Correlating the amount of urea, creatinine, and glucose in urine from patients with diabetes mellitus and hypertension with the risk of developing renal lesions by means of Raman spectroscopy and principal component analysis

Jeyse Aliana Martins Bispo
Elzo Everton de Sousa Vieira
Landulfo Silveira, Jr.
Adriana Barrinha Fernandes
Correlating the amount of urea, creatinine, and glucose in urine from patients with diabetes mellitus and hypertension with the risk of developing renal lesions by means of Raman spectroscopy and principal component analysis

Jeyse Aliana Martins Bispo,a,b Elzo Everton de Sousa Vieira,a,b Landulfo Silveira Jr., and Adriana Barrinha Fernandes

Abstract. Patients with diabetes mellitus and hypertension (HT) diseases are predisposed to kidney diseases. The objective of this study was to identify potential biomarkers in the urine of diabetic and hypertensive patients through Raman spectroscopy in order to predict the evolution to complications and kidney failure. Urine samples were collected from control subjects (CTR) and patients with diabetes and HT with no complications (lower risk, LR), high degree of complications (higher risk, HR), and doing blood dialysis (DI). Urine samples were stored frozen (−20°C) before spectral analysis. Raman spectra were obtained using a dispersive spectrometer (830-nm, 300-mW power, and 20-s accumulation). Spectra were then submitted to principal component analysis (PCA) followed by discriminant analysis. The first PCA loading vectors revealed spectral features of urea, creatinine, and glucose. It has been found that the amounts of urea and creatinine decreased as disease evolved from CTR to LR/HR and DI (PC1, p < 0.05), and the amount of glucose increased in the urine of LR/HR compared to CTR (PC3, p < 0.05). The discriminating model showed better overall classification rate of 70%. These results could lead to diagnostic information of possible complications and a better disease prognosis. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.

Keywords: Raman spectroscopy; hypertension; diabetes mellitus; renal complications.

Paper 130331PR received May 9, 2013; revised manuscript received Jun. 27, 2013; accepted for publication Jul. 8, 2013; published online Aug. 8, 2013; corrected Apr. 26, 2016.

1 Introduction

Diabetes mellitus (DM) is a heterogeneous group of chronic and degenerative metabolic disorders characterized by chronic hyperglycemia related to changes in the metabolism of carbohydrates, proteins, and fats as a result of defects in the insulin secretion, action (peripheral resistance in target tissues), or both.1 Hypertension (HT) is a chronic, multifactorial, in most cases asymptomatic disease that compromises the balance of the mechanisms involved in vessel vasodilatation and vasoconstriction, leading to increased blood pressure and capable of compromising tissue irritation and ultimately damaging irrigated organs. According to the World Health Organization (WHO), ~171 million people were diabetic in 2011, a number that will increase to ~366 million in 2030.2 Also, the WHO stated that there were about 600 million hypertensives worldwide in 2011.1 HT affects about 25% of the Brazilian population, reaching >50% in the elderly and surprisingly, 5% of the 70 million children and adolescents in Brazil are hypertensives.3

DM and HT are among the most common diseases in industrialized countries and the frequency of these diseases in the populations increase with age. It is estimated that 35% to 75% of the complications affecting diabetic patients can be attributed to HT since its prevalence is particularly higher in type 1 diabetic patients with clinical nephropathy and is already present in the pre-proteinuric phase in type 2 diabetes.4 Long-term complications of DM include retinopathy with potential loss of vision, nephropathy leading to renal failure, peripheral neuropathy with risk of foot ulcers and amputations, cardiovascular symptoms, and sexual dysfunction. Patients with diabetes have an increased incidence of cardiovascular and peripheral vascular atherosclerosis, cerebrovascular disease, HT, and abnormal lipoprotein metabolism.5

The Diabetes Control and Complications Trial (DCCT) and other studies in type 1 diabetic patients provided conclusive evidence that strict glycemic control in type 1 diabetes delays the onset of diabetic retinopathy.6 Although it is often claimed that these studies also demonstrated delay in the progression of diabetic nephropathy,7 a critical reading of the actual data demonstrates that strict glycemic control at the very best may reduce proteinuria or stabilize the development rate of renal lesion.8

HT is virtually always present in persons with end-stage diabetic renal disease and also contributes to focal sclerosis in diabetic laboratory animals.9 HT has been shown to be detrimental in essentially all forms of progressive renal disease, contributing to the progression of renal insufficiency. In case of permanent renal damage, the most appropriate treatment is hemodialysis. Removing the harmful wastes and extra salt and fluids helps...
to control blood pressure and keep the proper balance of chemicals like potassium and sodium in organism. The glomerular filtration rate (GFR) is an important component in the diagnosis and classification of chronic kidney disease. Clinically, the most used method for obtaining information on GFR is the 24-h urinary creatinine clearance, in which 24-h urinary creatinine excretion is divided by the serum creatinine concentration.

Raman spectroscopy is an optical technique that allows the analysis of urine samples and can be used to detect single or multiple urine components. Using dispersive or Fourier-Transform Raman techniques in the near-infrared spectrum, there is no need for additional chemical steps for analysis (separation, dilution, or mixture of other reagents) and may prove superior to current methods of testing urine nondestructively. Biochemical assays based on Raman spectroscopy could be used for testing body fluids such as blood, blood components, and metabolites in the serum for doped control. Detecting antibodies in cat’s serum and even monitoring heparin levels in blood during surgeries. Optical techniques may become a future alternative or even replace existing laboratory methods.

Urine samples may be obtained easily and in larger amounts than blood. The urine test provides diagnostic information about metabolic diseases (diabetes), urinary tract infections, and other diseases. McMurdy and Berger reported the first use of Raman spectroscopy to measure creatinine concentrations in unaltered urine samples from a multipatient population, with error of cross-validation of 4.9 mg/dL, compared to the error of the reference chemical method which was 1.1 mg/dL. Premasiri et al. used Raman spectroscopy and surface-enhanced Raman spectroscopy (SERS) to analyze the components present in human urine such as total nitrogen compounds, urea, creatinine, and rate of excretion (urea/creatinine). Wang et al. proposed a method for measuring the creatinine concentration in urine based on SERS-nanostructured substrates prepared without lithography. Park et al. used low-resolution Raman spectroscopy (785 nm) to detect minute amounts of glucose in diluted (10-fold) urine, with accuracy of 92% to classify abnormal (8 mg/dL) and normal urine samples according to their glucose concentrations.

The standard urine test for analysis of creatinine presents, as main drawback, the need for collecting urine samples for 24 h, being uncomfortable for the patients. The main advantages of Raman spectroscopy in detecting changes of urea and creatinine concentrations rely in the fact that urine sample can be evaluated in real time, rapidly and without need of reagents when compared to standard biochemical assay methods, so the analysis could be done in each sample collected along the day and even in a single collection analysis, bringing comfort to the patient and perhaps accuracy to the results.

The aim of this study was to identify, through Raman spectroscopy, the potential biomarkers (i.e., urea, creatinine, and glucose) in the urine of diabetic and hypertensive patients who have or do not have renal complications associated with these pathologies (severe renal damage hemodialysis, cardiac disease, cerebrovascular disease, and peripheral vascular disease) compared to normal (neither DM nor HT) patients, by evaluating the Raman spectral signature from each biomarker measured in the urine samples. Moreover, the Raman spectral information from these biomarkers were correlated to the degree of complications, in order to predict the disease evolution, by implementing a discriminating model based on principal component analysis (PCA) technique, where the most relevant spectral information from the dataset were extracted by PCA, and a discriminant analysis (DA) used to correlate the disease status with the spectral variations provided by PCA, thus obtaining a discrimination in classes according to the complications.

2 Materials and Methods

This study was approved by the Research Ethics Committee from UNICASTELO (protocol no. 8926). A total of 70 patients (40 women and 30 men) were enrolled, which were divided as follows: 18 normoglycemic and normotensive patients (CTR), 20 DM and HT patients with no apparent complications (low risk of renal disease, LR), 16 DM and HT patients who have other complications than diagnosed renal failure (high risk of renal disease, HR), and 16 DM and HT patients who have renal failure and are being submitted to blood dialysis (DI). The overall average age was 60.8 ± 9.0 years and the average age for each group was 61.2 ± 9.8, 65.2 ± 7.7, 60.0 ± 6.8, and 46.7 ± 14.1 years for CTR, LR, HR, and DI, respectively. Urines from single collection were obtained in the morning in fasting, bottled and stored in −20°C freezer until biochemical and spectral analysis. For spectroscopy, urine samples were unfrozen to reach room temperature and placed in an aluminum holder with a vessel of about 100 μL. Spectra were taken in the vessel by means of a Raman probe connected to a dispersive Raman spectrometer. The spectrometer (Lambda Solutions, MA, model P-1 Raman) is composed of a diode laser (830 nm) coupled to a Raman probe (Lambda Solutions, MA, model Vector Probe) that is used to illuminate sample and collect the scattered light. The probe is connected to a spectrophotograph with a Peltier-cooled, deep depleted/back illuminated charge-coupled device camera (−75°C), which collects high-resolution Raman spectrum from the sample in the fingerprint region (400 to 1800 cm−1). The laser power was adjusted to 300 mW and the integration/accumulation time to collect the Raman signal was set to 20 s. Triplicate spectra were obtained from each sample, which were averaged after preprocessing. For comparison purposes, the reference spectra of organic components of urine were obtained: glucose (Sigma-Aldrich Brasil, São Paulo, Brazil, ref. 49163), urea (Bio-Rad Laboratórios Brasil, Rio de Janeiro, Brazil, ref. 161–0730), and creatinine (Sigma-Aldrich Brasil, São Paulo, Brazil, ref. C4255).

It has been found that the integrity of the biochemical components of urine were maintained since no heating or burning of urine sample was observed. Moreover, a biochemical assay for glucose (UriGold strips, Gold Analisa Diagnóstica, MG, Brazil), urea, and creatinine (Urea CE and Creatinine colorimetric test, Labtest Diagnóstica, MG, Brazil) was conducted in the same urine that would be analyzed by Raman spectroscopy for correlation between Raman features and biochemical outcomes.

Spectra were preprocessed to remove the undesired background fluorescence using a seventh-order polynomial fitted over the spectral range of 600 to 1800 cm−1 and subtracted from the gross spectrum, as described elsewhere and normalized to the intensity of the Raman peak at 1640 cm−1 (weak Raman band of water), averaged and plotted with the aim of identifying spectral differences between the CTR and the three DM and HT groups that could be related to the renal disease status. One spectrum from LR group was withdrawn due to the poor signal-to-noise ratio.
After preprocessing, the spectra dataset was submitted to PCA, which is a multivariate statistical technique that transforms a set of original, correlated variables (spectra) to a new set of uncorrelated variables called principal components (PCs) (PC vectors and scores) based on the maximum variance. PCA can be used to group spectra according to their similar variances since each PC is orthogonal to each other, so a unique spectral characteristic is present in each component. The first PC vectors bring the most relevant spectral features presented in the dataset, whereas the PC scores bring the intensity of each PC vector in each spectrum of the dataset. Then one could use the scores to discriminate spectra from individuals of a population according to variation of the spectral characteristics (using a suitable DA), or use them to correlate the spectral information to a biochemical parameter of the sample (using a regression line). PCs were calculated from the average spectra using MATLAB 7.0.

DA is a classification method. It assumes that different classes generate data based on different Gaussian distributions (normally distributed). It undertakes the same task as multiple linear regression by predicting an outcome Y from weighted combinations of the predictors X but when the predictors X will, through the regression equation, produce categorical estimates Y instead of numerical ones (as done in multiple linear regression). First, the fitting function estimates the parameters of a Gaussian distribution for each class. Then, the trained classifier finds the class with the smallest misclassification cost.

DA involves the determination of a linear equation like regression that will predict which group the case belongs to. The form of the equation or function is

\[ D = v_1X_1 + v_2X_2 + \ldots + v_nX_n + c, \]  

where D is the discriminate function; v is the discriminant coefficient or weight for that variable; X is the respondent’s score for that variable; c is a constant; and n is the number of predictor variables. The objective of this function is to maximize the distance between the categories, i.e., come up with an equation that has strong discriminatory power between groups, where the v is calculated in order to maximize the distance between the means of the criterion (dependent) variable.

We used the “classify.m” function from Matlab’s statistics package, where the X independent variables were the PC scores of the four clinical groups, being the most relevant scores identified by calculating the analysis of variance (ANOVA) (5% significance level) and choosing the ones with lower significance. Several discriminant functions (classifiers) have been tested among the ones allowed by the “classify” function, such as linear, quadratic, and Mahalanobis. The better classifier was selected among the ones with higher discrimination capability using the first four PC scores.

3 Results

Figure 1 presents the Raman spectra of urine samples from the four groups evaluated by Raman spectroscopy (group 1: control, CTR; group 2: low risk, LR; group 3: high risk, HR; and group 4: dialysis, DI). For comparison purposes, Fig. 2 also presents the reference spectra of the organic components present in higher concentration in urine: glucose, urea, and creatinine, where the Raman bands of these compounds are clearly seen in the spectra of urine. These peaks have been reported in the recent literature. Being urea peaks at 1004 cm\(^{-1}\) (symmetrical C─N stretch) and 1161 cm\(^{-1}\) (attributed to NH\(_2\) modes), creatinine peaks at 608 cm\(^{-1}\) (N─CH\(_3\) stretching, C≡O deformation and ring vibrations), 680 cm\(^{-1}\) (C─NH\(_2\) and C≡O stretching, ring vibrations), 846 cm\(^{-1}\) (C─NH\(_3\) deformation and ring vibrations) and 910 cm\(^{-1}\) (C─C─N stretching), and glucose peak at 1128 cm\(^{-1}\) (C─O stretching). The urine spectra showed differences in the intensities of several peaks, indicating difference in the concentration of the urine’s biochemical components in each group. Differences were seen in the intensity of specific peaks depending on the group, such as the absence of the peak of glucose (1128 cm\(^{-1}\)) for the CTR and its presence in all disease groups, and the decreased intensity of the peaks of urea (1004 cm\(^{-1}\)) and creatinine (680 cm\(^{-1}\)) for the diseased groups compared to CTR.

The PCA was applied to the spectra of urine, aiming the reduction of dimensionality of the dataset by concentrating the spectral variation on the first PCs, and using these first PCs loading vectors to find differences in the biochemical constitution of the urine of CTR, LR, HR, and DI groups related to the degree of kidney complications. It has been found that the first four PCs accounted for >98% of all the spectral variance (PC1 = 89.4%, PC2 = 4.6%, PC3 = 2.9%, PC4 = 1.3%). The plot of the first PC vectors (Fig. 2 left) indicated that PC1 presented spectrochemical characteristics of creatinine and urea (peaks at 680, 846, 1004, and 1161 cm\(^{-1}\)); PC2 indicated the spectrochemical characteristics of urea (peak at 1004 cm\(^{-1}\)) and remnant background fluorescence; PC3 and PC4 showed marked features of glucose (main peak at 1128 cm\(^{-1}\)).

PC scores were averaged for each group and the first four were plotted in Fig. 2. These scores represent the intensities of the first four PC vectors in each urine spectrum, so the intensity could be correlated to the amount of each biochemical presented in each group. ANOVA indicated that PC1, PC2, and PC3 presented significant differences among columns (p < 0.05). PC1 showed a statistically significant difference in the LR and HR groups compared to CTR and DI group (p < 0.05), indicating a decrease of urea and creatinine concentration in the patients with complications including renal failure. Despite the nonsignificance (p > 0.05), the HR group showed a decreased amount of urea and creatinine compared to the LR group in PC1 plots. PC2 followed the same profile of PC1 as it presented (negative) spectral features of urea but with
significance between CTR and HR/DI, and LR and DI. PC3 vector, which presented (negative) spectral features of glucose, showed to be increased in the urine of patients of LR and HR compared with CTR group \((p < 0.05)\). Additionally, the LR group presented higher concentration of glucose when compared to the HR group \((p < 0.05)\). Despite that nonsignificance was found, the DI group presented higher concentration of glucose than CTR and lower than LR and HR groups. It was found that PC4 scores do not present statistically significant differences between the groups, despite the presence of peaks related to glucose.

A discriminant model has been developed based on DA to classify spectra of urine samples in one of each clinical group based on the PCA scores. PCA can be used to group spectra according to their similar variances since each PC is orthogonal to each other, so a unique spectral characteristic is present in each component. The PC scores carry the intensity of each PC vector of each spectrum of the dataset. Then one could use the scores to discriminate spectra from individuals of a population according to variation of the spectral characteristics (using a suitable DA), or use them to correlate the spectral information to a biochemical parameter of the sample (using a regression line). The discriminant model used was the “quadratic discriminant analysis” (QDA) under the Matlab classify function. Unlike linear discriminant analysis (LDA), however, in QDA, there is no assumption that the covariance of each of the classes is identical. This leads to a quadratic decision surface instead of the linear when the Gaussians for each class are assumed to share the same covariance matrix.

Figure 4 presents the results of the classification using QDA, where integers from 1 to 4 in the ordinate axis represent the class number for each group (1, CTR; 2, LR; 3, HR; and 4, DI), where 70% of correct overall classification applied to the \(X\) variables: PC1, PC3, and PC4. Table II summarizes the classification of the QDA model in each group and the number of correct classification for each group.
4 Discussion

Raman spectroscopy is an optical technique with advantages over traditional biochemical techniques and may be used in future for urinalysis and to predict the complications arising from diabetes and HT in patients without signs of complications, with advantages of rapid analysis in single samples. It is critical that future research using Raman spectroscopy with biological material may be performed and compared with the conventional techniques that have already been adopted for analysis of urinary glucose, urea, and creatinine.

PC1 score (Fig. 2) showed that the amount of urea and creatinine decreased in the groups with diabetes/HT (LR and HR) and even more in the group submitted to DI; this indicates a decrease of urea and creatinine concentration in the patients with complications, being found that urea is a marker that has a strong importance in the metabolic changes induced by DM and HT diseases. The absence of the peak of glucose ($1128 \text{ cm}^{-1}$) for the CTR and its presence in all other disease groups suggests that diabetic patients are not receiving adequate therapy because of this glucosuria. Despite the nonsignificance of PC4 (glucose features), its inclusion in the LDA model helped the discrimination of the normal, CTR group.

Despite nonsignificance, the HR group showed a decreased amount of urea and creatinine compared to the LR group. These results may suggest that patients in the LR group, which apparently had no complications, are evolving for a higher risk of a serious renal disease. Thus, the therapy used to control DM and HT needs to be adjusted to become more effective.

PC3 (glucose features) were shown to be increased in groups LR and HR compared to CTR group, and the LR group presented higher concentrations of glucose when compared to the HR group. Despite the nonsignificance, the DI group presented higher concentrations of glucose than CTR and lower...
than LR and HR groups. These data indicate that both LR and HR groups are not receiving adequate treatment since the presence of glucose in urine generally reflects the inability of the tubule to retain the tubular glucose due to the specific lesion. The renal threshold (the ability of the tubules to reabsorb glucose), which is called as renal glycosuria, can be associated with other functional disorders of the tubular cells.

According to Basto and Kirsztajn, optimal management of chronic kidney disease (CKD) is based on three pillars: (1) early diagnosis of disease, (2) immediate referral for nephrological treatment, and (3) implementation of measures to preserve renal function. The early diagnosis of disease is frequently difficult because of the absence of symptoms in patients in the early stages of CKD, thus requiring that clinicians may suspect all patients, especially for those with risk factors for CKD.

Functional change in the GFR is an important characteristic in the diagnosis and classification of CKD. Even with reduced number of nephrons, the stable or near-normal GFR can be explained by the occurrence of increased filtration pressure or glomerular hypertension. This is sometimes observed in early diabetic nephropathy, where the GFR is increased up to 40% above the normal value.

There is no evidence that the GFR is lessened by any kind of strict glycemic control. On the other hand, HT has been found to be a major factor in the prediction of progression of diabetic nephropathy along with microalbuminuria and hyperglycemia.

The best way to evaluate the GFR is by determining the clearance of exogenous substances such as inulin, $^{125}$I-iothalamate, ethylenediaminetetraacetic acid, technetium-labeled diethylene triamine pentaacetic acid, or iohexol. As these compounds are excreted from the body via glomerular filtration, and they do not undergo further secretion and/or reabsorption when passing through the renal tubules, they are considered an ideal filtration marker. Since these exogenous substances need to be infused, the evaluation of these clearances is challenging, requiring a long time for the procedure, and has been restricted to specific pathological conditions or research, where more simple clearance techniques do not provide sufficient information to guide the medical decision.

In clinical practice, the GFR is assessed by measuring specific endogenous biomarkers such as urea and creatinine. Urea is not completely reliable since its levels are susceptible to changes due to other conditions not related to GFR. For instance, a high protein diet, intense gastrointestinal hemorrhage, tissue breakdown, and therapy with corticosteroids can lead to an increase in plasma urea, whereas a low protein diet and liver disease can lead to its reduction. Also, since 40% to 50% of filtered urea may be reabsorbed by the tubules, this proportion is reduced in advanced renal failure.

The most used method for obtaining information on GFR is the creatinine clearance in 24-h urine, in which the urinary creatinine excretion is divided by the serum creatinine concentration. The creatinine clearance does not present all the criteria for an ideal marker of GFR since creatinine is excreted via glomerular filtration and also via secretion in the proximal tubule. The major problem with creatinine clearance is the requirement for urine collection over 24 h, where the collections are often inaccurate, particularly in some clinical situations (elderly patients, cognitive impairment). At present, determination of GFR by creatinine clearance is recommended even in extremes of age and body size, obesity, severe malnutrition, vegetarian diet, disease of skeletal muscle, paraplegia or quadriplegia, acute renal failure, and adjustment of dosage of potentially nephrotoxic drugs.

The course of CKD is often asymptomatic until the disease reaches its advanced stages, usually diagnosed when the patient seeks medical attention after presenting one or more disease complications and/or other comorbidities such as DM and HT. It remains unclear which patients with CKD will progress to end-stage renal disease and which ones are at greater risk of needing renal dialysis. However, it is important that interventions would be implemented earlier for stabilizing the progression of renal disease and preventing the occurrence of end-stage renal disease. Furthermore, it is important to emphasize that successful treatment of the underlying disease, such as DM and HT, is also needed for preventing the end-stage renal disease.

Early diagnosis and immediate nephrology referral are key steps in management because enable predialysis education, allow implementation of preventive measures that delay or even halt progression of CKD to end-stage renal disease, as well as decrease initial morbidity and mortality. Thus, Raman spectroscopy is a rapid and reliable method that can be useful in future diagnosis of complications from DM and HT by evaluating the amount of urea, creatinine, and glucose in single urine samples and using them to prevent complications.

5 Conclusion

Dispersive near-infrared Raman spectroscopy proved to be a promising tool for analysis of renal biomarkers (urea, creatinine, and glucose) by analyzing urine of diabetic and hypertensive patients without and with complications compared to control ones and correlating the changes in the spectral features. In this study, PCA scores showed a significant decrease in urea and creatinine concentrations in the LR, HR, and DI groups compared to CTR group and higher glucose concentration in the urine of the LR and HR groups compared to CTR group. QDA showed that the PCs PC1, PC2, and PC4, which bring spectral features of the urine, creatinine, and glucose, could be used to classify the urine spectra in one of the groups, with high discrimination capability (70% overall classification rate, 89% for the CTR, and 81% for both the HR and DI groups). This result would, in turn, make the Raman spectroscopy a technique to predict the renal function status and possible kidney failure due to DM and HT diseases using a single spectrum of urine.

Acknowledgments

L. Silveira Jr. acknowledges FAPESP (São Paulo Research Foundation, Brazil) for the partial financial support (Proc. no. 2009/018788-5 and 2012/20666-0).

References

Bispo et al.: Correlating the amount of urea, creatinine, and glucose in urine from patients...