SPECTROGRAMS COMPARISON BETWEEN NORMAL AND ABNORMAL SIGNAL OF IN VIVO MAMMARY TISSUE, OBTAINED BY BACK-SCATTERING INFRARED LIGHT

Janeth Fernández P, Mónica Andrea Mendoza Arias, Gomez-Torrado A, David A. Miranda

Universidad Industrial de Santander, Cra 27 Cl 9, Bucaramanga, Colombia.

Email: dalemir@uis.edu.co

Abstract. High breast cancer incidence and mortality rates in Colombian women motivated the research group CIMBIOS to develop methodologies for early detection of cancer. To this aim, a method to study alterations in breast tissue has been proposed by CIMBIOS. Colombian invention patent application No NC2017 / 0003413 describes its method. A technological design called FEDSA emerged from this process. With FEDSA, the breast was irradiated in vivo (near infrared), and the back-scattering light by tissue was measured. Twenty-four women with seventeen normal and seven abnormal diagnoses were studied with this device, and intensity as a function of time was obtained for them. Data pre-processing to eliminate white noise was implemented, and fast Fourier transforms with two sampling windows, 512 and 1024, was applied. With the squared magnitude of the FFT, 1104 spectrograms were constructed, or 48 spectrograms per patient. Abnormal and normal spectrogram were labeled according to their clinical history, and these data were ordered by sampling window and frequency. Differences between normal and abnormal tissue in the breast quadrants were found with the statistical method ANOVA. In the breasts, outer upper and lower quadrant, and inner upper right in abnormal and normal women were different with p<0.05. These results agree with the clinical history of patients with abnormal classification.

1. Introduction

Breast cancer incidence rates are the highest of any disease affecting Colombian women [1]. Breast self-exam and mammography are the most widely implemented techniques to identify masses or visible abnormalities. Adult women of all ages are encouraged to perform breast self-exams at least once a month because forty percent of breast cancer diagnoses result from women feeling a lump (National Breast Cancer Foundation UUEE) [2].

Mammography uses ionizing radiation (X-rays), whose physical nature limits its implementation in young women. For women aged over 35, once or twice a year is suggested, depending on each patient. For this reason, optical techniques have been worked with light in the infrared. Radiation light in performing breast cancer detection is booming, because: 1) it does not alter or damage tissues, and 2) it can help to measure
physical properties such as absorption or dispersion coefficient in tissue. These coefficients have typical values for both normal and abnormal tissue [3-5].

Many experimental techniques have been built in order to understand and detect the carcinogenesis process. PWS (Partial Wave Spectroscopy) is an optical technique that can perform pre-screening of lung, ovary and colon cancer [6]. This technique measures a parameter of disorder that describes the degree of structural heterogeneity. Differences between healthy and transformed tissue can be measured with this technique. PWS is the first technology to measure the field cancerization effect or field effect (FCE) [7]. This effect was formulated by Slaughter in 1953 to explain the origin of multiple oral cavity tumors [8]. Currently, all abnormalities in normal tissue caused by abnormal cells have been related by FCE [9].

FED-SA (Field Effect Detection by Spectral Analysis) was designed in order to measure abnormalities in biological tissue based on the field cancerization effect detection. This technique was proposed based on amplification of biochemical anomalies due transformed cells [10]. This phenomenon, associated with FCE, implies an increase in the molecular expression due to transformed cell. This increase changes the molecular sizes distribution in the tissue and with a measurement dependent of particles size, dynamic light backscattering, could be possible infer the tissue cancer state. Colombian invention patent application NC2017 / 0003413 described its method [11]. In this investigation, a spectral analysis with spectrograms was proposed as a guide to obtain the best method of analysis of the technique. Then, spectrograms of the signals obtained with the FED-SA technique were compared in order to find differences between the tissue previously classified as normal and abnormal.

2. Methods and materials

2.1 FED-SA (Field Effect Detection by Spectral Analysis) technique

This method, proposed based on Colombian invention application patent No NC2017 / 0003413 for detecting abnormalities of biological cells, comprises the steps: a) Illuminating tissue with infrared light. b) Detecting a retro-scattered light signal by the tissue with two detectors. c) Transforming the light signal of step b) into a digital signal with a data acquisition unit. d) Processing the digital signal of step c) and plotting a spectrum representing frequency, time, space, intensity, and combinations thereof. e) Comparing the spectrum of step d) with a predetermined spectrum. The design, building and measure protocol of FED-SA was proposed by CIMBIOS group in the investigation “Study of field cancerization in breast tissue”.

2.2 Participants and ethical approval

FED-SA measured data from obtained from twenty-four women aged 23 to 60 years. Mean age was $41 \pm 12.36$ years. Verbal and written consent was obtained from all study participants. The study protocol by Universidad Industrial of Santander Ethics Committee was approved.

2.3 Measure protocol

The ethical committee determined FED-SA to be non-invasive, posing minimal risk. The biological safety of women was guaranty by the use of dressing rooms with the required dressing gowns, latex gloves, and masks. Between measurements, correct functioning of the sensor (detection system) was verified. In order to comply with the principle of data confidentiality, the people involved in the research were assigned a PNXXX code (Patient No. XXX), and the coding of quadrants was used, adding R (right breast) and L (left breast).
A unit into the university was suitable for measurements. Participants visited the unit between 8:00 am and 4:00 pm. All assessments were conducted with the participant in sitting position after a 20-minute rest period. The illumination and detection systems were placed by gently pressing the tissue. Each breast was divided into four quadrants. For right and left breast OURQ (outer upper right quadrant), IURQ (inner upper right quadrant), OLRQ (outer lower right quadrant), ILRQ (inner lower right quadrant), OULQ (outer upper left quadrant), IULQ (inner upper left quadrant), OLLL (outer lower left quadrant), ILLQ (inner lower left quadrant). Measurements were made on the right breast following the order of the quadrants, and then on the left breast.

2.4 Experimental measurements
The illumination and detection system has one transmitter (NIR) and two detectors. The acquisition system has three channels for data acquisition, every channel acquired at 333 KHz. The third channel subtracts (differential) the signal of detectors. For each quadrant, measurements lasted 10 seconds. For this time, each channel detection contains 2.7 million data, for a total of 16 million data obtained per patient.

2.5 Pre-processing and calculation of spectrograms
Voltage as a function of time from the detection system for twenty-four female volunteers was obtained. The backscattering light that emerged from the tissue is the value of voltage (directly proportional to intensity) which is shown in figure 1a. White noise was observed within the entire biological signal. A method to minimize this noise from all the signals was proposed: let \( x = x[t] \) be a set of discrete data, where \( t \) is a vector that represents the time sequence, sees figure 1b. Then:

a) Dataset \( x \) corresponds to the digital sample of a signal that contains an information component, \( x_a \), and a noise component, \( e \), that is \( x = x_a + e \)

b) The information component has an average equal to zero, that is, \( <x_a> = 0 \)

c) The noise of the signal is white noise, which means that it does not protrude for any special frequency, but is distributed across the spectrum.

A new function \( y \) with the original data was constructed in order to minimize noise. An easy way to minimize white noise is to average the data; however, to minimize noise, it is necessary to make several measurements (several datasets). Therefore, in this investigation, a theoretical model was used in data preprocessing. Let \( x = x[t] \) and \( y = y[t] \), two discrete random signals such that they contain an information component \( (x_a, y_a) \) and a white noise \( (e_x, e_y) \). Then, \( x = x_a + e_x, y = y_a + e_y \). Hence signal \( y \) from signal \( x \) is obtained when \( e_y << e_x \). Signal "\( y \)" can be constructed as the result of the set average of sub-sampled signals obtained from "\( x \)", figure 1c.

Signal \( x[t] \) was sampled at the frequency \( f_s \). Then the sampling frequency associated with \( y[t] \) will be \( f_y = f_s \). In order to calculate the spectrograms, Fast Fourier Transform (FFT) in signal \( y[t] \) was implemented.

2.6 Data Analysis
For all participants, 1104 spectrograms were calculated. For each spectrogram, four clusters or ranges of frequency “(0-10 KHz), (10-100 KHz), (100-150 KHz) and (150 -160 KHz)” were created. In every range of frequency, the average value was calculated and a new group of data with this value and label (normal or abnormal) was created. For this new set, a one-way ANOVA statistical analysis was performed. Python programming language for the elimination of noise, calculation of spectrograms, creation of frequency bands, and the application of ANOVA analysis was used.
Figure 1. Back-scattering intensity as a function of time (a), and its amplification in 0.001 s (b). The entire signal into n equal parts (for example 100 data) was divided and $x_1, x_2, x_3, ..., x_{100}$ in each window were identified and the new vector $y$ averaged was found (c).

3. Results and discussions

Figure 2 shows the normal and abnormal spectrograms of two patients on zone OURQ (outer upper right quadrant). The frequency as a function of time for the two detectors and the differential signal is represented. The spectrograms are shown little differences by visual comparison, but it was not possible to identify a pattern between abnormal and normal spectrograms by only visual inspection. The classification between normal and abnormal data was achieved by organized as a cluster of frequencies each spectrogram. For these new data set, ANOVA statistical analysis was done. Data statistical distribution was normal. Principal results for $p$ and $F$ are presented in table 1.
Figure 2. Spectrograms for abnormal and normal patients in the OURQ (outer upper right quadrant). Frequency as a function of time for the two detectors and differential signal is represented.

Table 1. ANOVA statistical analysis to the frequency cluster of the spectrograms conformed by the data normal and abnormal patients.

| Quadrant | Frequency [kHz] | OLR | OLL | OUR | IUR |
|----------|-----------------|-----|-----|-----|-----|
|          | 0-10            | 0-10| 10-100 | 10-100 |
| p        | 0.0477          | 0.0212 | 0.0428 | 0.0334 |
| F        | 4.4             | 6.16 | 4.62 | 5.15 |
For the statistical analysis, it was hypothesized that the frequency data clusters for normal and abnormal patients are not different. The results with p<0.05 reject the hypothesis. This statistical analysis suggests that the right breast of abnormal patients differs from the normal ones in three quadrants primarily and in the left breast, it only finds significant differences in one quadrant.

The results show that data coming from frequencies higher than 100 kHz do not have significant differences. The penetration length in NIR is approximately 0.5 centimeters, and the differences found in frequency were assigned to the back-scattered light to living organisms within the tissue larger than 100 micrometers [12]. Fat cells have a diameter between 15-250 micrometers.

However, abnormal cells can modify cell signaling into the tissue in order to generate oxygen and block antibodies for its growth, with the help of normal cells [13-14]. Therefore, the field cancerization effect is not a localized effect, and it is possible that normal cells amplify the changes in the tissue or in the organ [10]. So, fat cells can contain information about anomalies into the breast.

4. Conclusion
The results of this study showed that the clustering of the spectrogram of dynamic light backscattering let differentiate between normal and abnormal breast tissue. The data was obtained by FED-SA technique in a group of 24 women, of whom 17 belong to the control group, and seven were diagnosed by mammography with possible alterations. Statistical analysis to clusters obtained from spectrograms of data showed significant differences between women with alterations and the control group. The values obtained from p and F showed three quadrants in the right mammary tissue with possible alterations, and a quadrant in the left breast, in accordance with the clinical information: the clinical history of the patients shows a patient with microcysts grouped in the right upper external quadrant; the second patient with suspicion in upper and lower external quadrants of left breast; third patient with microcysts in upper quadrants of both breasts; fourth patient with alterations in upper internal and external quadrants of breasts.

5. Acknowledgment
We are grateful to COLCIENCIAS for the financial support and to the research vice-chancellor of the “Universidad Industrial de Santander.”

References
[1] Torre L A, Bray F, Siegel R L, et al 2015 Ca. Cancer. J. 65 87-108
[2] National Breast Cancer Foundation, INC. about Breast Cancer [Internet] 2016 [revised 2018 October 28]. Available from: https://www.nationalbreastcancer.org/breast-cancer-faqs
[3] Peters V G, Wyman D R, Patterson M S and Frank G L 1990 Phys. Med. Biol. 35 1317-1990.
[4] Suzuki K, Yamashita Y, Ohta K, and Chance B 1994 Inves. Radiol. 29 410–14.
[5] Soliman H, Gunasekara A, Rycroft M, Zubovits J, Dent R, Spayne J, Yaffe M j and Czarnota G J 2010 Clin. Cancer. Res. 1078–0432
[6] Hemant K R, Hensing T and Backman V 2011 Future. Oncol. 7 1–3
[7] Damania D, Hemant K R, Subramanian H, Weinberg D S, et al 2012 *Cancer. Res.* 3807

[8] Slaughter D P, Southwick H W and Smejkal W 1953 *Cancer.* 6 963–68

[9] Dakubo G D 2015 *Cel. Dev. Biol.* 4 2

[10] Fernández J, Méndez-Sánchez S C, Gonzalez-Correa C A and Miranda D A 2016 *Med. Hypotheses.* 97 107–11

[11] Miranda D A and Fernández J *Método y dispositivo para detectar anormalidades en células biológicas* application patent Colombia No NC2017 / 0003413 april 7 2017

[12] Barolet D 2008 *Semin. Cutan. Med. Surg.* 27 227-38

[13] Liao K L, Bai X F and Friedman A 2014 *Plos. One.* 9 e91844

[14] Liao K L, Bai X F and Friedman A 2014 *Plos. One.* 9 e110126