Basic Study
Assessment of periportal fibrosis in *Schistosomiasis mansoni* patients by proton nuclear magnetic resonance-based metabonomics models

Rodrigues ML et al. Fibrosis assessment by metabonomics

Abstract
BACKGROUND
The evaluation of periportal fibrosis (PPF) is essential for a prognostic assessment of patients with schistosomiasis mansoni. WHO Protocol Niamey defines patterns of fibrosis from abdominal ultrasonography, $^1$H nuclear magnetic resonance (NMR)-based Metabonomics has been employed to assess liver fibrosis in some diseases.

AIM
To build $^1$H NMR-based metabonomics models (MM) to discriminate mild from significant periportal PPF and identify differences in the metabolite profiles.

METHODS
A prospective cross-sectional study was performed on schistosomiasis patients at a University Hospital in Northeastern Brazil. We evaluated 41 spectra of serum samples from 10 patients with mild PPF (C Niamey pattern) and 31 patients with significant PPF (D/E/F Niamey patterns). MM were built using partial least squares-discriminant analysis (PLS-DA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA) formalisms.
RESULTS

PLS-DA and OPLS-DA resulted in discrimination between mild and significant PPF groups with R2 e Q2 values of 0.80 and 0.38 and 0.72 and 0.42 for each model, respectively. The OPLS-DA model presented accuracy, sensitivity and specificity values of 92.7%, 90.3% and 100% to discriminate significant PPF. The metabolites identified as responsible by discrimination were: N-acetylglucosamines, alanine, glycolaldehyde, carbohydrates and valine.

CONCLUSION

The MMIs were able to discriminate mild from significant PPF patterns in patients with schistosomiasis mansoni through identification of differences in serum metabolites profiles.

Key Words: Metabolomics; Portal hypertension; Schistosoma mansoni; Biomarkers; Neglected disease; Nuclear magnetic resonance

Rodrigues ML, da Luz TPSR, Pereira CLD, Batista AD, Domingues ALC, Silva RO, Lopes EP. Assessment of Periportal Fibrosis in Schistosomiasis mansoni patients by Proton nuclear magnetic resonance-based metabonomics models. World J Hepatol 2022; In press

Core Tip: In this study, we demonstrated a metabolic signatures and metabolic pathway disturbances that allowed to discriminate mild from significant periportal fibrosis in 41 patients with Schistosomiasis mansoni. Partial least squares-discriminant analysis (PLS-DA) and OPLS metabonomics models provided a clear separation between the groups. PLS-DA model presented accuracy, R2 and Q2 values equals to 0.85, 0.80 and 0.38, respectively. While OPLS model showed R2 and Q2 values equals to 0.717 and 0.417, respectively. We also identified some metabolites responsible by
discrimination, which are associated to changes related to liver function and amino acids metabolism.

INTRODUCTION

Schistosomiasis is a neglected disease that still occurs around the world and affects about 240 million people in 78 countries\cite{1}. In Brazil, it is caused by *Schistosoma mansoni*. It has been considered an endemic disease in the state of Pernambuco, with cases reported in 102 of 185 cities\cite{2,3}.

Periportal fibrosis (PPF), known as Symmers’ fibrosis, is induced by helminth eggs deposition in the portal vein and its branches. This fibrosis can extend to the peripheral intrahepatic branches, without promoting hepatocyte necrosis, being one of the causes of non-cirrhotic portal hypertension\cite{4}.

Ultrasonography (US) scan is used for diagnosis and assessment of PPF by the Niamey-Belo Horizonte Protocol, the WHO Standard Protocol. This protocol classifies 6 PPF patterns from A (no-fibrosis) up F (very advanced fibrosis), plus mixed patterns, such as the D/C or E/C patterns\cite{5,6}.

Although the US exam enables PPF diagnosis and measurement, there are limitations for its use, including the inter-observer variation and the low sensitivity to the diagnosis of initial forms of the disease, especially if the examiner has no experience in applying the Niamey-Belo Horizonte protocol. Additionally, the device can be difficult to access in some poor regions. Due to these difficulties, alternative strategies are being studied. Some serum biomarkers, alone or in association (indexes scores), have been used for this purpose. Some authors reported an inversely proportional relationship between platelet count and PPF pattern, as well as a directly proportional relationship between liver enzymes serum levels and PPF patterns\cite{5,6}.

Metabonomics is an area of knowledge that uses multivariate statistical formalisms applied to spectra data of biofluids to obtain a multiparametric response to external stimuli, such as pathogens\cite{7,8}. The biofluid, properly stored, can be analyzed at a center distant from the collection site. Batista *et al.*\cite{9} used $^1$H nuclear magnetic resonance
(NMR)-based metabonomics for liver fibrosis assessment in patients with chronic hepatitis C. The method proved to be useful in the diagnosis of significant and advanced fibrosis in these patients. Gardini et al.\textsuperscript{[10]} developed the profile of the serum metabolome of patients with hepatocellular carcinoma, in early from advanced stages, and found a discrimination between groups. They found that 1H-NMR metabolomics profiling could discriminate early from advanced hepatocellular carcinoma. The multivariate statistical formalisms most commonly used in metabonomics assays are: principal components analysis (PCA), for exploratory analyses, since it does not depend on class information and investigates if there are outlier samples; and partial least square-discriminant analyses (PLS-DA) or orthogonal PLS-DA (OPLS-DA) which use the class information to build metabonomics models (MM) that discriminate among samples from different groups\textsuperscript{[11]}. In the present study, we aimed to build 1H NMR-based MM to discriminate mild from significant PPF in patients with schistosomiasis mansoni and identify differences in the profiles of the endogenous metabolites.

MATERIALS AND METHODS

Study design
This is a phase II diagnostic validation test, a cross-sectional study performed with adult patients who were diagnosed with schistosomiasis mansoni, aiming to assess PPF patterns by 1H NMR-based metabonomics.

Patients
Patients aged 18 years or over diagnosed with schistosomiasis mansoni were included from the Schistosomiasis Clinic of the Gastroenterology Service of the Hospital das Clínicas, Universidade Federal de Pernambuco (Recife, Pernambuco, Brazil), between March and December 2019. Schistosomiasis diagnosis was based on the clinical history of contact with water sources in endemic areas, report of previous treatment with praziquantel, associated with finding of PPF by US scan. Exclusion clinical criteria were: presence of fatty liver disease, cirrhosis or hepatocellular carcinoma, portal vein
thrombosis, HIV, hepatitis B or C virus coinfection, or history of drug-induced liver injury or alcohol.

*Abdominal ultrasonography scan*

All patients were submitted to US scan after overnight fasting of about 8 h, by the same examiner. According to the Niamey-Belo Horizonte Protocol, PPF pattern was defined as follows: C (peripheral fibrosis), D (central fibrosis), E (advanced fibrosis) and F (very advanced fibrosis) patterns. Patients without or with a doubtful PPF (A and B pattern) were excluded of the study. All US exams were performed using a US Siemens Acuson S2000 instrument equipped with a 6C1 Ultrasound probe (Siemens Medical Solutions, Mountain View, CA, United States).

*Liver function tests*

Blood samples were collected from a peripheral vein, after US scan. Serum was obtained after centrifugation (3500 rpm) using a Centurion-Laborline equipment. Liver function tests including alanine aminotransferase and aspartate aminotransferase, gamma-glutamyl transferase (GGT), alkaline phosphatase (ALP), lipid profile (total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol and triglycerides) and glycemia were carried out using Wiener Lab® kits in a Wiener Lab® autoanalyzer (Wiener Lab Group, Argentina). Part of the samples were stored at minus 40 °C until the NMR analysis.

*Statistic of the clinical and laboratory data*

To investigate the distribution of demographic and clinical or laboratory data between groups, univariate tests were performed using GraphPad Prism 6 software (GraphPad Software, Inc., La Jolla, CA, United States) with unpaired Student's *t*-test, Mann-Whitney, and Fisher’s exact as appropriate. A *P*-value < 0.05 was set as the level of statistical significance.


$^1$H NMR spectra and Metabonomics modelling

All $^1$H NMR spectra were recorded using a VNMRsys400 spectrometer operating at 400 MHz. After thawing, serum samples were prepared by mixing 400 µl of serum and 200 µl of D$_2$O into NMR tubes of 5 mm id. $^1$H NMR spectra were performed using a sequence of radiofrequency pulses with presaturation of the water signal hyphenated to the Carr-Purcell-Meiboom-Gill pulse sequence, which was employed as a T2 filter. The following parameters were used: spectral window of 6.4 kHz, saturation delay of 2.0 s, acquisition time of 1.704 s, 90° RF pulse, temperature of 27 °C, 88 cycles, tau equal to 0.0004 s, bigtau equal to 0.07 s and 128 scans. The line broadening used was 0.3 Hz. Baseline and phase distortions were corrected manually. The signal attributed to the methyl group of lactate (δ 1.33 ppm) was used as a chemical shift reference. Using MestreNova 9.0 software, the region between δ 4.004 and 0.772 ppm was binned into 808 bins (each 0.004 ppm-wide). The matrix was built with 41 rows (cases) and 809 variables (bins of $^1$H NMR spectra plus class variable), and then was submitted to multivariate analysis. The models based on PCA, PLS-DA and OPLS-DA were constructed using MetaboAnalyst online platform 4.0[12,13]. In the preprocessing step, each sample was normalized by sum (cumulative intensity of the spectrum). This was performed to compare the spectral data, avoiding problems with sample dilutions, for example[14]. In addition, data were pre-processed using autoscaling. The validation of the PLS-DA and OPLS-DA models was based on two methods: (1) the leave-one-out cross validation method (LOOCV), where the optimal number of latent variables for the PLS-DA model was determined, thus providing the basis for the computation of the predictive ability (Q2), determination coefficient (R2), and the classification accuracy of the model; and (2) the permutation test, which made 2000 permutations of the class label to verify the accuracy of metabonomics models. PLS-DA and OPLS-DA models provided a quantitative measure of the discriminating power of each spectral bin. Variable importance in the projection (VIP) score was used. VIP is a weighted sum of squares of the PLS loadings. These weights are based on the amount of explained variance of the dependent variable in each PLS dimension. A VIP score cut-off equal to
1 was used. Discriminatory signals were attributed to metabolites using HMDB platform and also based on the literature[15-18]. Accuracy, sensitivity and specificity values were obtained from a confusion matrix that was constructed considering classification of OPLS-DA model.

RESULTS

Forty-four patients were included, but three were excluded because their samples proved to be outliers. Thus, 41 patients with PPF were selected to the study, being 10 patients with C; 12 patients with D; 17 patients with E and two patients with F patterns, according to the Niamey-Belo Horizonte Protocol. These patients were divided into two group: mild PPF (C pattern) and significant PPF (D/E/F patterns)[19]. Table 1 shows clinical and demographic data of the patients.

Figure 1 shows a typical 1H NMR spectrum of serum obtained in the study with assigned peaks. Exploratory analyses, by PCA, failed to indicate separation between the groups (data not shown). Thus, MM were developed using supervised methods: PLS-DA and OPLS-DA formalisms. Figure 2 shows (A) score plot and (B) performance of MM constructed using PLS-DA formalism. Regarding accuracy, the best performance was achieved when four latent variables were used, resulting in accuracy, R2 and Q2 values equal to 0.85, 0.80 and 0.38, respectively.

Figure 3 shows (A) variables importance for projection (VIP) score plot and (B) permutation test from PLS-DA model which presented P-value equal to 0.0245 after 2000 classes permutations. The spectral region responsible for discrimination was between δ 1.975 and δ 2.011 ppm, which is attributed to the methyl group of N-Acetylglucosamines. The serum level of N-Acetylglucosamines is higher in Significant PPF group than in Mild PPF group. In addition to this region, three more discriminatory bins can be observed—δ 3.544; 3.692; and 3.808 ppm, which were assigned to carbohydrates. According to the VIP score plot, the serum level of these carbohydrates is higher in the Mild PPF group. The VIP score full table presents other discriminatory bins, such as δ 1.502 ppm (VIP score = 1.98) and δ 3.492 ppm (VIP score
= 1.83), which were assigned to alanine and glycolaldehyde, both higher in the Significant PPF group.

Figure 4 shows (A) score plot and (B) results of permutation test of metabonomics models using OPLS-DA formalism. The OPLS-DA MM presented R2Y and Q2 values equal to 0.717 and 0.417, respectively, with P-values from permutation test less than 0.01. We identified four discriminatory bins, as follows: δ 1.030 ppm, assigned to valine; δ 1.046 ppm (not assigned); δ 1.446 ppm attributed to alanine; and δ 3.692 ppm assigned to carbohydrates. The serum levels observed for valine, carbohydrates and for the unidentified metabolite (signal at δ 1.046 ppm) were higher in the Mild PPF group, while the alanine serum level was higher in the Significant PPF group.

Table 2 presents a summary of metabolites that differentiated the Mild PPF samples from the Significant PPF samples, as well as the chemometric formalism employed, the chemical shift (δ) of each metabolite and the group in which these metabolites had higher serum level.

Table 3 shows the confusion matrix obtained from the OPLS-DA MM. Accuracy, sensitivity, specificity, positive predictive and negative predictive values are equal to 92.7%, 90.3%, 100%, 100% and 76.9%, respectively.

DISCUSSION

PPF is a mark of schistosomiasis mansoni disease. Assessment of PPF intensity is crucial to determine disease morbidity and prognosis. In addition, significant PPF is associated with non-cirrhotic portal hypertension and its consequences. Generally, US scan and serum biomarkers have been used for PPF assessment in schistosomiasis patients[220]. Among serum biomarkers for PPF assessment, liver enzymes and platelet count alone or combined with ALP, as in the Coutinho-index, have been given importance in the literature[247]. In agreement with these authors, laboratory data found in this study indicate a higher serum level of GGT in particular, and lower platelet count in patients with significant PPF. Köpke-Aguiar et al[221] also reported higher serum levels of GGT in
the more severe cases. Pereira et al[22] and Lambertucci[23] also reported lower platelets count, as well as increased spleen size in more severe schistosomiasis mansoni disease. Significant PPF in the schistosomiasis patients is induced by increase in number of eggs in the intrahepatic portal veins due repeated infections and by the exacerbated host’s immune response[24]. The presence of adult worms in mesenteric and portal vessels, as well as the presence of their eggs promotes immunological stimulation and induces primary splenomegaly by reticuloendothelial system hyperplasia, which leads to pancytopenia by hypersplenism[19,22]. Splenomegaly also causes blood hyperflow in the splenic vein, which contributes to presinusoidal portal hypertension[25]. In addition, the changes in liver hemodynamics, triggered by portal hyperflow, could promote increase of GGT serum level, as observed in our study[26].

Usually, metabonomics studies begin with exploratory analyses by PCA. However, this initial analysis failed to discriminate between the groups. Therefore, discriminant analyses formalisms were employed, resulting in efficient separation between mild PPF and significant PPF groups. Serum levels of valine (δ 1.030 ppm) and alanine (δ 1.446 and 1.502 ppm) observed in the spectral data suggest that the changes in portal vein flow could trigger disorders of amino acids metabolism in hepatocytes, as reported by Li et al[27]. Changes in amino acid serum levels were important for discrimination in both the PLS-DA and OPLS-DA MM. These findings are in agreement with studies that correlate changes in serum levels of these amino acids in response to liver fibrogenesis caused by schistosomiasis in mice[15,28]. Balog et al[29] also reported an association between valine and alanine serum levels with the disease progression.

An increase of N-acetylg glucosamines serum level was observed in the significant PPF group, while the carbohydrate serum level was higher in the mild PPF group. Glucosamines are products of glucose metabolism, which are capable of suppressing the production of metalloproteinases. Therefore, the N-acetylg glucosamines and glucose serum levels observed could be associated with liver damage, which requires glucose consumption and production of N-acetylg glucosamines[30].
In clinical practice, the monitoring of patients with schistosomiasis mansoni is done by US scan. Hence, it is necessary to transfer patients from rural zone to hospital unit or to bring the device to the field. In addition, an experienced examiner is necessary, since the US scan is operator-dependent.

The metabolic profile presented in this study can be strategic for monitoring the patients in endemic regions, through blood samples collected and transported to a reference laboratory. $^1$H NMR-based metabolomics produce a “metabolic fingerprint”, providing systemic metabolic information about patients. It can help to identify those with more severe forms of schistosomiasis. The main limitation of the study was the sample size and the disproportionate among PPF standards.

In the present study, we used $^1$H NMR-based metabolomics from the serum of patients with schistosomiasis mansoni to discriminate those with highest intensity of PPF. Moreover, the chemometric formalisms used enabled the identification of some metabolites associated with the discrimination, such as: alanine, glycolaldehyde and N-acetylglucosamines, which presented higher serum levels in the significant PPF group; while valine and carbohydrates presented lower serum levels in the most severe cases.

CONCLUSION
The $^1$H NMR-based metabolomics models were able to discriminate mild from significant PPF patterns in patients with schistosomiasis mansoni through identification of differences in serum metabolites profile. We intend to expand the study in the coming years in order to confirm the results and best understand the metabolic pathways associated to observed discrimination.

ARTICLE HIGHLIGHTS
Research background
Classification of the pattern of periportal fibrosis (PPF) is essential in the prognostic evaluation of patients with schistosomiasis mansoni.
Research motivation

There is a need for novel methods minimally invasive and new biomarkers for the diagnosis schistosomiasis mansoni.

Research objectives

To develop metabolic models, based on $^1$H nuclear magnetic resonance spectra, that allow the classification of the pattern of PPF and its associated metabolites in patients with schistosomiasis mansoni.

Research methods

Metabonomics models (MMs) were built to differentiate requirements with mild PPF and significant PPF. An analysis of the performance of MMs was performed for the prediction of PPF, using ultrasonography as a reference standard and the description of the main metabolites present in each PPF group and their relationship with serum markers.

Research results

The partial least squares-discriminant analysis (PLS-DA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA) formalisms discriminated spectral regions between the groups as follows: carbohydrates and valine, more concentrated in those of the group with level FPP; N-Acetylglucosamines, Alanine, Glycolaldehyde more concentrated in the samples of the group with significant PPF. OPLS-DA showed accuracy, sensitivity, specificity, were equal to 92.7%, 90.3%, 100%, for the diagnosis of significant PPF.

Research conclusions

The constructed MMs were able to discriminate between mild and significant FPP in patients with schistosomiasis with good accuracy.
Research perspectives

This technique will be able to detect even low-intensity infections, overcoming the limitations of current diagnostic techniques, with the use of a single serum sample. These models can be inserted in the propaedeutic arsenal in clinical practice for the measurement of PPF in remote areas.