JOURNAL OF VIRAL HEPATITIS
JVH

Journal of Viral Hepatitis, 2010, **17,** 854–858 doi:10.1111/j.1365-2893.2009.01252.x

Hepatitis C virus infection diagnosis using metabonomics

M. M. G. Godoy,[1] E. P. A. Lopes,[1] R. O. Silva,[2] F. Hallwass,[2] L. C. A. Koury,[1] I. M. Moura,[1] S. M. C. Gonçalves[2] and A. M. Simas[2] [1]*Departamento de Medicina Clínica, Hospital das Clínicas, Centro de Ciências da Saúde, Universidade Federal de Pernambuco, Recife, PE, Brazil; and* [2]*Departamento de Química Fundamental, Centro de Ciências Exatas e da Natureza, Universidade Federal de Pernambuco, Recife, PE, Brazil*

Received August 2009; accepted for publication October 2009

SUMMARY. Metabonomics based on nuclear magnetic resonance (NMR) can reveal the profile of endogenous metabolites of low molecular weight in biofluids related to disease. The profile is identified a 'metabolic fingerprint' like from the pathological process, why this metabonomics has been used as a diagnostic method. The aim of the present study was to apply metabonomics to identify patients infected with the hepatitis C virus (HCV) through an analysis of ^{1}H NMR spectra of urine samples associated with multivariate statistical methods. A pilot study was carried out for the diagnostic test evaluation, involving two groups: (i) 34 patients positive for anti-HCV and HCV-RNA and negative for anti-HBc (disease group); and (ii) 32 individuals positive for anti-HBc and negative for HBsAg and anti-HCV. The urine samples were analyzed through ^{1}H NMR, applying principal component analysis and discriminant analysis for classification. The metabonomics model was capable of identifying 32 of the 34 patients in the disease group as positive and 31 of the 32 individuals in the control group as negative, demonstrating 94% sensitivity and specificity of 97% as well as positive and negative predictive values of 97% and 94%, respectively, and 95% accuracy ($P < 0.001$). In conclusion, the metabonomics model based on ^{1}H NMR spectra of urine samples in this preliminary study discriminated patients with HCV infection with high sensitivity and specificity, thereby demonstrating this model to be a potential tool for use in medical practice in the near future.

Keywords: ^{1}H NMR, HCV, hepatitis C, metabonomics, spectroscopy.

INTRODUCTION

Hepatitis C virus (HCV) infection is a public health problem in both developed as well as developing countries [1]. Recent data indicate that HCV accounts for 70% of cases of chronic hepatitis and 30% of liver transplants performed in the developed countries [2,3].

In the majority of situations, it is difficult to establish the natural history of HCV infection because of its silent onset and its course with few symptoms. After contamination, some patients progress toward a cure of the virus, but most often, the infection evolves toward chronicity, with persistence of the antibody (anti-HCV) and viremia (HCV-RNA) [4–8].

Routinely used tests for both the screening and diagnostic confirmation of HCV infection are divided into two groups: enzyme immunoassay (EIA) for the detection of anti-HCV and molecular tests for the identification of HCV-RNA. The study of anti-HCV with third-generation EIA offers high sensitivity and specificity, achieving a correct diagnosis in approximately 97% of cases. The detection of HCV-RNA, which reflects the infection in course, is routinely performed using polymerase chain reaction (PCR) [9–12]. In practice, both the study of anti-HCV and HCV-RNA are carried out with blood samples, which requires the puncturing of a peripheral vein. Furthermore, the more prolonged time for PCR, the higher the potential risk of contamination and the high cost of the procedure limit its application in daily practice [13]. Thus, a simpler, faster, more inexpensive test would be important for identifying and screening patients infected with HCV.

In recent years, there has been increasing interest in metabonomics technology, which offers the capability of associating changes in the profile of metabolites in a given organic biofluid with environmental, pathological or toxicological aggression [14–17]. Nuclear magnetic resonance

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; DA, discriminant analysis; EIA, enzyme immunoassay; GGT, gamma-glutamyl transferase; HCV, hepatitis C virus; MNV, maximal normality value; NMR, nuclear magnetic resonance; PCA, principal components analysis; PCR, polymerase chain reaction; PLS, partial least squares; SIMCA, simple classification analysis.

Correspondence: Michele Maria Gonçalves de Godoy, Hospital das Clínicas, Bloco A, Departamento de Medicina Clínica, Av. Prof. Moraes Rego, s/n, Recife, PE, 50670-901, Brazil. E-mail: michelegodoy@oi.com.br

(NMR) spectroscopy has become one of the main tools for the analysis of such changes, as the spectrum can accurately demonstrate metabolites and their concentrations [18–23].

The profile of NMR spectra is complex, containing hundreds of information. For an efficient interpretation, the use of statistical tools is necessary, such as principal components analysis (PCA), partial least squares (PLS), discriminant analysis (DA) and simple classification analysis (SIMCA). The statistical chemometric treatment of spectral data allows analysis in a simple, fast, reliable manner [24–26].

Metabonomics can reveal organic alterations through NMR spectral patterns in biofluids, such as blood and urine or even tissues. Such patterns serve a 'metabolic fingerprint' like from the pathological process [27–30]. Therefore, clinical studies seek to identify biomarkers that undergo alterations when unbalance occurs. Papers on experiments and clinical trials that employ metabonomics to assess infections by *Schistosoma mansoni* and HIV, for example, have been recently published, thereby placing this tool in propaedeutic practice [30,31].

The aim of the present study was to identify patients with HCV infection through an analysis of the ^{1}H NMR spectra of urine samples associated with multivariate statistical data reduction tools.

MATERIALS AND METHODS

Study population

A pilot study was carried out for the assessment of the diagnostic test and was made up of two groups:

(a) patients positive for anti-HCV and HCV-RNA, and negative for anti-HBc (disease group),

(b) individuals positive for anti-HBc, and negative for HBsAg and anti-HCV (control group).

Between January and December 2007, patients with HCV infection who fulfilled the eligibility criteria were evaluated consecutively based on demand at the Hepatitis Outpatient Clinic at the Hospital from Universidade Federal de Pernambuco (UFPE). The individuals in the control group were sent from the blood bank to the clinic in question. The study received approval from the Ethics Committee of the UFPE Health Sciences Center.

Individuals in treatment for hepatitis C include those with chronic alcohol use in the previous 15 days, those with HCV infection with extrahepatic manifestations, such as diabetes mellitus, thyroiditis; arthritis or glomerulonephritis, and those with other illnesses, such as schistosomiasis, HIV/AIDS, auto-immune disease or malignant tumour were excluded from the study.

Eighty individuals were included, 48 were anti-HCV positive and 32 made up the control group. Among the 48 patients who were anti-HCV positive, 14 (29%) were excluded (in 10, the HCV-RNA study resulted as negative, and in four the anti-HBc was positive). Thus, 66 individuals were evaluated – 34 in the disease group and 32 in the control group.

Sample analysis

At the Central Laboratory of the Hospital (UFPE), 10 mL of peripheral vein blood and 50 mL of urine were collected from each participant. In the blood, serum levels of the following components were determined: alanine aminotransferase (ALT), aspartate aminotransferase (AST) and gamma-glutamyl transferase (GGT) – all by the automatic kinetic method (Abbott®). HBsAg, anti-HBs, anti-HBc and anti-HCV were determined using EIA (Abbott®). HCV-RNA in patients who were anti-HCV positive was determined using the PCR method.

The urine samples were analyzed using reagent strips for urine (Labstix®) at the Central Laboratory of the Hospital (UFPE) and sent to the Analytical Facilities of the Department of Fundamental Chemistry (UFPE) for NMR evaluation.

^{1}H NMR spectroscopy

The ^{1}H NMR spectra were acquired by the Varian Unit Plus spectrometer operating at 300 MHz, using NMR tubes of 5 mm. To 400 μL of *in natura* urine, was added 200 μL of buffer solution (Na_2HPO_4/NaH_2PO_4 – 0.2 mol/L). After homogenization, the ^{1}H NMR spectrum was acquired, using a sequence of pulses with presaturation of the water signal.

The following parameters were employed: spectral window of 5 kHz, acquisition time of 6.4 s, 90° pulse, saturation delay of 3.0 s and 32 scans. All spectra were processed using line broadening of 0.3 Hz, and the signal at δ 3.06 ppm, attributed to the methyl group of creatinine, was used as a reference for chemical shift. Distortions from the baseline and phase were corrected manually.

^{1}H NMR spectra (δ 0.0–10.0 ppm) were automatically reduced to 154 regions of equal width (δ 0.05 ppm), excluding the region between δ 4.2 and 6.4 ppm, to eliminate the water and urea signals. The spectrum was normalized, based on the intensity of each region using the *VNMR6.1C* software program (Varian, Inc.).

Statistical analysis and reduction of the ^{1}H NMR spectrum data

The chi-square test of Fisher's exact test was used for the frequency analysis between groups. The Student's *t*-test was used for comparisons between mean values. The level for the rejection of the null hypothesis was equal to or less than 1% ($P \leq 0.01$).

Principal component analysis was performed for the evaluation of the spectra, which resulted in 65 (n–1) principal components (PCs). This analysis was unable to separate the two groups, but the matrix produced by the PCs, and contributions from each sample were used for the DA.

The forward stepwise procedure was used for the selection of the variables ($P \leq 0.05$). Sixteen of the 65 PCs were selected, as these PCs had a significant association with the problem in question (disease and control groups). Two functions were constructed: *P* and *N* to classify the samples as positive and negative, respectively. The *F* test for the function was performed on 16 variables and with 49 ($n-17$) degrees of freedom. An *F* value of 9.632 was found, which means that the model achieved statistical significance ($P < 0.001$), as the $F_{16.49}$ value for the 95% confidence interval was 0.470.

The number of samples used (66) did not allow the separation of the calibration and test groups, as this would greatly reduce the number of degrees of freedom. Thus, the decision was made to perform a cross-validation of the metabonomics model. As such, 66 different models were constructed with the same 16 PCs. One sample was removed from the model at a time; a model was generated without it; and, finally, the removed sample was predicted from such a model.

The original data were standardized by calculating the mean and standard deviation per sample, subtracting the mean of the amplitude for each sample and dividing the result by the standard deviation. As 66 samples were evaluated, the matrix had 66 lines (cases) and 155 columns (154 of the NMR spectrum plus one, referring to the diagnosis).

RESULTS

Forty-seven per cent of the patients in the disease group and 78% of the individuals in the control group were men ($P = 0.02$). Mean age was 48 ± 14 years in the disease group and 41 ± 12 years in the control group ($P = 0.035$). Mean body mass index was 26.8 ± 3.1 kg/m^{-2} in the disease group and 26.5 ± 5.9 kg/m^{-2} in the control group ($P = 0.769$).

Table 1 displays the mean values of serum AST, ALT and GGT divided by the maximal normality value (MNV) of the patients. The mean enzyme values were significantly higher in the disease group.

Among the 34 urine samples from the disease group, the metabonomics model obtained from the NMR spectra identified 32 as positive (94%) and two as negative (6%). In addition, among the 32 samples from the control group, the model identified 31 as negative (97%). Table 2 demonstrates that the sensitivity of the test was 94% and specificity was 97%; the positive and negative predictive values were 97% and 94%, respectively; and accuracy was 95% ($P < 0.001$).

In the cross-validation of the metabonomics model, removing one sample at a time and reconstructing 66 models to test the sample excluded, the same results obtained in the model were found among the 34 samples of the disease group. Among the 32 samples in the control group, there was agreement in 28 from the 31 samples considered negative and the one sample considered positive in the metabonomics model. Accuracy of the cross-validation was 91%.

Figure 1 illustrates the classification of the urine samples based on the *P* function score of the metabonomics model of the 34 patients in the disease group and the 32 patients in the control group. There was a distinction between the groups, with the samples from the disease group achieving higher values for the *P* function, whereas the control group achieves lower values. The overlap region (grey zone) reveals two samples that should have been classified as positive and four that should have been classified as negative. Among the latter, the metabolic model classified three from the four correctly as negative, but the cross-validation did not achieve the same result.

DISCUSSION

Alterations in homeostasis trigger responses in the organism, which are reflected in endogenous metabolites in biofluids and consequent changes in NMR spectra [27,32].

Table 1 Mean values of serum levels divided by the maximal normality value (MNV) of the hepatic enzymes for 66 individuals – 34 patients with hepatitis C virus (HCV) infection (disease) and 32 with positive anti-HBc (control)

		Groups		
Enzymes	Mean ± SD	Disease	Control	*P*
N	66	34	32	
AST per MNV	1.2 ± 1.3	1.6 ± 1.6	0.7 ± 0.2	0.002
ALT per MNV	1.3 ± 1.5	1.7 ± 1.8	0.6 ± 0.2	0.001
GGT per MNV	1.3 ± 1.1	1.5 ± 1.2	0.8 ± 0.9	0.010

AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma-glutamyl transferase; MNV, maximal normality value.

Table 2 Results of urine samples based on scores obtained through the metabonomics analysis of the ^{1}H NMR spectra model of 66 individuals – 34 patients with hepatitis C virus (HCV) infection (disease) and 32 individuals with positive anti-HBc (control)

	Groups		
Model	Disease (%)	Control (%)	Total
Positive	32 (94)	1 (3)	33
Negative	2 (6)	31 (97)	33
Total	34	32	66

Sensitivity = 94%, Positive predictive value = 97%.
Specificity = 97%, Negative predictive value = 94%.
Accuracy = 95%, $P = 0.001$.

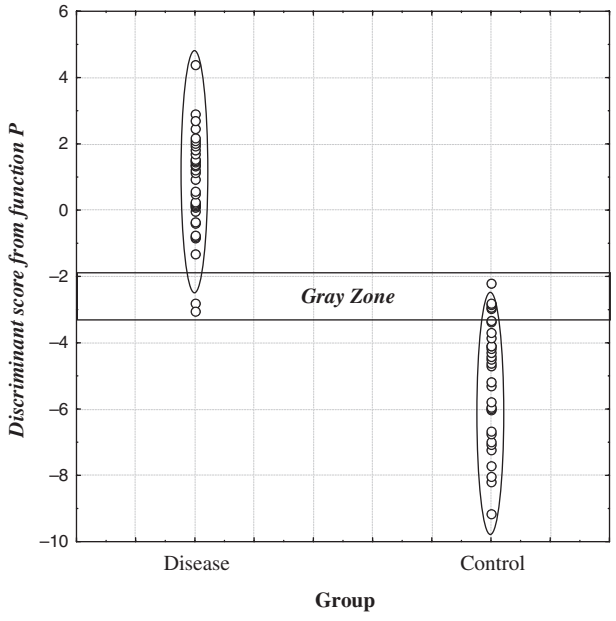

Fig. 1 Classification of urine samples based on scores obtained through the metabonomics analysis of the ^{1}H NMR spectra model of 66 individuals – 34 patients with hepatitis C virus (HCV) infection (disease) and 32 individuals with positive anti-HBc (control).

Metabonomics has been used in experiments with animal samples, whereas few studies have used tissues or biofluids from human beings so far. This occurs because of the greater facility in diet control, climatic conditions and age with animals, resulting in minimal variation [33,34]. Although variations are greater in human samples, metabonomics has been used in drug toxicity assessment and the illnesses diagnosis [35].

In the present study, demographic differences were found between the two groups, such as a lower frequency of men and greater mean age of the patients in the disease group. These findings may be related to the fact that control individuals came from the blood bank and were healthy young men. These differences, however, were nonsignificant ($P > 0.01$) and possibly did not affect the profile of the urine metabolites.

The findings of this pilot study reveal the potential of metabonomics for the fast, noninvasive diagnosis of HCV infection using a urine sample. All the patients in the disease group exhibited anti-HCV and progressed with viremia. However, the metabonomic model obtained from the NMR spectra must not have detected the antibody or the viral RNA. It was not the antibody, as the control group also exhibited IgG serum protein – anti-HBc. Further, it was not the RNA, because of its high molecular weight. The tests have registered alterations in the metabolic profile related to the hepatic aggression caused by the HCV. Indeed, the serum levels of the AST, ALT and GGT enzymes were greater in the disease group (Table 1), as the patients in this group were infected with a hepatotropic virus. Moreover, studies involving patients with HCV infection with normal liver function tests are underway.

Although the majority of studies are carried out with equipments that have large magnetic fields, such as NMR spectrometers operating at 600 MHz, the present study used a 300 MHz NMR, which was capable of discriminating the two groups with high sensitivity and specificity. A study evaluating patients with HIV/AIDS also used 300 MHz equipments [31]. Subsequent studies employing spectrometers with larger magnetic fields could provide greater resolutions and will be able to generate even better results when compared to these preliminary data.

Due to the complexity and amount of information obtained from the urine analysis, statistical methods were used for the interpretation of NMR spectroscopic data. Principal component analysis was unsuccessful in discriminating the two groups (control and disease). However, the results of the principal component analysis were used in the DA, with the selection of the 16 PCs that best classified the samples.

In conclusion, the present pilot study involving NMR spectrometry on urine samples for the diagnosis of HCV infection demonstrated its potential for application in clinical practice, having achieved 95% accuracy. In the cross-validation employed to check the metabonomics model, the test still continued to have a good performance. Further studies with larger numbers of patients, as well as with patients with hepatitis B and C and other noninfectious liver diseases, are underway. Moreover, the chemical structures of substances that produce differences in urine spectra of the two groups are being analyzed and should be revealed in upcoming studies.

This study was funded in part by the following Brazilian agencies: CNPq, FACEPE, FINEP.

REFERENCES

1 Alter MJ. Epidemiology of hepatitis C virus infection. *World J Gastroenterol* 2007; 13: 2436–2441.
2 Sherman M, Shafran S, Burak K *et al.* Management of chronic hepatitis C: consensus guidelines. *Can J Gastroenterol* 2007; 21: 25C–34C.
3 Davis GL, Albright JE, Cook SF, Rosenberg DM. Projecting future complications of chronic hepatitis C in the United States. *Liver Tranpl* 2003; 9: 331–338.
4 Seeff LB, Miller RN, Rabkin CS *et al.* 45-Year follow-up of hepatitis C virus infection in healthy young adults. *Ann Intern Med* 2000; 132: 105–111.
5 Lauer GM, Walker BD. Medical progress: hepatitis C virus infection. *N Engl J Med* 2001; 345: 41–52.
6 Seeff LB. The history of the ''natural history'' of hepatitis. *Liver Int* 2009; 29: 89–99.
7 Niederau C, Lange S, Heintges T *et al.* Prognosis of chronic hepatitis C: results of a large, prospective cohort study. *Hepatology* 1998; 28: 1687–1695.

8 Alter HJ, Seeff LB. Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. *Semin Liver Dis* 2000; 20: 17–35.
9 Lopes EPA, Granato CH, Lanzoni V *et al.* Evaluation of an enzyme immunoassay for hepatitis C virus antibody detection using a recombinant protein derived from the core region of hepatitis C virus genome. *Mem Inst Oswaldo Cruz* 2000; 95: 717–720.
10 Gretch DR. Diagnostic tests for hepatitis C. *Hepatology* 1997; 26: 43S–47S.
11 Ferreira-Gonzalez A, Shiffman ML. Use of diagnostic testing for managing hepatitis C virus infection. *Semin Liver Dis* 2004; 24: 9–18.
12 Pawlotsky JM, Lonjon I, Hézode C *et al.* What strategy should be used for diagnosis of hepatitis C virus infection in clinical laboratories? *Hepatology* 1998; 27: 1700–1702.
13 Zaaijer HL, Cuypers HT, Reesink HW, Winkel IN, Gerken G, Lelie PN. Reliability of polymerase chain reaction for detection of hepatitis C virus. *Lancet* 1993; 341: 722–724.
14 Robertson DG. Metabonomics in toxicology: a review. *Toxicol Sci* 2005; 85: 809–822.
15 Lindon JC, Nicholson JK, Holmes E, Everett JR. Metabonomics: metabolic processes studied by NMR spectroscopy of biofluids. *Concepts Magn Reson* 2000; 12: 289–320.
16 Robertson DG, Reily MD, Singler RE, Wells DF, Paterson DA, Braden TK. Metabonomics: evaluation of NMR and pattern recognition technology for rapid *in vivo* screening of liver and kidney toxicants. *Toxicol Sci* 2000; 57: 326–337.
17 Nicholson JK, Connelly J, Lindon JC, Holmes E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discover* 2002; 1: 153–161.
18 Moyec LL, Racine S, Toumelin PL *et al.* Aminoglycoside and glycopeptide renal toxicity in intensive care patients studied by proton magnetic resonance spectroscopy of urine. *Crit Care Med* 2002; 6: 1242–1245.
19 Brindle JT, Nicholson JK, Schofield PM, Grainger DJ, Holmes E. Application of chemometrics in ^{1}H NMR spectroscopic data to investigate a relationship between human serum metabolic profiles and hypertension. *Analyst* 2003; 128: 32–36.
20 Mäkinen VP, Soininen P, Forsblom C *et al.* Diagnosing diabetic nephropathy by 1H NMR metabonomics of serum. *Magn Reson Mater Phy* 2006; 19: 281–296.
21 Marchesi JR, Holmes E, Khan F *et al.* Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 2007; 6: 546–551.
22 Weljie AM, Dowlatabadi R, Miller BJ, Vogel HJ, Jirik FR. An inflammatory arthritis-associated metabolite biomarker pattern revealed by 1H NMR spectroscopy. *J Proteome Res* 2007; 6: 3456–3464.
23 Coen M, O'Sullivan M, Bubb WA, Kuchel PW, Sorrel T. Proton nuclear magnetic resonance-based metabonomics for rapid diagnosis of meningitis and ventriculitis. *Clin Infect Dis* 2005; 41: 1582–1590.
24 Lindon JC, Holmes E, Nicholson JK. Pattern recognition methods and applications in biomedical magnetic resonance. *Prog NMR Spectroscopy* 2001; 39: 1–40.
25 Keun HC. Metabonomic modeling of drug toxicity. *Pharm Therap* 2006; 109: 92–106.
26 Coen M, Holmes E, Lindon JC, Nicholson JK. NMR-based metabolic profiling and metabonomic approaches to problems in molecular toxicology. *Chem Res Toxicol* 2008; 21: 9–27.
27 Lindon JC, Holmes E, Nicholson JK. So what's the deal with metabonomics? *Anal Chem* 2003; 75: 384A–391A.
28 Odunsi K, Wollman RM, Ambrosone CB *et al.* Detection of epithelial ovarian cancer using 1H-NMR-based metabonomics. *Int J Cancer* 2005; 113: 782–788.
29 Holmes E, Tsang TM, Tabrizi SJ. The application of NMR-based metabonomics in neurological disorders. *NeuroRx* 2006; 3: 358–372.
30 Wang Y, Holmes E, Nicholson JK *et al.* Metabonomic investigations in mice infected with *Schistosoma mansoni*: an approach for biomarker identification. *Proc Natl Acad Sci* 2004; 101: 12676–12681.
31 Hewer R, Vorster J, Steffens FE, Meyer D. Applying biofluid ^{1}H NMR-based metabonomic techniques to distinguish between HIV-1 positive/AIDS patients on antiretroviral treatment and HIV-1 negative individuals. *Pharm Biom Anal* 2006; 41: 1442–1446.
32 Nicholson JK, Lindon JC, Holmes E. Metabonomics: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999; 29: 1181–1189.
33 Lenz EM, Bright J, Wilson ID, Morgan SR, Nash AF. A ^{1}H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. *J Pharm Biomed Anal* 2003; 33: 1103–1115.
34 Lenz EM, Bright J, Wilson ID *et al.* Metabonomics, dietary influences and cultural differences: a ^{1}H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J Pharm Biomed Anal* 2004; 36: 841–849.
35 Bollard ME, Stanley EG, Lindon JC, Nicholson JK, Holmes E. NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. *NMR Biomed* 2005; 18: 143–162.