**Significance of the $^{13}$C-caffeine breath test for patients with cirrhosis**

**Dimitrios Konstantinou, Ekaterini Margariti, Emilia Hadziyannis, Dimitrios Pectasides, George V. Papatheodoridis**

Hippokration General Hospital; Laikon General Hospital, Athens, Greece

---

**Abstract**

**Background** The $^{13}$C-caffeine breath test (CBT) is a non-invasive, quantitative test of liver function which has been shown to correlate inversely to the Child-Pugh score. The aim of the study was to determine the utility of CBT in the assessment of cirrhosis and its correlation to the model for end-stage liver disease (MELD) score.

**Methods** Thirty-nine patients, 29 with cirrhosis and 10 with chronic liver disease without cirrhosis, and 8 healthy volunteers were included. Cirrhotic patients were graded according to Child-Pugh and MELD scores. All participants underwent CBT and laboratory tests on the same day. The results of the CBT were expressed as percentages of changes over baseline values (Δ‰) per 100 mg caffeine.

**Results** The mean single 15-min, 30-min, 45-min and 1-h CBT results, as well as cumulative CBT values differed significantly between healthy controls or chronic liver disease patients and cirrhotics (1-h CBT: 3.22±1.06 or 3.56±2.80 vs. 1.69±2.52, P≤0.01). In contrast, the CBT results at any time point or cumulative values did not correlate with MELD or Child-Pugh scores. Receiver operating characteristics (ROC) analysis showed that the 30-min CBT values were more accurate in differentiating cirrhotics from chronic liver disease patients (area under ROC curve: 0.871).

**Conclusions** CBT can reliably differentiate the patients with decompensated cirrhosis from non-cirrhotic patients with chronic liver diseases. However, in patients with decompensated cirrhosis, CBT results do not seem to be associated with the Child-Pugh and MELD scores.

**Keywords** Caffeine breath test, MELD, Child-Pugh, cirrhosis

Ann Gastroenterol 2014; 27 (1): 1-7

---

**Introduction**

Any chronic liver disease can lead to cirrhosis, characterized histologically by extensive fibrosis and the conversion of normal liver architecture into structurally abnormal nodules [1]. The severity of cirrhosis may range from a completely asymptomatic compensated phase to symptomatic decompensated liver disease presenting with ascites, variceal bleeding, hepatic encephalopathy and/or non-obstructive jaundice [2-4]. Since cirrhosis represents the end-stage of any chronic liver disease, it has been mandatory to elaborate reliable and accurate predictors of its severity and prognosis. To date, two widely accepted prognostic scores are mainly used in clinical practice worldwide: the Child and the subsequently modified Child-Pugh score [5,6], and the more recently developed MELD (model for end-stage liver disease) score [7,8], which has replaced the Child-Pugh score in prioritizing patients on the waiting list for liver transplantation [9,10].

Several $^{13}$C-breath tests have been recently evaluated for the assessment of liver function in a variety of chronic liver diseases. Among them, the caffeine breath test (CBT) appears to be promising, since caffeine has high oral bioavailability, almost exclusive hepatic metabolism principally via demethylation by cytochrome P450 1A2 to CO$_2$ [11], low extraction ratio, low plasma binding capacity and may be considered innocuous when used at test doses [12,13]. Thus, CBT values...
can reflect the functional hepatic reserve and represent a good marker of the severity of liver function impairment.

Park et al found that CBT values correlate inversely and reproducibly with the degree of hepatic dysfunction, as estimated by Child-Pugh score. CBT values were lower in subjects with cirrhosis of varying etiology than in non-cirrhotic patients with chronic hepatitis B or C [14]. They also showed that the 13C-CBT results can assess the severity of fibrosis in chronic hepatitis B patients and may show improvement in liver function in response to long-term lamivudine therapy [15]. In two further studies, it was shown that the CBT results can reliably predict severe hepatic fibrosis in patients with non-alcoholic fatty liver disease [16,17]. It should be emphasized that the existing studies assessing the performance of CBT have included only patients with chronic liver disease and well-compensated cirrhosis, but not cases with decompensated cirrhosis. The aim of the current study was to determine the utility of the CBT in assessing cirrhosis and examine its relationship with MELD score.

**Patients and methods**

**Participants**

In total, 66 subjects were screened and 47 of them were included in the study (Fig. 1): 29 patients with cirrhosis consecutively admitted to our Department between September 2011 and March 2012, 10 non-cirrhotic patients with chronic liver disease visiting our outpatient liver clinics during the same period and 8 healthy volunteers as controls. The diagnosis of cirrhosis was based on histological findings and/or clinical signs of decompensation which included ascites, variceal bleeding, hepatic encephalopathy and non-obstructive jaundice (bilirubin >3 or 10 mg/dL for non-cholestatic or cholestatic chronic liver disease, respectively). The absence of cirrhosis in the 10 patients with chronic liver disease was based on histological findings from liver biopsies within the last year. Healthy volunteers had no abnormal clinical or biochemical (aminotransferases, alkaline phosphatase and gamma-glutamyl-transpeptidase) findings, reported daily alcohol use <20 g and no recent medication use, had no evidence of hepatic steatosis at ultrasonography and were negative for hepatitis B surface antigen and antibodies against hepatitis C virus. Subjects with a history or symptoms and signs of significant cardiopulmonary disease were excluded as were those with hepatocellular carcinoma, signs of sepsis within the last week and those enrolled in the waiting list for liver transplantation. No patient should have had active variceal bleeding within the last week. The study was approved by the hospital ethics committee and all participants gave written informed consent before enrollment.

Clinical and laboratory data were collected on the day of the CBT. A complete medical history was recorded and complete physical examination was performed for all study participants. Patients with cirrhosis were scored using the Child-Pugh [6] and MELD score [7]. All subjects were classified according to their smoking habits into 3 categories: current smokers, ex-smokers defined as cases that ceased smoking at least 12 months prior to study enrollment, and non-smokers. Since there have been reports suggesting that CBT results may be affected by smoking, all associations of CBT values were evaluated in the total patient population and separately in ex- or non-smokers (after the exclusion of current smokers).

**Caffeine breath test**

All study participants underwent 13C-CBT. Subjects abstained from caffeine-containing products and alcohol for at least 24 h before testing. All non-essential medications were withheld for 48 h before testing. After an overnight fast, subjects ingested 2 mg/kg of [3-methyl-13C]-caffeine [99% 13C], obtained in powder form from Cambridge Isotope Laboratories (Cambridge, MA) and dissolved in 30 mL of water, followed by a 40 mL water wash of the container. The quantity of caffeine consumed was approximately equivalent to 2 cups of coffee. Subjects were instructed to sit quietly for 15 min before and throughout the CBT to prevent an effect of physical activity on endogenous CO2 production [18].

![Figure 1 Flow chart of screened and enrolled patients](image_url)

*COPD, chronic obstructive pulmonary disease; HCC, hepatocellular carcinoma; LT, liver transplantation*
Breath samples were obtained during prolonged expiration into 10 mL glass vials via straws. Samples were collected just before and at 15, 30, 45 and 60 min after caffeine ingestion. The $^{13}$C-enrichment of expired CO$_2$ was determined by continuous flow isotope ratio mass spectrometry using the automated 13-carbon breath analyzer (FanCi2 breath tester, Olympus). The average predosage measurement was subtracted from the average post-dosage measurement relative to the international $^{13}$C/$^{12}$C Pee-Dee-Belemnite standard (ratio=0.0112372; on the basis of a fossil limestone of the Pee-Dee-Formation, SC). These results were expressed as percentages of changes over baseline values ($\Delta$‰) per 100 mg caffeine. The cumulative $^{13}$C-enrichment over a particular time interval was calculated by averaging the measured enrichments over that time interval. For example, the cumulative 1-h value represented the average $^{13}$C enrichment of 4 breath samples taken every 15 min during the first hour after administration of $^{13}$C-caffeine. The single 1-h CBT value represented the $^{13}$C enrichment measured 1 h after substrate dosing. Accordingly, the single one-point CBT value represented the $^{13}$C enrichment measured up to that point (15 min, 30 min, 45 min). Because the doses of labeled caffeine were adjusted to body weight, the $^{13}$C enrichment values were recorded without body weight normalization [14].

Statistical analysis

All analyses were carried out using the SPSS Statistics 18.0 (SPSS Inc, an IBM Company). Quantitative data were expressed as mean values ± standard deviation or as median (range) values. The normality of distribution of continuous quantitative variables was tested using the Shapiro-Wilk test. Comparisons between groups of quantitative variables with normal or not normal distribution were performed using the t-test or Mann-Whitney test, respectively. Comparisons of categorical variables were performed using the corrected chi-square or the two-sided Fischer’s exact test, as appropriate. Correlations between quantitative variables were assessed using Pearson’s correlation coefficient or Spearman’s correlation coefficient, as appropriate. Receiver operating characteristics (ROC) analysis was applied to assess the performance and area under the ROC curve (AUROC) of each of the CBT time values. A two-tailed P value <0.05 was considered to be statistically significant.

Results

Participants’ characteristics

The main baseline characteristics of the participants are presented in Table 1. The mean age of the healthy controls, chronic liver disease cases and patients with cirrhosis was 38±17, 38±15 and 60±16 years, respectively (P=0.001). There were 26 (55%) males and 21 (45%) females, while current smokers were 23 (49%), ex-smokers 7 (15%) and non-smokers 17 (36%) cases. Among the patients with cirrhosis, the mean Child-Pugh score was 9±2 (range: 5-13) and the mean MELD score 10±5 (range: 6-22), whereas the majority (90%) of cirrhotics had Child class B or C. Twelve of the patients with cirrhosis were taking proton pump inhibitors and/or ciprofloxacin at the time of CBT.

CBT results

The mean single 1-h CBT (CBT-1h) values were significantly lower in the 29 patients with cirrhosis (1.69±2.52) than in the 10 patients with chronic liver disease (3.56±2.80, P=0.01) or in the 8 healthy controls (3.22±1.06, P=0.004), but they showed no significant difference between healthy controls and chronic liver disease patients (P=0.73). Similar differences were also observed for the results of single points CBT at 15, 30 and 45 min and the cumulative results of CBT, as shown in Fig. 2A. The exclusion of smokers did not change the results of these comparisons (Fig. 2B). Similarly, the exclusion of patients under proton pump inhibitors and/or ciprofloxacin did not change the results of these comparisons.

Among patients with cirrhosis, the Child-Pugh score was not found to correlate to CBT results (Table 2). However, CBT results (at 15, 30, 45, 60 min and cumulative) were

<table>
<thead>
<tr>
<th>Table 1 Main baseline characteristics of the 47 study participants</th>
<th>Healthy controls, N=8</th>
<th>Chronic liver disease, N=10</th>
<th>Cirrhosis, N=29</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (males), n (%)</td>
<td>5 (63)</td>
<td>3 (30)</td>
<td>18 (62)</td>
</tr>
<tr>
<td>Age, years</td>
<td>37±17*</td>
<td>37±15*</td>
<td>59±16*</td>
</tr>
<tr>
<td>Smoking habits, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non/Ex-smokers</td>
<td>4 (50)</td>
<td>6 (60)</td>
<td>14 (48)</td>
</tr>
<tr>
<td>Current smokers</td>
<td>4 (50)</td>
<td>4 (40)</td>
<td>15 (52)</td>
</tr>
<tr>
<td>Cause of liver disease, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBV infection</td>
<td>-</td>
<td>4 (40)</td>
<td>2 (7)</td>
</tr>
<tr>
<td>HCV infection</td>
<td>-</td>
<td>5 (50)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>-</td>
<td>0</td>
<td>18 (63)</td>
</tr>
<tr>
<td>Other</td>
<td>-</td>
<td>11 (10)</td>
<td>32 (11)</td>
</tr>
<tr>
<td>Unknown</td>
<td>-</td>
<td>0</td>
<td>5 (18)</td>
</tr>
<tr>
<td>Child class, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-</td>
<td>-</td>
<td>3 (11)</td>
</tr>
<tr>
<td>B</td>
<td>-</td>
<td>-</td>
<td>12 (42)</td>
</tr>
<tr>
<td>C</td>
<td>-</td>
<td>-</td>
<td>14 (49)</td>
</tr>
<tr>
<td>MELD score</td>
<td>-</td>
<td>-</td>
<td>10.3±4.9</td>
</tr>
</tbody>
</table>

MELD, model for end-stage liver disease

*P<0.05 for cirrhotics vs. patients with chronic liver disease or vs. healthy controls

1One patient with non-alcoholic steatohepatitis; 2two patients with primary biliary cirrhosis and one patient with autoimmune hepatitis

Annals of Gastroenterology 27
significantly lower in patients with Child class B or C than in patients with chronic liver disease or cirrhosis of Child class A or in healthy controls (P<0.02 for all comparisons). In contrast, there was no difference between CBT cumulative results or the CBT results at any single time point between patients with cirrhosis of Child class B or C (P≥0.70 for all comparisons) (Fig. 3A and 3B respectively). The exclusion of smokers did not change the previous results.

ROC curve analyses showed that the CBT results at 30 min were more accurate in differentiating cirrhotics from chronic liver disease patients (AUROC: 0.871) with a cut-off value of 1.24 offering sensitivity of 78% and specificity 100%.

MELD score did not correlate to CBT results at any single time point or to the cumulative results (Table 2). In addition, CBT results at any time point (15, 30, 45, 60 min) or cumulative CBT values did not differ significantly between patients with MELD score ≤9 and patients with MELD >9 (P>0.47 for all comparisons) (Fig. 3C). The differences and the correlations remained non-significant even after the exclusion of smokers (P>0.10 for all comparisons).

The CBT results at any time point or the cumulative CBT values did not correlate to age or gender (P>0.05 for all comparisons). The correlations between CBT and the laboratory data are summarized in Table 3. The international normalized ratio correlated inversely with the CBT cumulative values (r=-0.599, P≤0.001). This correlation was enhanced after the exclusion of smokers (r=-0.816, P≤0.001). Moderate correlations were also observed between albumin and bilirubin and the CBT cumulative results.

**Discussion**

The current study shows that the CBT results can be helpful in the differentiation of patients with advanced liver disease. In particular, CBT values were significantly lower in patients with cirrhosis compared to non-cirrhotic patients with chronic liver disease or healthy controls, but they did not differ between patients with chronic liver disease and healthy controls. Our findings further support previous findings on the usefulness of CBT in the non-invasive evaluation of severe chronic liver disease [14,19].

CBT has several advantages that render it an appealing non-invasive method for the evaluation of the severity of liver disease. CBT is relatively easy to perform, and it is performed non-invasively, which makes it more acceptable to patients. Furthermore, CBT is not affected by smoking, which is a common behavior among patients with liver disease. Finally, CBT can be performed quickly, and it does not require the use of complex equipment.

**Table 2** Associations of caffeine breath test (CBT) results with Child-Pugh and model for end-stage liver disease (MELD) scores in 29 patients with cirrhosis

<table>
<thead>
<tr>
<th></th>
<th>Child-Pugh</th>
<th>MELD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r  P</td>
<td>r  P</td>
</tr>
<tr>
<td>CBT at 15 min</td>
<td>-0.193 0.32</td>
<td>-0.17 0.40</td>
</tr>
<tr>
<td>CBT at 30 min</td>
<td>-0.225 0.25</td>
<td>-0.06 0.78</td>
</tr>
<tr>
<td>CBT at 45 min</td>
<td>-0.182 0.35</td>
<td>0.015 0.94</td>
</tr>
<tr>
<td>CBT at 1 h</td>
<td>-0.292 0.12</td>
<td>-0.059 0.76</td>
</tr>
<tr>
<td>CBT cumulative</td>
<td>-0.215 0.30</td>
<td>-0.097 0.64</td>
</tr>
</tbody>
</table>

**Table 3** Correlations of caffeine breath test cumulative values with serum liver tests in all patients with liver disease and in non/ex-smokers

<table>
<thead>
<tr>
<th></th>
<th>All patients</th>
<th>Non/Ex-smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r  P</td>
<td>r  P</td>
</tr>
<tr>
<td>INR</td>
<td>-0.599 &lt;0.01</td>
<td>-0.816 &lt;0.01</td>
</tr>
<tr>
<td>Platelets</td>
<td>0.086 0.68</td>
<td>0.017 0.95</td>
</tr>
<tr>
<td>ALT</td>
<td>0.202 0.24</td>
<td>0.112 0.64</td>
</tr>
<tr>
<td>AST</td>
<td>-0.10 0.56</td>
<td>-0.252 0.29</td>
</tr>
<tr>
<td>GGT</td>
<td>-0.201 0.25</td>
<td>-0.346 0.14</td>
</tr>
<tr>
<td>Albumin</td>
<td>0.365 0.03</td>
<td>0.369 0.12</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>-0.420 0.01</td>
<td>-0.493 0.03</td>
</tr>
</tbody>
</table>

INR, international normalized ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl-transpeptidase
are generally depressed in patients with liver failure or severe cirrhosis [23,24]. In particular, CYP1A2 and its related catalytic activity are significantly reduced in both cholestatic and non-cholestatic types of liver disease, whereas the activities of other P450 proteins appear to undergo less pronounced and variable alterations [25]. The CBT, therefore, provides a convenient assay of microsomal P450 1A2 metabolic activity and yields quantitative information concerning the degree of hepatic functional derangement [26].

Our findings confirm previous reports [11-14,27], but several issues need to be considered when interpreting our results. First, the cirrhotic patients included in our study had advanced liver disease (Child class B or C in 26 of 29 cases) and were usually hospitalised for refractory ascites, variceal bleeding, jaundice, hepatic encephalopathy and spontaneous bacterial peritonitis or other infections. At the time of the CBT, they were all hemodynamically stable and presented no sign or laboratory data consistent with infection or sepsis. Four cirrhotic patients were under ciprofloxacin. Second, the effect of smoking on caffeine metabolism was different in our study from previous reports. Cigarette smoking has been previously shown to increase caffeine clearance [28,29] and affect the CBT results [14]. In our study, the mean CBT results were not significantly different between current and non/ex-smokers, but the absence of a statistical difference could be due to a type II error associated with the small number of our patients. Furthermore, there may be confounding factors, either genetic or environmental, that contribute to a certain extent to variations in CYP 1A2 levels. In the case of genetic factors, Rasmussen et al identified a strong correlation in CYP1A2 activity between identical (monozygous) twins and suggested that genetic factors account for approximately 75% of the observed variation in CYP1A2 activity [30]. In the case of environmental factors, certain drugs (omeprazole, ciprofloxacin), cruciferous vegetable (e.g. broccoli), strenuous exercise and infectious diseases are all known to influence CYP1A2 expression [31]. In our study, we tried to eliminate the above factors by restricting exercise, excluding certain food ingredients from the diet, suspending all the unnecessary drugs and excluding patients with documented infection or sepsis. However, some cirrhotic patients were taking drugs like proton pump inhibitors and/or ciprofloxacin [32], which might have affected the CYP1A2 activity. Finally, a more provocative hypothesis is that P450 1A2 function in our patients was compromised due to the liver disease to such a degree that smoking cannot induce a significant increase in its activity, delineating a lower residual liver function. This lack of inducibility was shown in aminopyrine breath test [33], but needs to be tested in CBT, as well.

Third, we found no correlation between Child-Pugh or MELD score and the CBT values. MELD score is a valid prognostic score for intermediate term mortality in patients with cirrhosis [34,35]. Child-Pugh score is a simple descriptive marker of the severity of cirrhosis with an almost equivalent to the MELD score prognostic value [36]. Our conclusion is in agreement with the results of Lewis et al [37] who...
used the salivary caffeine clearance to evaluate the hepatic functional reserve in cirrhosis. The authors concluded that, while caffeine clearance correlated strongly with the Child-Pugh score in outpatients with mild cirrhosis, there was no correlation between the same parameters in hospitalized patients with more severe decompensated liver disease. Although the salivary caffeine clearance is a less reliable assay [38], the similarity with our results suggests that the hepatic functional compromise and the microsomal P450 1A2 derangement are not reflected by the Child-Pugh and presumably MELD scores. This observation remains to be validated in future larger studies.

Given its simplicity and non-invasive nature, there may be indications that CBT might be useful in clinical practice. One indication for CBT might be the long-term follow up of patients with early-stage cirrhosis, which may progress to a more advanced stage without substantial changes in clinical and/or laboratory findings. The CBT might be also used to improve the predictive value of current clinical scores. Such a combined score might improve the assessment of the severity of cirrhosis.

In conclusion, the present study shows that the CBT can reliably differentiate the patients with cirrhosis. Since CBT can be an easy to use, safe, non-invasive method for the evaluation of liver function impairment, larger prospective studies are required to assess its predictive role in the long-term outcome of patients with cirrhosis.

### Summary Box

**What is already known:**

- Reliable and accurate predictors of the severity and prognosis of patients with chronic liver disease are required in daily clinical practice
- Several 13C-breath tests have been recently evaluated for the assessment of liver function in a variety of chronic liver diseases
- In particular, the 13C-caffeine breath test (CBT) has provided promising results for the evaluation of such patients

**What the new findings are:**

- CBT values were found to reliably differentiate the patients with cirrhosis
- CBT values did not differ between non-cirrhotic patients with chronic liver disease and healthy controls
- In patients with decompensated cirrhosis, CBT results did not seem to be affected by the smoking habits or to be associated with the Child-Pugh or Meld score

### References

22. Nelson DR, Zeldin, Hoffman SM. Comparison of cytochrome
P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* 2004;14:1-18.


35. Al Sibae MR, Cappell MS. Accuracy of MELD scores in predicting mortality in decompensated cirrhosis from variceal bleeding, hepatorenal syndrome, alcoholic hepatitis, or acute liver failure as well as mortality after non-transplant surgery or TIPS. *Dig Dis Sci* 2011;56:977-987.

