the digital symbol test, on which the patient was exposed to several symbols, each one corresponding to a number. The patient’s task was to fill boxes with the corresponding symbols of the numbers as shown in less than 120 seconds. The number of correct symbols drawn corresponded to be the patient score. The second subtest, number connection A (NCT-A), evaluates the amount of time that the patient needed to connect the numbers 1 to 25 in the correct order. On the third subtest, number connection B (NCT-B), the task was to connect numbers and letters in alternating order; meaning 1 to A, from A to 2, from 2 to B and so on. The time needed for all connections was again measured. The fourth subtest was the Serial Dotting (SD) test, on which the patient had to make a dot in the center of a circle as quick and precisely as possible. The time needed was also measured. The fifth subtest was the line tracing (LT) test, with the goal to draw a continuous line in between the two given lines without touching or crossing the printed lines. The time and the number of mistakes (number of times when the lines were touched or crossed) were also measured.

The PHES tests were obtained with previous consent of the Neurological Department of The Hannover Medical School, which provided its use over a symbolic cost. A manual on how to interpret the results of each test was provided. Each test has a normative data and a conversion scale according to the patients age and scores.

Statistical analysis

Quantitative variables were described by mean, standard deviation, median, minimum and maximum values. To evaluate the quality of the Stroop EncephalApp, were estimated values of sensibility, specificity, accuracy, positive predictive value and negative predictive value, using as gold standard the PHES. For the multivariate analysis of the factors related to the presence of CHE on the PHES test, a logistic regression model was adjusted using the Wald test to evaluate the significance of the variables.
The measure of association estimated was the odds ratio. $P$ values <0.05 indicated statistical significance. The data were analyzed by the computerized system IBM SPSS Statistics v.20.0. Armonk, NY: IBM Corp.

**RESULTS**

A total of 99 patients (mean age 57.6 years) previously diagnosed with liver cirrhosis were enrolled in the study. Of those, 17 were excluded from the final results due to not achieving the minimal score at mini mental state examination (MMSE) (n=11 patients) or not completing the three tests (n=6 patients). Thus, 82 individuals were included in the final analysis. The TABLE 1 show the main characteristics of our study population.

### TABLE 1. Socio-demographic characteristics of our study population (n=82).

| Variable                      | Female (n=29) | Male (n=53) | P*        |
|-------------------------------|---------------|-------------|-----------|
| Age (years)                   | 58.4±9.7      | 56.8±9.9    | 0.484     |
| Educational level (years)     | 12.3±4.8      | 12.5±4.2    | 0.833     |
| Cirrhosis etiology, n**       |               |             |           |
| NASH/HCV/HBV/Alcohol/Autoimmune/Other | 15/7/0/1/5/1 | 14/15/6/10/2/6 |           |
| Covert Hepatic Encephalopathy, n (%) |             |             |           |
| Absence                       | 19 (65.5)     | 34 (64.2)   |           |
| Presence                      | 10 (34.4)     | 19 (35.8)   | 1         |

Results presented by median ± stand deviation or frequency (percentage). *

* Fisher test; P<0.05. ** NASH: nonalcoholic steatohepatitis; HBV: hepatitis B virus; HCV: hepatitis C virus; Autoimmune: primary biliary cholangitis or primary sclerosing cholangitis or autoimmune hepatitis; Other: cryptogenic or hemorrhagic.

The population was considered homogenous for age and years of education ($P=0.484$ and $P=0.833$, respectively). From the total individuals, 29 woman and 53 men were analyzed. Among these patients, 35.3% (n=29) were diagnosed with CHE according to PHES test results. The most prevalent cause of liver cirrhosis among the women was NASH (51.7%, n=15), followed by hepatitis C (24.1%, n=7). Among the men, the causes were, from most to least prevalent: hepatitis C (28.3%, n=15), followed by NASH (26.4%, n=14), alcoholic cirrhosis (18.8%, n=10), hepatitis B (11.3%, n=6) and others (11.3%, n=6).

Regarding the performance of the patients at both Stroop and PHES tests, no significant statistic dissimilarity was observed between gender on most of the tests, except for the line-tracing test at the final phase of PHES, on which women accomplished better results. The comparison between men and women for each test phase is shown on TABLE 2.

Considering the PHES test as the gold-standard method for the diagnosis of CHE, both specificity and sensitivity of Stroop App were analyzed. Among the 82 cirrhotic patients included in the analysis, 29 had positive results for CHE according to PHES, of which 28 were equally diagnosed with MHE by the Stroop App, demonstrating that the app has a high rate of sensitivity (96.6%). The 53 cirrhotic patients that had PHES results negative for CHE, only 26 were equally classified according to the Stroop App results, denoting that the app has low specificity when compared to the PHES (49.1%). The quality of the Stroop App on predicting results (sensitivity, specificity, accuracy, false positive and false negative probability) when compared to the gold standard method is described on TABLE 3.

In order to evaluate the impact of age, gender and years of study of the 53 individuals over the possibility of having CHE diagnosed by the PHES, a multivariate analysis was performed (TABLE 4). Values of the statistic results and also the estimated values of the odds ratio (OR) are shown with their own confidence gaps of 95% (IC 95%). Logistic regression model was adjusted for the PHES (presence or absence of encephalopathy), including age, gender and schooling years as explicative variables. For each variable, on

### TABLE 2. Patients performance on psychometrics tests.

| Variable                       | Female (n=29) | Male(n=53) | P*        |
|-------------------------------|---------------|-------------|-----------|
| Stroop EncephalApp test        |               |             |           |
| Off time/s                    | 89.9±30.9     | 88±27.9     | 0.781     |
| On time/s                     | 10.9±24.3     | 98.8±31.9   | 0.550     |
| Off time + On time/s          | 196.3±66.8    | 184.6±60    | 0.420     |
| No. runs for off state        | 6 (5–9)       | 5 (5–14)    | 0.310     |
| No. runs for on state         | 5 (5–10)      | 5 (5–8)     | 0.292     |
| Off time – On time/s          | 14.2 (3–50)   | 8 (27–60.8) | 0.179     |

Results presented by median ± stand deviation. * Mann- Whitney U test; P<0.05. NCY number connection.

### TABLE 3. Stroop app predicting results.

| StroopApp                      | PHES (gold-standard) | Gender |
|-------------------------------|----------------------|--------|
|                               | CHE absent | CHE present |
| CHE absent                    | 26         | 1         |
| CHE present                   | 27         | 28        |
| Total                         | 53         | 29        |
| StroopApp Results             | CI 95%     |           |
| Sensitivity                   | 96.6%      | 89.9–100% |
| Specificity                   | 49.1%      | 35.6–62.5%|
| Accuracy                      | 65.9%      | 55.6–76.1%|
| FP +                          | 50.9%      | 37.5–64.4%|
| FP -                          | 3.4%       | 0–10.1%   |

FP+: False positive rate. FP-: False negative rate. PHES: Psychometric Hepatic Encephalopathy Score.

### TABLE 4. Multivariate analysis.

| Variable          | Classification | $P^*$ | OR | IC95%       |
|-------------------|----------------|-------|----|-------------|
| Age (years)       |                | 0.205 | 1.03 | 0.98–1.09  |
| Gender            | Female         | 0.739 | 1.19 | 0.42–3.32  |
|                   | Male           | 0.026 | 0.88 | 0.78–0.99  |

* Wald test: P<0.05.
the presence of any other of those included, the null hypothesis of lack of any association between the variant and the probability of having MHE was tested against the alternative hypothesis that there could be any sort of association. Age and gender were not associated with the possibility of testing positive for CHE on PHES test ($P=0.205$ and $P=0.739$, respectively). Schooling years were associated with a better performance of the tests (OR=0.88; 95%CI: 0.78–0.99)

**DISCUSSION**

In the current study, we evaluated the applicability of the Stroop App on CHE screening in cirrhotic patients in a private hospital in Southern Brazil. Besides its clinical relevance, diagnosing CHE remains challenging[18]. PHES is considered as the current gold-standard method for detecting CHE worldwide, but its basic requirements to be executed such as time and financial costs makes it unavailable for routine use[14]; therefore, easier and more convenient tools for CHE diagnosis are being pursued[15,16].

Previous studies had already demonstrated the efficacy of the Stroop App in diagnosing CHE in cirrhotic patients. The first one was a cross-sectional study which evaluated 125 individuals with liver cirrhosis and the test demonstrated 78% sensitivity and 90% specificity in detecting CHE[10]. In sequence, the Stroop was tested in 167 cirrhotic patients and it had 89.1% sensitivity and 82.1% specificity[11]. When compared to PHES, a multi-center analysis of 437 cirrhotic patients showed 72% and 88% sensitivity and specificity for the Stroop app, respectively[12].

The results of our study showed that 35% of the individuals (n=29) were diagnosed with CHE according to PHES results. This number is inferior to other studies evaluating the prevalence of MHE in Brazil, which demonstrated prevalence of 50%-51,18. After evaluation through the Stroop App test, the prevalence observed rose up to 67%, indicating that the app has a high sensitivity (96.6%) but low specificity (49.1%) in diagnosing CHE. These results are similar to what Zeng et al. found in a 144 patients analysis in China[13].

This low specificity of the app can be explained by a few reasons. The first one is that while PHES can be composed by five different subtests, all capable of assessing together the psychomotor speed, visual coordination, memory and attention of the patient, the Stroop App focuses only on evaluating the psychomotor speed and the alert status of the individual, through the calculation of the time between the visual stimulation and the motor cortex reaction facing the symbols and words of the app.

Considering this fact, a learning curve may be considered for the reproducibility of the test in order to avoid high rates of false positives as found in our study (50.9%). A previous study had already shown that repeating the test at regular intervals of time is capable of providing better results among cirrhotic patients with no history of CHE[19]; a similar follow-up in our patients would certainly show the veracity or not of the false positives for the Stroop test. Furthermore, both Stroop and PHES tests aim to evaluate the neurocognitive performance of the patients. For that reason, their results may suffer variations according to non-mensurable factors such as fatigue, stress, anxiety about obtaining a good test result and other aspects that could cause distraction during the exam.

The needed time for test application was 3.1 minutes for Stroop App and 7.1 minutes for PHES. No significant difference between men and women was observed during the performance of the tests, aside from the Line Tracing Test on which women showed better results ($P=0.035$). The mean time processing of the psychomotor speed analysis of our patient population through the Stroop test was 196.3 and 184.6 seconds for women and men, respectively (Ontime + Offtime, TABLE 2). Other studies found similar results[10,12].

On the multivariate analysis (TABLE 4), it was observed that high levels of education were associated with better performance during the tests. The odds ratio estimated value (OR=0.88; 95CI% 0.78–0.99) indicates that each additional year of study may reduce in 12% the chances of having MHE by the PHES test. In another interpretation, this could also mean that years of study were only related to a better performance during the test application.

No significant association between age and gender with the possibility of diagnosing MHE was observed. A brief observation is that some difficulty was noted in elderly patients to perform the tests, probably due to the aging process itself, causing them to present lower response reflexes when compared to younger patients. The elderly patients who had better results were those who were already familiarized with smartphones use or other tools for cognitive exercises.

This is the first study in Brazil to evaluate the use of Stroop EncephalApp as a screening tool for CHE. Our results reinforce the hypothesis that this application can be a useful tool in the follow-up of cirrhotic patients, due to its easy application, good sensitivity rates and free availability, allowing a better investigation of this specific neurological disorder caused by liver cirrhosis.

Despite this, this study has some considerable limitations. The lack of skills regarding the use of smartphones from some patients is a factor that can make the test application harder for these individuals, possibly increasing the final test time for them. Besides, the high schooling rate of the population analyzed can distinguish results from the real scenario of the general population in Brazil. Furthermore, the parameters for Stroop and PHES were defined according to studies based on North American and German populations, due to the absence of reference values in Brazil. Future studies should focus on the validation of the PHES and Stroop App in Brazilian individuals, defining adequate analysis parameters for our population and controlling possible confounding factors based on its specific rates of schooling, gender and age.

**CONCLUSION**

Covert hepatic encephalopathy has a 35% prevalence in our population, and the Stroop App is a useful tool with good sensitivity for screening CHE in the analyzed sample.

**ACKNOWLEDGEMENTS**

We thank all the patients and doctors from the Hepatology service of Hospital Nossa Senhora das Graças in Curitiba/PR for the opportunity to develop this study, especially Dr. Daphne Benatti Morsoletto, Dr. Helen Perussolo Alberton and Dr. Claudia Pontes Ivantes. We are also grateful to Prof. Marcia Olandoski from the Epidemiology and Statistics Group of PUCPR for the comments and suggestions.

**Authors’ contribution**

Machado Júnior PAB: data collection and writing of the text.
Ziliotto RD: data collection and writing of the text. Ferreira APVN: data collection and writing of the text. Buson TM: data collection and writing of the text. Couto CM: statistical analysis, literature research, writing of the text and revision of the manuscript. Pissaia Junior A: advisor, design of the study, data interpretation and final revision of the manuscript.

Machado Júnior PAB, Ziliotto RD, Ferreira APVN, Buson TM, Couto CM, Pissaia Junior A. Uso do Stroop EncephalApp para o screening de encefalopatia hepática mínima em pacientes cirróticos no sul do Brasil. Arq Gastroenterol. 2020;57(4):399-403.

RESUMO – Contexto – A encefalopatia hepática mínima (EHM) é uma complicaçãono-siquiátrica da cirrose cuja detecção é dificultada pela falta de ferramentas práticas. O Stroop EncephalApp é um aplicativo de smartphones capaz de detectar a doença, entretanto sua aplicabilidade na população brasileira ainda não é conhecida. Objetivo – Avaliar o uso do Stroop EncephalApp para diagnóstico e avaliação de EHM em uma população de pacientes cirróticos no Brasil. Métodos – Através de um estudo observacional transversal, 99 indivíduos sabidamente cirróticos foram recrutados do ambulatório de hepatologia de um hospital privado em Curitiba/PR. Primeiramente, foram aplicados o mini exame do estado mental (MEEM) para excluir indivíduos com demência; após, foram aplicados o Escore Psicométrico da Encefalopatia Hepática (PHES), atual padrão-ouro para diagnóstico de EHM, e posteriormente o Stroop EncephalApp, ajustando para idade, sexo e anos de formação acadêmica, buscando avaliar a eficiência do aplicativo em detectar a doença e comparar seus resultados com o atual padrão-ouro. Foram excluídos do estudo indivíduos com demência, pontuação insuficiente no MEEM, analfabetos, daltônicos, e com histórico de abuso de álcool/drogas ilícitas nos últimos 3 meses e paciente com episódios prévios ou atuais de encefalopatia hepática. A análise estatística foi realizada pelo SPSS 2.0 e a significância adotada em 5%. Resultados – Um total de 82 indivíduos foram incluídos na análise final. Destes, 29 foram diagnosticados com EHM (35,36% de prevalência) através do PHES e 28 obtiveram o mesmo resultado após a aplicação do Stroop (96,6% de sensibilidade). Cinquenta e três pacientes obtiveram um resultado negativo para EHM através do PHES, sendo que 27 desses obtiveram um resultado positivo para EHM através do Stroop. Na análise multivariada, níveis elevados de escolaridade estiveram associados com melhor desempenho durante a execução dos testes. Não houve associação significativa entre idade e sexo com a probabilidade de apresentar encefalopatia através do PHES. Conclusão – O Stroop EncephalApp é uma ferramenta viável e com boa sensibilidade para o screening de EHM, mas possui baixa especificidade na população estudada.

DESCRITORES – Cirrose hepática. Encefalopatia hepática. Teste de Stroop. Telefone celular, instrumentação. Psicometria.

REFERENCES

1. Völsch J, Amodio P, Bajaj J, Cordoba J, Ferenci P, Mullen KD, 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;60:715-35.

2. Feltracco P, Cagnin A, Carollo C, Barbieri S, Ori C. Neurological disorders in liver transplant candidates. Pathophysiology and clinical assessment. Transplant Rev (Orlando). 2017;31:193-206.

3. Weissenborn K. Diagnosis of minimal hepatic encephalopathy. J Clin Exp Hepatol. 2015;5:S54-9.

4. Prakash RK, Kanna S, Mullen KD. Evolving concepts: The negative effect of minimal hepatic encephalopathy and role for prophylaxis in patients with cirrhosis. Clin Ther. 2013;35:1485-73.

5. Nardone R, Taylor AC, Höller Y, Brigo F, Lochner P, Trinka E. Minimal hepatic encephalopathy: A review. Neurosci Res. 2016;111:1-12.

6. Bajaj JS, Etemadian A, Haferzullah M, Saedan K. Testing for minimal hepatic encephalopathy in the United States: An AASLD survey. Hepatology. 2007;45:833-4.

7. Nabi E, Bajaj JS. Useful tests for hepatic encephalopathy in clinical practice. Curr Gastroenterol Rep. 2014;16(1).

8. Patidar KR, Bajaj JS. Covert and overt hepatic encephalopathy: Diagnosis and Management. Clin Gastroenterol Hepatol. 2015;13:2048-61.

9. Barone M. Neurophysiologic and psychometric tests in the diagnosis of low-grade encephalopathy. Gastroenterology. 2015;148:463.

10. Bajaj JS, Thacker LR, Heuman DM, Fuchs M, Sterling RK, Sanyal AJ, et al. The stropo smartphone application is a short and valid method to screen for minimal hepatic encephalopathy. Hepatology. 2013;58:1122-32.