



## Analysis of energy and major components in chromatographic signals for the diagnosis of prostate cancer

Análisis de energía y de componentes principales en señales cromatográficas para el diagnóstico de cáncer de próstata

Ángelo Joseph Soto-Vergel<sup>1</sup>, Luis Enrique-Mendoza<sup>2</sup>, Byron Medina-Delgado<sup>3\*</sup>

<sup>1</sup>Magíster en Educación Matemática, [angelo.soto@unipamplona.edu.co](mailto:angelo.soto@unipamplona.edu.co), ORCID: 0000-0001-5093-0183, Universidad de Pamplona, Pamplona, Colombia.

<sup>2</sup>Magíster en Ingeniería Biomédica, [luis.mendoza@unipamplona.edu.co](mailto:luis.mendoza@unipamplona.edu.co), ORCID: 0000-0002-2012-9448, Universidad de Pamplona, Pamplona, Colombia.

<sup>3</sup>Doctor en Ciencias, [byronmedina@ufps.edu.co](mailto:byronmedina@ufps.edu.co), ORCID: 0000-0003-0754-8629, Universidad Francisco de Paula Santander, Cúcuta, Colombia.

**How to cite:** A. Soto-Vergel, L. Mendoza y B. Medina-delgado, "Analysis of energy and major components in chromatographic signals for the diagnosis of prostate cancer". *Respuestas*, vol. 24, no. 1, pp. 74-83, 2019.

Received on May 25, 2018; Approved on October 15, 2018

### ABSTRACT

**Keywords:**

Energy analysis  
Principal component analysis  
Prostate cancer  
Chromatography  
Signal processing

The prostate exam is an early detection tool to prevent prostate cancer and the main diagnostic tools for obtaining signs are generally invasive. This article tries chromatographic signals from the urine of prostate cancer patients and control patients as a non-invasive examination proposal. For this purpose, methodologically, urine samples are taken, digitized in chromatograms, treated with mathematical techniques and classified. The mathematical techniques are time normalization, dead time elimination, baseline correction, noise elimination, and peak alignment. Classification techniques analyze energy, in the domain of time and frequency, and the main components in sedimentation graphs and scores. As a result, the chromatographic signal is characterized and identifies the characteristic curve that represents the signal of prostate cancer patients and control patients. The data structure shows a cluster distribution of 88.88% of the vectors for the control patients. In the case of prostate cancer patients, the distribution of data is in clusters around the area defined by control patients. This characterization demarcates signal classification regions to diagnose possible prostate cancer patients, validating the relationship between the chromatographic signal and cancer.

### RESUMEN

**Palabras clave:**

Análisis de energía  
Análisis de componentes principales  
Cáncer de próstata  
Cromatografía  
Procesamiento de señales

El examen de próstata es una herramienta de detección temprana para prevenir el cáncer de próstata y los principales instrumentos diagnósticos para obtener indicios son generalmente invasivos. Este artículo analiza señales cromatográficas provenientes de la orina de pacientes con cáncer de próstata y pacientes control como propuesta de examen no invasivo. Para tal efecto, metodológicamente, se toman muestras de orina, se digitalizan en cromatogramas, se tratan con técnicas matemáticas y se clasifican. Las técnicas matemáticas son normalización de tiempo, eliminación del tiempo muerto, corrección de línea base, eliminación de ruido y alineación de picos. Las técnicas de clasificación analizan la energía, en el dominio del tiempo y frecuencia, y los componentes principales en gráficas de sedimentación y puntuaciones. Como resultado se caracteriza la señal cromatográfica e identifica la curva característica que representa la señal de los pacientes con cáncer de próstata y pacientes control. La estructura de los datos muestra una distribución de conglomerado, del 88,88 % de los vectores, para los pacientes control. Para el caso de los pacientes con cáncer de próstata la distribución de los datos es en conglomerados alrededor de la zona delimitada por los pacientes control. Esta caracterización demarca regiones de clasificación de señales para diagnosticar posibles pacientes con cáncer de próstata, validando la relación existente entre la señal cromatográfica y el cáncer.

\*Corresponding author.

E-mail address: [byronmedina@ufps.edu.co](mailto:byronmedina@ufps.edu.co) (Byron Medina Delgado)

Peer review is the responsibility of the Universidad Francisco de Paula Santander.

This is an article under the license CC BY-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



## Introduction

Prostate cancer is one of the cancers that most affects the male gender today; more than 5% of every million people are affected by this disease; In addition, the early detection tools available to prevent it and the main diagnostic instruments to obtain evidence are generally invasive, with the rectal examination and serum concentration of the specific prostate antigen being the best known. In this sense, [1] identified factors that may be related to the non-performance of the exam such as: fear of cancer, shame, discomfort, pain, low educational level, disinformation of the exam, distrust of medical professionals and concern that the rectal touch may affect masculinity; factors that are expected to be mitigated with this research, taking advantage of the increasing use of new technologies, where applications have been developed to improve health conditions worldwide [2], seeking to make the procedures as effective and as invasive as possible.

Computer-assisted diagnostic systems, which use signal processing techniques, have been widely used to diagnose diseases such as upper limb sarcopenia [3], cardiovascular diseases [4], [5], Parkinson [6] - [8], to mention a few. Likewise, prostate cancer has also been tried to diagnose using image processing techniques from the chemical treatment of a biopsy [9], [10]; others have used machine learning techniques to improve the validity of the diagnosis [11], [12], however, the method remains invasive in obtaining the sample for the analysis of the information contained therein.

However, it is possible to obtain information on prostate cancer non-invasively through chromatography, a procedure defined as the method by which chemical components are separated from a sample, which is represented by a one-dimensional signal with which it is possible to analyze delay, energy or concentration times; allowing the qualitative and quantitative identification of chemical components based on their distribution for characterization [13].

As presented, this article tries urine samples from a chromatographic process to obtain one-dimensional signals, analyzes the differentiating characteristics

by applying signal processing techniques and identifies whether the signal corresponds to a patient with prostate cancer or a control patient (no prostate cancer).

Processing techniques for the characterization of chromatographic signals include time normalization, dead time elimination, baseline correction, noise elimination, signal alignment, energy analysis for feature extraction and principal component analysis for classification.

This document presents the materials and methods, describing the methodology implemented and exposes the results obtained with their respective analyses.

## Materials and methods

Figure 1 shows the research methodology implemented in the sampling stages, database consolidation, signal conditioning using mathematical processing techniques and chromatogram classification.



Figure 1. Research methodology

The sampling includes the search for patients diagnosed with prostate cancer and control patients and the chemical preparation of the sample. The purpose of this stage is to have the appropriate urine sample of each patient, ready to enter the chromatogram, a process that occupies 61% of the total time dedicated to this procedure. Therefore, an adequate sample preparation is a determining factor

in the correct performance of a chromatography, since it guarantees the integrity of the results and eliminates the contaminants that affect the chromatogram [14].

The consolidation of the database contemplates the digitization of the one-dimensional signals in text files, one per patient, whose data correspond to the intensities of the sample in millivolts. Each patient has an associated chromatogram, which is constructed iteratively until the appropriate resolution is obtained [13], [15].

Figure 2 graphically depicts the text file of a chromatogram with its attributes, where the x-axis constitutes the time in minutes and the axis and the intensity of the sample in millivolts. The recording time in each chromatogram is defined in seven minutes.

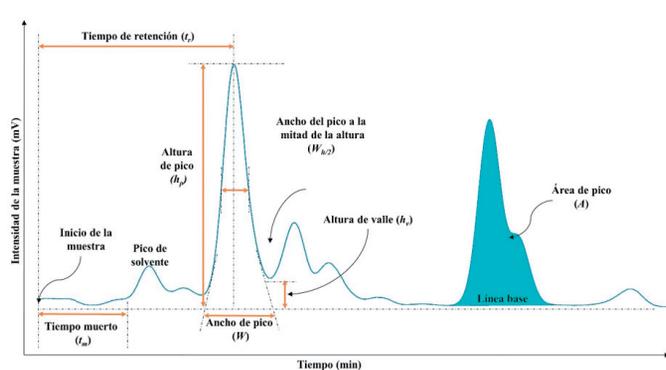


Figure 2. Chromatogram and its attributes

The attributes of the chromatogram are: peak height ( $h_p$ ) which is the portion of a component defined by its height and width at baseline; the dead time ( $t_m$ ) which is the time to register the first component; retention time ( $t_r$ ) considered as the time elapsed between the start of the sample and the appearance of peaks, used to identify and characterize signals; the peak - valley relationship ( $h_p/h_v$ ) used as a classification criterion when it is not possible to separate two consecutive peaks in the baseline, where  $h_v$  is the height at the lowest point of the curve that separates the minor and major peaks above the extrapolated baseline; and finally, the peak width at half the height ( $W_{(h/2)}$ ) and the peak area ( $A$ ), parameters used to describe signal characteristics in a given context [15].

The signal conditioning adapts the data, through

mathematical processing techniques, so that the characteristics of the signal can be correctly classified and validated; To do this, it uses processes such as: normalization of time, elimination of dead time, correction of baseline, elimination of high and low frequency noise, and alignment of peaks [16] - [18].

The chromatogram classification extracts the energy of each peak, in the time and frequency domain, to form characteristic vectors for each signal, which are used to classify the signals with principal component analysis. This last analysis is commonly used to validate if the chosen characteristics of a signal, in this case, the energy, are correct for classification [19] - [22].

## Results and Discussion

This section is structured based on the methodology of Figure 1, presents the results of the processing of chromatographic signals from urine samples and exposes its analysis.

### Sampling

Chromatographic signals were taken at the Quality Control Laboratory of the University of Pamplona, based on the analysis of urine samples. Eighteen shots were taken, nine corresponding to prostate cancer patients and nine control patients.

### Database Consolidation

Eighteen text files were obtained in the .txt format, one file for each patient, which constitute the database of chromatographic signals from urine samples. Figure 3 graphically shows the chromatogram of a patient with prostate cancer and a control patient.

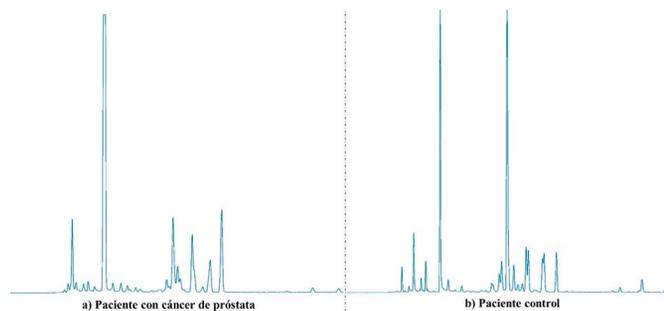


Figure 3. Graphical representation of the chromatogram digitization

The control patient records a prominent peak in the chromatogram, while the patient with prostate cancer differs by presenting two prominent peaks. That is, there is an indication in the chromatographic signal that relates a possible prostate cancer.

**Signal Conditioning**

Figure 4 shows the order of the processes used for signal conditioning. The processes are time normalization, dead time elimination, baseline correction, low and high-frequency noise elimination, and peak alignment. Each process implements threads sequentially represented horizontally. Next, each process is described.

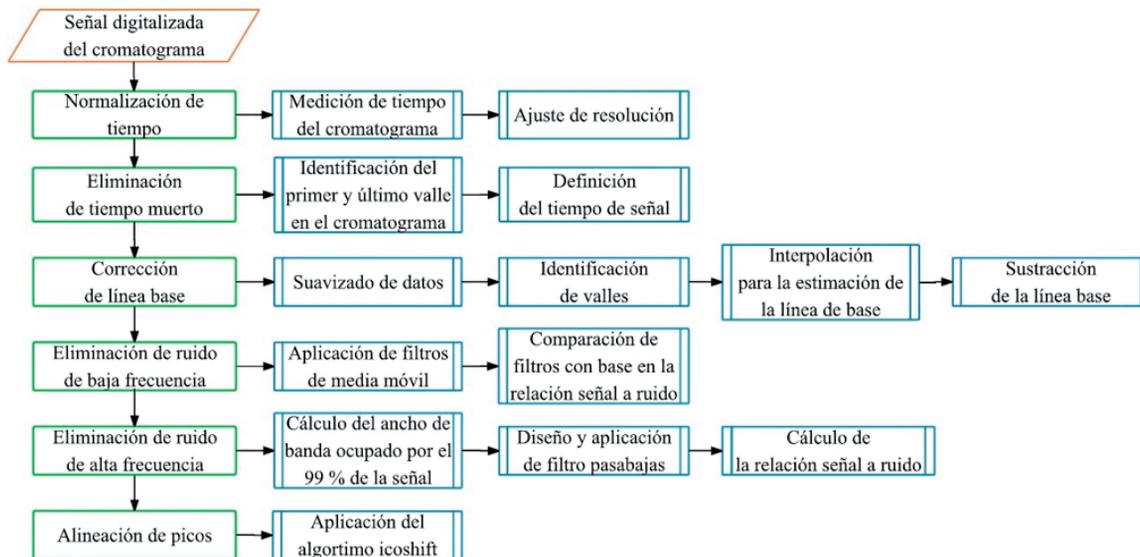


Figure 4. Mathematical techniques for signal conditioning

**Time normalization**

It measures the time it takes to generate the chromatogram to set it as a limit and generates a vector between zero and the measured value, whose resolution depends on the number of samples in each chromatogram, to correlate the intensity of the sample with a unique retention time. For this project, the normalized time was 7.0 minutes per chromatogram.

**Downtime Elimination**

Determine the time that the chromatogram information contains, identifying the instant at which the first valley of the first peak and the second valley of the last peak of the signal appear. The application

of this process to the project chromatograms obtained a minimum signal time of 2.3 minutes and a maximum signal time of 3.0 minutes. For this reason, the maximum signal time of 3.0 minutes is chosen, to avoid the loss of information.

**Baseline Correction**

Corrects errors in the signal offset concerning the axis of zeros for the correct reading of the intensities of the sample. For this, an algorithm is applied that softens the signal using a weighted moving average filter using (1), identifies the valleys using the criterion of the second derivative using (2) to form a vector with these values, interpolates a curve

between two Consecutive valleys for the estimation of the baseline between them, using a second-order spline approximation by (3) and, subtracts the values of the peak baseline to correct the signal by (4).

$$y_i = \sum_{i=1}^n c_i y_{i-1} \quad (1)$$

Where:

$y_i$  is the present value of the signal

$y_{i-1}$  is the past value of the signal

$c_i$  is the weighting factor

$n$  is the maximum number of samples

$$y'' = y_{i+2} - 2y_{i+1} + y_i > 0 \quad (2)$$

Where:

$Y''$  is the second derivative of the sample  $y_i$

$y_i$  is the present value of the signal

$y_{i+1}$  is the future value of the signal

$y_{i+2}$  is the future value of  $y_{i+1}$

$$p(v) = \begin{cases} p_0(v) & \text{si } v \in [v_0, v_1] \\ p_1(v) & \text{si } v \in [v_1, v_2] \\ \dots & \dots \\ p_{i-1}(v) & \text{si } v \in [v_{i-1}, v_i] \end{cases} \quad (3)$$

Where:

$p(v)$  is the set of quadratic polynomials of the baseline of each peak

$p_i(v)$  is the quadratic polynomial of the single peak baseline

$v_i$  is each valley identified to the right of the peak

$v_{i-1}$  is each valley identified to the left of the peak

$$y_i = y_i - p_i(v) \quad (4)$$

Where:

$y_i$  is the present value of the signal

$p_i(v)$  is the quadratic polynomial of the peak baseline

Figure 5 shows the graphic representations of the original signal (blue color) and the signal with the corrected baseline (orange color). The baseline correction process does not alter the chromatogram peaks, evidencing the proper functioning of the algorithm.

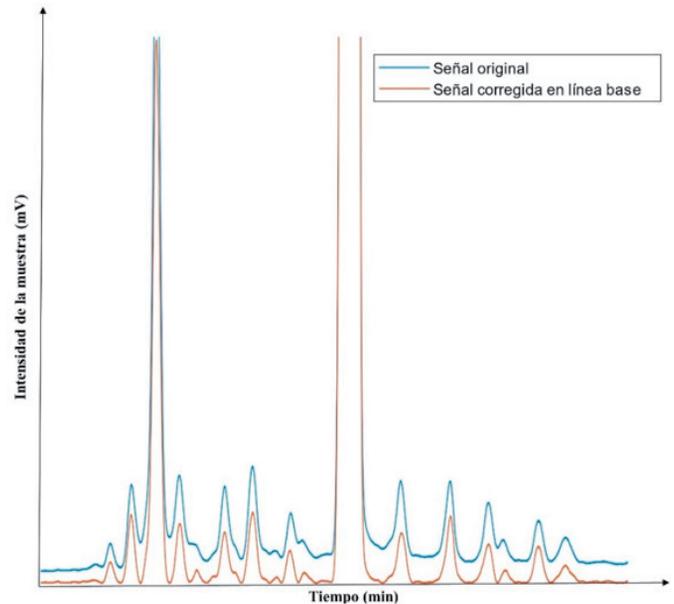


Figure 5. Graphical representation of baseline correction

### Low frequency noise elimination

The data smoothing technique is used that helps reveal characteristics and components of the signal that can be hidden by noise, making it difficult to calculate parameters such as areas and heights [23]. In this process the algorithms are compared considering the signal to noise ratio.

The first algorithm implemented is rectangular smooth or boxcar which is an algorithm without weighted smoothing replaces each point in the signal with the average of  $m$  adjacent points, where  $m$  is a positive and odd integer so that the coefficients balance  $x$  of the peaks and other characteristics in the smoothed signal. This project defines  $m=3$  resulting (5).

$$S_j = \frac{y_{j-1} + y_j + y_{j+1}}{3} \quad (5)$$

Where:

$S_j$  is the signal softened

$y_j$  is the present value of the original signal

$y_{j-1}$  is the last value of the original signal

$y_{j+1}$  is the future value of the original signal

Another algorithm used is the triangular smooth that implements a weighted smoothing function. For this project it is defined  $m=5$  resulting (6).

$$S_j = \frac{y_{j-2} + 2y_{j-1} + 3y_j + 2y_{j+1} + y_{j+2}}{9} \quad (6)$$

In both cases, the integer in the denominator is the sum of the coefficients in the numerator, which results in a smooth unit gain that has no effect on the signal where it is a straight line and that preserves the area of the peak.

The *pseud Gaussian*  $\sigma$  and the *w width* are also used, which iterate three and four times the rectangular smooth of three points, respectively.

Table I shows the signal to noise ratio for the smoothing algorithms applied, where it is possible to conclude that the best results were obtained by applying four passes of a *rectangular smooth* three-point.

Table I. Signal to noise relationship by applying softening algorithms

Signal type	Improvement of the signal to noise ratio			
	Rectangular	Triangular	Pseudo - Gaussian	w -width
Control patients	32,33 %	32,47 %	32,48 %	32,53 %
Patients with prostate cancer	51,22 %	51,27 %	51,32 %	51,46 %

Figure 6 shows the result of applying the *w width* algorithm, which eliminates the signal jumps, contributing to the identification of peaks and valleys for the calculation of the area of the peaks, in the energy analysis.

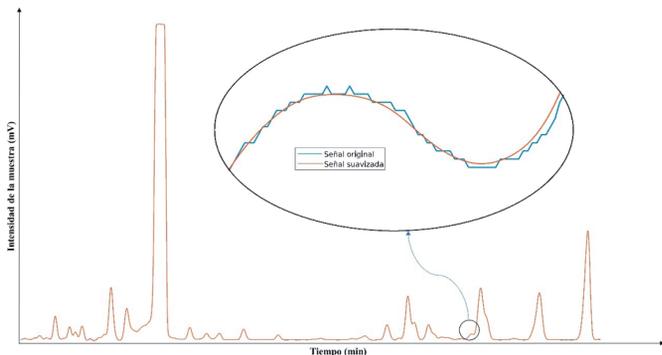


Figure 6. Graphical representation of applying the *w width* algorithm

### High frequency noise elimination

Analyze the area of interest in the frequency domain. Design a high-pass low pass filter, average the filter

values, complete the vector with the calculated average data, and shift the data to compensate for the filter delay. With the application of the filter, the signal to noise ratio of the smoothed signal is improved by 5.14% for the chromatograms of the control patients and by 6.09% for the patients with prostate cancer.

### Peak alignment

In the digitized signal of the chromatogram, misalignment and shifting of the data are common due to calibration errors of the instruments or of the environment where the procedure is performed. For this reason, the *icoshift* algorithm developed by [24] is implemented, consisting of three parts: definition of intervals, maximization of the cross-correlation of each interval by a Fast Fourier Transform (FFT) engine and reconstruction of the signal. This algorithm takes a reference signal and aligns other signals according to the peaks marked as representative of it, performing the process iteratively until the alignment of all the peaks is achieved. This procedure provides value in the construction of the characteristics matrix for the analysis of main components.

Figure 7 graphically depicts the effect on chromatographic signals when applying an iteration with the *icoshift* algorithm. The figure on the left shows the non-aligned chromatograms and the figure on the right shows the aligned chromatograms; this is done in order to have the representative peaks of each component at the same retention time (location).

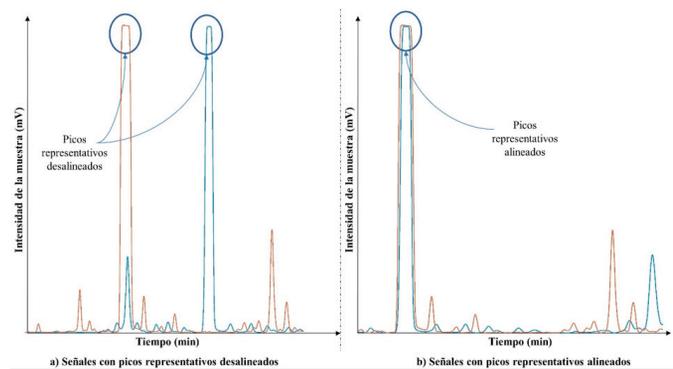


Figure 7. Graphical representation of applying the *icoshift* algorithm

### Chromatogram Classification

Figure 8 shows the order of the processes used for the classification of chromatograms. The processes are energy analysis and principal component analysis.

Each process implements threads sequentially represented horizontally. Next, each process is described.

the time domain, in the frequency domain applying FFT and, applying the Discrete Cosine Transform (DCT).



Figure 8. Mathematical chromatogram classification techniques

### Energy analysis

This process improves the resolution and symmetry of the peaks, identifies the peaks and valleys, delimits the area of each peak, calculates its energy and builds the characteristic matrix.

The sharpening technique improves the resolution by (7) and the symmetry by (8) and; it contributes to the precision of the measured areas to identify the peaks and valleys and, to delimit the area of each peak with the perpendicular drop method.

$$R_j = y_j - k_1 y'' + k_2 y'''' \quad (7)$$

Where:

$R_j$  is the signal with the improved resolution

$y_j$  is the original signal

$y''$  is the second derivative of the original signal

$y''''$  is the fourth derivative of the original signal

$k_1, k_2$  are weighting factors of the second and fourth derivative respectively

$$S_j = y_j + k_1 y' \quad (8)$$

Where:

$S_j$  is the signal with the improved symmetry

$y_j$  is the original signal

$y'$  is the first derivative of the original signal

$k_1$  It is a weighting factor

The energy of each peak is calculated using (9), where the summation limits are the data between the beginning and end valleys of each peak. This calculation is performed with each chromatogram in

$$E_p = \sum_{i=v_i}^{v_{i+1}} y_i^2 \quad (9)$$

Where:

$E_p$  is the energy of each peak

$p=1,2,3...n$  is the energy of each peak

$y_i^2$  is the intensity of the sample squared

$v_i$  is the valley to the left of each peak

$v_{i+1}$  is the valley to the right of each peak

As a result, three characteristic matrices are obtained, one for each analysis, formed by the energy of each peak respecting its position. In the case of absence of component, the position in the feature vector is completed with zero.

### Principal component analysis

It uses a correlation algorithm between variables that uses statistical methods to describe a set of data in terms of new uncorrelated variables, called components, by the linear combination of the original variables seeking to reduce the dimensionality of the data [25].

Figure 9 shows the analysis of main components in a sedimentation plot from the energy data calculated in the time domain, with a unique characteristic curve, whose classification does not show overlap between patients. The main components 1 and 2 are those that contain the most information of the characteristic chromatograms, followed by component 3.

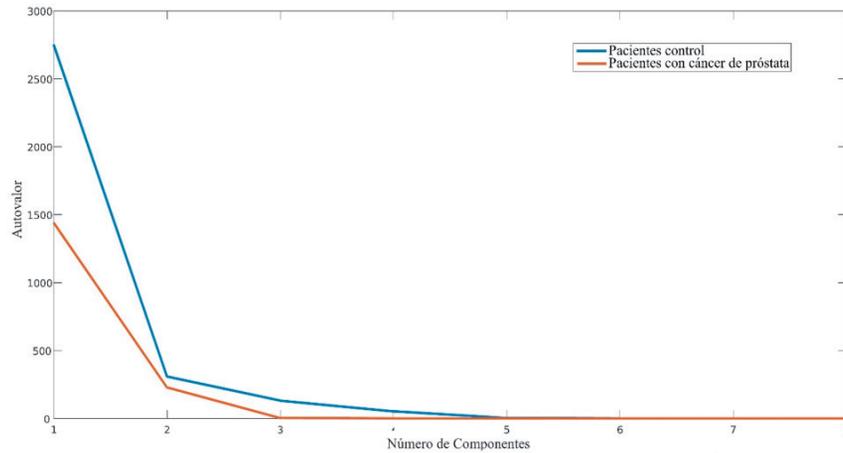


Figure 9. Sedimentation chart of main components - Energy analysis in the time domain

Figure 10 shows the analysis of main components in a sedimentation plot from the energy data calculated in the frequency domain, FFT and DCT, which did not show a single characteristic curve, showing overlap between patients in their classification.

score graph, with the main components 1 and 2 as axes, with a conglomeration of 88.88% for the control patients. Patients with prostate cancer are presented as isolated values. This analysis shows that control and cancer patients have differences in their chromatograms, evidencing the relationship between signals and prostate cancer.

Figure 11 shows another alternative for analyzing the data calculated in the time domain using the

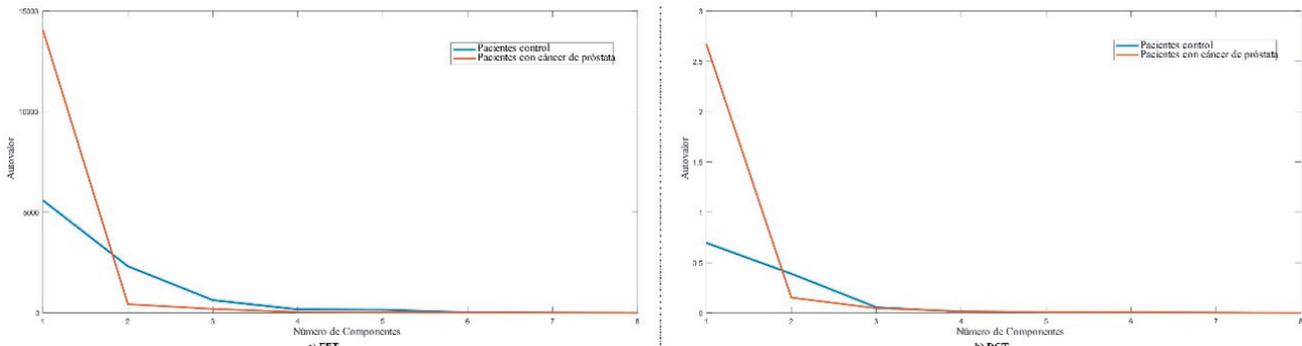


Figure 10. Sedimentation graph of main components - Energy analysis in the frequency domain

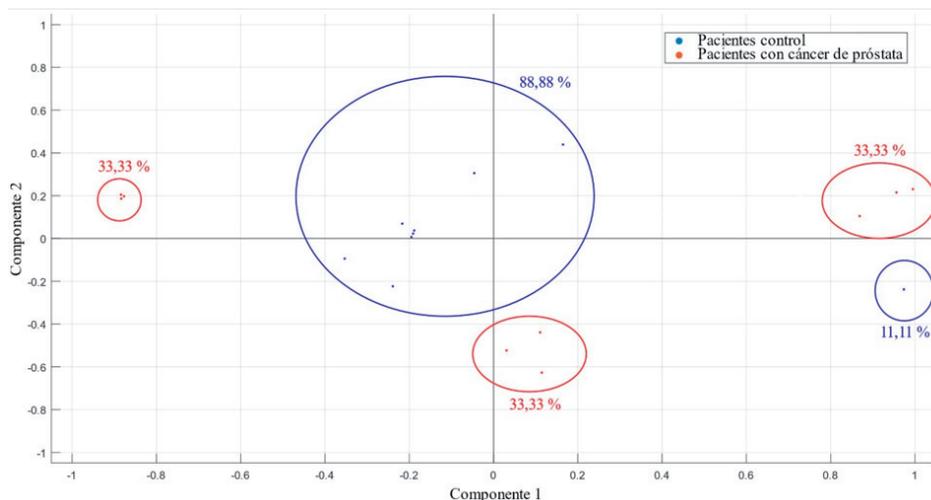


Figure 11. Score chart of main components

## Conclusions

The sequence of the mathematical techniques of signal processing applied to the chromatograms improved the signal-to-noise ratio is 37.67% for control patients, and in 57.55% for patients with prostate cancer. This improvement contributes to the accuracy in the identification of peaks and valleys, the analysis of the energy and main components.

The sedimentation graph has a unique behavior of the main components corresponding to control patients and prostate cancer patients, validating the energy analysis of the peaks of each signal in the time domain as a differentiating factor.

In the score graph, the structure of the data shows a cluster distribution of 88.88% of the vectors for the control patients. The data representing 11.11% is considered atypical and involves an error in the inclusion of the chromatogram in the control group, which could be presented in the urine sample.

In the case of prostate cancer patients, the distribution of the data is uniform in three groups of 33.33% of the vectors around the area defined by the control patient vectors. This representation delimits signal classification regions to diagnose possible prostate cancer patients.

The results show evidence to apply the extraction of significant peaks, as a pattern extraction technique and to find other characteristics that differentiate and accentuate the classification of chromatograms of prostate cancer patients and control patients.

## References

- [1] Á. Fajardo-Zapata and G. Jaimes-Monroy, "Conocimiento, percepción y disposición sobre el examen de próstata en hombres mayores de 40 años", *Revista de la Facultad de Medicina*, vol. 64, no. 2, pp. 223–228, 2016.
- [2] D. Glujovsky, A. Bardach, S. García-Martí, D. Comandé and A. Ciapponi, "PRM2 EROS: A New Software For Early Stage Of Systematic REVIEWS", *Value in Health*, vol. 14, no. 7, pp. A564, 2011.
- [3] I. Rivera et al., "Diseño de dispositivos para el diagnóstico de sarcopenia en miembro superior", *Memorias del Congreso Nacional de Ingeniería Biomédica*, vol. 2, no. 1, pp. 174–177, 2017.
- [4] L. Garrido-Martínez and R. I. González-Fernández, "Revista cubana de informática médica", *Revista Cubana de Informática Médica*, vol. 15, no. 2, pp. 153–164, 2015.
- [5] E. Dugarte-Dugarte et al., "Algoritmo de bajo costo de procesamiento para la detección de potenciales tardíos ventriculares (PTV)", *Revista Conocimiento Libre y Licenciamiento*, vol. 8, no. 15, pp. 73–93, 2017.
- [6] P. A. Stack-Sánchez, G. Dorantes-Méndez and A. R. M. Rodríguez, "Caracterización del temblor Parkinsoniano mediante dimensión fractal en señales de acelerometría", *Memorias del Congreso Nacional de Ingeniería Biomédica*, vol. 5, no. 1, pp. 190–193, 2018.
- [7] M. E. Bedoya-Vargas, J. C. Vásquez-Correa and J. R. Orozco-Arroyave, "Time-frequency representations from inertial sensors to characterize the gait in Parkinson's disease", *TecnoLógicas*, vol. 21, no. 43, pp. 53–69, 2018.
- [8] I. G. Bravo, P. A. S. Sánchez, G. D. Méndez and A. R. M. Rodríguez, "Evaluación del movimiento a través de acelerometría en pacientes con enfermedad de parkinson", *Memorias del Congreso Nacional de Ingeniería Biomédica*, vol. 4, no. 1, pp. 138–141, 2017.
- [9] E. Payá-Bosch, "Desarrollo de un sistema de extracción avanzada de características en imagen histológica para la identificación automática del cáncer de próstata", *Universidad Politécnica de Valencia*, 2017.
- [10] B. Zapote-Hernández, J. Cruz-Santiago, E. González-Vargas and A. Jaramillo-Núñez, "Concordancia diagnóstica entre los métodos visual e informático en la detección de metástasis por gammagrafía ósea en cáncer de

- próstata”, *Anales de Radiología México*, vol. 15, no. 2, pp. 111–119, 2016.
- [11] L. Hussain et al., “Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies”, *Cancer Biomarkers*, vol. 21, no. 2, pp. 393–413, 2018.
- [12] J. Wang, C.J. Wu, M.L. Bao, J. Zhang, X.N. Wang and Y.D. Zhang, “Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer”, *European Radiology*, vol. 27, no. 10, pp. 4082–4090, 2017.
- [13] B. Patiño-Domínguez, “Determinación de parámetros operacionales necesarios en el empaquetado de columnas de cromatografía”, *Universidad Da Coruña*, 2016.
- [14] R. Majors, "Sample preparation fundamentals for chromatography", *Canada: Agilent Technologies*, 2013.
- [15] J. Cazes, "Encyclopedia of Chromatography", *3ra ed. New York*, 2009.
- [16] A. Medina-Santiado, “Sistema de diagnóstico de señales biomédicas con redes neuronales artificiales”, 2015.
- [17] J. A. Navarro-Acosta and J. P. Nieto-González, “Detección y diagnóstico de fallas para la dinámica lateral de un automóvil utilizando máquinas de soporte vectorial multiclase”, *Research in Computing Science*, vol. 73, pp. 167–179, 2014.
- [18] M. A. Melara-Estrada, “Introducción a la transformada Wavelet y la la teoría de análisis de señales”, *Universidad de El Salvador*, 2015.
- [19] A. Sheinker and M. B. Moldwin, “Magnetic anomaly detection (MAD) of ferromagnetic pipelines using principal component analysis (PCA)”, *Measurement Science and Technology*, vol. 27, no. 4, pp. 045104, 2016.
- [20] Z. Chen, Q. Zhu, Y. C. Soh and L. Zhang, “Robust Human Activity Recognition Using Smartphone Sensor via CT-PCA and Online SVM”, *IEEE Transactions on Industrial Informatics*, 2017.
- [21] J. G. Rueda-Bayona, C. J. Elles-Pérez, E. H. Sánchez-Cotte, Á. L. González-Ariza and G. D. Rivillas-Ospina, “Identificación de patrones de variabilidad climática a partir de análisis de componentes principales, Fourier y clúster k-medias”, *Tecnura*, vol. 20, no. 50, pp. 55–68, 2016.
- [22] P. Arroyo, I. Suárez, J. Lozano, J. Herrero and P. Carmona, “Nariz electrónica personal para la detección de contaminantes en el aire”, *Actas las XXXIX Jornadas Automática*, pp. 894–899, 2018.
- [23] T. O’Haver, “A Pragmatic Introduction to Signal Processing with applications in scientific measurement”, *University of Maryland at College Park*, 2018.
- [24] F. Savorani, G. Tomasi and S. B. Engelsen, “Alignment of 1D NMR Data using the iCoshift Tool: A Tutorial”, in *Magnetic Resonance in Food Science: Food for Thought*, pp. 14–24, 2013.
- [25] A. Kassambara, "Practical guide to principal component methods in R : PCA, (M)CA, FAMD, MFA, HCPC", *factoextra. STHDA*, 2017.