



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
PÓS-GRADUAÇÃO EM GENÉTICA E BIOQUÍMICA**

**ATR-FTIR spectroscopy analysis of saliva components as a
diagnostic and prognostic tool for Breast Cancer:
a preliminary study**

Aluna: Izabella Cristina Costa Ferreira

Orientadora: Profa. Dra. Yara Cristina de Paiva Maia

UBERLÂNDIA - MG

2017



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
PÓS-GRADUAÇÃO EM GENÉTICA E BIOQUÍMICA**

**ATR-FTIR spectroscopy analysis of saliva components as a
diagnostic and prognostic tool for Breast Cancer:
a preliminary study**

Aluna: Izabella Cristina Costa Ferreira

Orientadora: Profa. Dra. Yara Cristina de Paiva Maia

**Dissertação apresentada à
Universidade Federal de Uberlândia
como parte dos requisitos para
obtenção do Título de Mestre em
Genética e Bioquímica (Área
Genética)**

UBERLÂNDIA - MG

2017

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

F383a Ferreira, Izabella Cristina Costa, 1992
2017 ATR-FTIR spectroscopy analysis of saliva components as a
diagnostic and prognostic tool for breast cancer: a preliminary study /
Izabella Cristina Costa Ferreira. - 2017.
77 f. : il.

Orientadora: Yara Cristina de Paiva Maia.
Dissertação (mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Genética e Bioquímica.
Disponível em: <http://dx.doi.org/10.14393/ufu.di.2018.140>
Inclui bibliografia.

1. Bioquímica - Teses. 2. Mamas - Câncer - Teses. 3. Saliva - Teses.
4. Biomarcadores tumorais - Teses. I. Maia, Yara Cristina de Paiva. II.
Universidade Federal de Uberlândia. Programa de Pós-Graduação em
Genética e Bioquímica. III. Título.

CDU: 577.1

Angela Aparecida Vicentini Tzi Tziboy – CRB-6/947



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
PÓS-GRADUAÇÃO EM GENÉTICA E BIOQUÍMICA

**ATR-FTIR spectroscopy analysis of saliva components as a
diagnostic and prognostic tool for Breast Cancer:
a preliminary study**

ALUNO: Izabella Cristina Costa Ferreira

COMISSÃO EXAMINADORA

Presidente: Profa. Dra. Yara Cristina de Paiva Maia (Orientadora)

Examinadores: Dra. Angela Aparecida Servino de Sena Priuli (Membro Externo)
Prof. Dr. Robinson Sabino da Silva (Membro Interno)
Profa. Dra. Juliana Franco Almeida (Suplente Membro Externo)
Profa. Dra. Vivian Alonso Goulart (Suplente Membro Interno)

Data da Defesa: 31 / 07 / 2017

As sugestões da Comissão Examinadora e as Normas PGGB para o formato da Dissertação/Tese foram contempladas

Profa. Dra. Yara Cristina de Paiva Maia

À minha família que sempre me apoiou e incentivou na realização dos meus sonhos e crescimento pessoal e profissional.

AGRADECIMENTOS

Agradeço a Deus por me fortalecer e iluminar a cada dia da minha caminhada, e à minha família por me ajudar na realização dos meus sonhos.

Agradeço ao Prof. Dr Luiz Ricardo Goulart Filho, pelos incentivos, contribuições e por disponibilizar o Laboratório de Nanobiotecnologia para a realização deste estudo.

Agradeço à Profa. Dra. Yara Cristina de Paiva Maia pela brilhante e dedicada orientação, pelo apoio, motivação, paciência e ensinamentos.

Agradeço ao Centro de Pesquisa de Biomecânica, Biomateriais e Biologia Celular – CPBio por disponibilizar o aparelho de espectroscopia FTIR para realização do estudo e à Mestranda Emília pela colaboração.

Agradeço também aos demais professores, alunos de pós-graduação e de graduação do Laboratório de Nanobiotecnologia-INGEB, que compartilharam conhecimentos e experiências de grande importância.

Agradeço aos funcionários e pacientes do Hospital de Clínicas de Uberlândia que colaboraram com a pesquisa, em especial aos funcionários do setor de Ginecologia e Obstetrícia e ao Dr. Donizette William.

Agradeço também aos amigos que me apoiaram e ajudaram de diversas formas na realização dessa etapa.

*“Deus nunca disse que a jornada seria fácil,
mas Ele disse que a chegada valeria a pena.”*

(MAX LUCADO)

SUMÁRIO

ABSTRACT	10
LIST OF ABBREVIATIONS AND ACRONYMS	11
1 INTRODUCTION	12
1.1 General aspects of Breast Cancer (BC).....	12
1.2 Fourier transform infrared (FTIR) spectroscopy: technical principles, advantages and applications	19
1.3 Saliva: a promising biological fluid	30
2 OBJECTIVES	33
2.1 General Objectives	33
2.2 Specific Objectives	33
3 MATERIAL AND METHODS	34
3.1 Ethical aspects and study subjects.....	34
3.2 Sample collection and preparation	34
3.3 ATR-FTIR spectroscopy.....	35
3.4 Spectral data preprocessing	35
3.5 Statistical analysis	36
4 RESULTS	37
4.1 Study subjects characterization	37
4.2 FTIR analysis of saliva spectra between breast cancer, benign and control patients	39
4. 3 FTIR analysis of saliva spectra within the group of breast cancer patients.....	48
5 DISCUSSION	53
6 CONCLUSION.....	57
REFERENCES	58
ATTACHMENTS	68

RESUMO

A crescente incidência mundial de câncer de mama e a ausência de métodos confiáveis e suficientes para detecção precoce exigem a busca de técnicas mais efetivas. Uma potencial técnica candidata é a reflexão total atenuada - infravermelho de transformação de Fourier (ATR-FTIR), que consiste em uma espectroscopia vibratória que pode efetivamente fornecer informações sobre a estrutura e composição química de materiais biológicos a nível molecular. A maioria dos trabalhos sobre a aplicação de FTIR para detecção de câncer de mama utiliza tecido e sangue. No entanto, são necessários mais métodos não invasivos, por exemplo, usando saliva. Este estudo tem como objetivo investigar as diferenças nos espectros entre os grupos analisados, bem como a influência específica das características clínicas relevantes dos pacientes com câncer de mama. Além disso, são descritos os possíveis modos vibracionais e moléculas que contribuem para as diferenças dos espectros. As amostras de saliva foram coletadas antes da cirurgia de 10 pacientes com câncer de mama confirmado por exame clínico, histológico e patológico; 10 pacientes com doença benigna da mama; e 10 sem achados patológicos, o grupo controle. As amostras de saliva foram processadas e liofilizadas durante a noite. Os espectros foram medidos em um espectrômetro FTIR VERTEX 70 / 70v acoplado com diamante de platina ATR. A espectroscopia ATR-FTIR foi capaz de discriminar a saliva de pacientes com câncer de mama da saliva de pacientes com doença benigna da mama e controles. Verificaram-se maiores níveis de absorbância em pacientes com câncer de mama no número de onda 1041 cm^{-1} , com acurácia razoável e na área de $1433\text{-}1302,9\text{ cm}^{-1}$, com boa acurácia. Este aumento nos níveis de absorbância entre o câncer de mama e os outros dois grupos de pacientes foi associado a mudanças nos modos vibracionais de ácidos nucleicos, proteínas, lipídios e carboidratos. As alterações nas bandas de absorção no grupo do câncer de mama revelaram-se dependentes do fenótipo do tumor e relacionadas principalmente a proteína e ácido nucleico. Portanto, a espectroscopia FTIR foi capaz de mostrar alterações bioquímicas nos componentes da saliva como resultado da carcinogênese da mama que causa diferentes modos vibracionais nas biomoléculas. Este estudo é o primeiro a gerar espectros FTIR da saliva e derivar “*fingerprints*” químicas para fins de diagnóstico e prognóstico do câncer de mama. É importante notar que, diferentemente de outros métodos que pesquisam biomarcadores na saliva, o FTIR detecta mudanças

em um nível multi-molecular, sendo uma ferramenta promissora para o diagnóstico precoce e prognóstico do câncer de mama.

Palavras-chave: Câncer de Mama; Saliva; Espectroscopia ATR-FTIR

ABSTRACT

The increasing worldwide incidence of breast cancer and the absence of sufficient and reliable methods for early detection require search for more effective techniques. A potential candidate technique is the attenuated total reflection-Fourier transform infrared (ATR-FTIR), which consists on a vibrational spectroscopy that can effectively provide information concerning the structure and chemical composition of biological materials at the molecular level. Most of the work on the application of FTIR for breast cancer detection use tissue and blood. However, more non-invasive methods are required, for example, using saliva. This study aims to investigate differences in the spectra between the analyzed groups of patients, as well as the specific influence of the relevant clinical characteristics of breast cancer patients. Moreover, the possible vibrational modes and molecules that contribute to the spectral differences are described. Saliva samples were collected before surgery from 10 patients with confirmed breast cancer by clinical, histological, and pathologic examination; 10 patients with benign breast disease; and 10 without pathological findings, the control group. Saliva samples were processed and lyophilized overnight. The spectra were measured in a FTIR spectrometer VERTEX 70/70v coupled with platinum diamond ATR. ATR-FTIR spectroscopy was capable to discriminate breast cancer saliva from benign breast disease and control. It was found higher absorbance levels in breast cancer patients at wavenumber 1041 cm^{-1} , with reasonable accuracy, and in the area of $1433\text{-}1302.9\text{ cm}^{-1}$ region, with good accuracy. These increase in absorbance levels between breast cancer and the other two groups of patients was associated to changes in vibrational modes of nucleic acids, protein, lipids and carbohydrates. Changes in absorptions bands within breast cancer group were found to be dependent of the tumor phenotype and related mainly to protein and nucleic acid. Therefore, the FTIR spectroscopy was capable to show biochemical changes in saliva components as result of breast carcinogenesis that cause different vibrational modes in the biomolecules. This study is the first to generate FTIR spectra from saliva and derive chemical fingerprints for the purpose of diagnosis and prognosis of breast cancer. It is important to note that differently from other methods that search biomarkers in saliva, FTIR detect changes at a multi-molecular level, being a promising tool for early diagnosis and prognosis of breast cancer.

Keywords: Breast cancer; Saliva; ATR-FTIR spectroscopy.

LIST OF ABBREVIATIONS AND ACRONYMS

AJCC	American Joint Committee for Cancer
ATR	Attenuated Total Reflectance
AUC	Area Under the Curve
BC	Breast Cancer
DCIS	Ductal Carcinoma In Situ
EGFR	Epidermal Growth Factor Receptor
ER	Estrogen Receptor
FIR	Far-infrared
FTIR	Fourier transform infrared
HER2	Human Epidermal Growth Factor Receptor 2
IHC	Immunohistochemistry
IR	Infrared
IARC	International Agency for Research on Cancer
IDC	Invasive Ductal Carcinoma
MIR	Mid-infrared
MRI	Magnetic Resonance Imaging
NIR	Near-infrared
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PR	Progesterone Receptor
WHO	World Health Organization
TNM	Primary Tumor [T], Regional Lymph Nodes [N], Distant Metastases [M]
δ	Bending
ν	Stretching
as	Asymmetric
sym	Symmetric

1 INTRODUCTION

1.1 General aspects of Breast Cancer (BC)

Cancer is a group of diseases characterized by uncontrolled growth and spread of abnormal cells. It manifests as hundreds of types and subtypes, collectively affecting most organs and tissues (HANAHAHAN, 2014; KALIA, 2015). Normal cells progressively evolve into a neoplastic state, they acquire a succession of biological capabilities, or hallmarks, during the human tumors development, which occurs by multiple steps. Incipient cancer cells need to acquire traits that allow them to become tumorigenic and ultimately malignant (HANAHAHAN e WEINBERG, 2011).

The hallmarks are an organizing principle for rationalizing the complexities of neoplastic disease. They include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, genome instability, inflammation, reprogramming of energy metabolism and evading immune destruction. Tumors also contain a repertoire of recruited cells, apparently normal, that create a "tumor microenvironment" contributing to the acquisition of hallmarks traits (HANAHAHAN e WEINBERG, 2011).

According to the World Cancer Report 2014 from International Agency for Research on Cancer (IARC)/World Health Organization (WHO), cancer is one of the leading causes of morbidity and mortality worldwide, with approximately 14 million new cases and 8.2 million deaths related to cancer in 2012. Among the most incident cancers in the world, breast cancer was in second place (1.7 million) and, in the female population worldwide, was the type with the highest incidence and highest mortality, both in developing and developed countries (STEWART, 2014). In Brazil, 57.960 new cases and deaths of breast cancer were expected for 2016 (INCA, 2016).

Breast cancer is a complex and heterogeneous disease caused by several factors that contribute to its carcinogenesis, progression, metastasis and relapse (KOH *et al.*, 2014). Among these factors, it can be mention steroid hormones and

their receptors, growth factors, oncogenes and tumor suppressor genes, significant family history, life style, dietary and environmental (KEEN e DAVIDSON, 2003; SHARMA *et al.*, 2010). The breast cancer dissemination involves a succession of clinical and pathological stages beginning with carcinoma in situ, progressing to invasive lesion and culminating in metastatic disease (GHERSEVICH e CEBALLOS, 2014).

Investigate breast anatomy and histology is essential to understand the origins of breast cancer (**Figure 1**). The breast includes glandular (secretory) tissue composed by a ductal system that extend radially between the stromal tissue, formed by adipose (fatty) and fibrous connective tissue. The glandular tissue is composed of 15 to 20 lobes that comprise 20 to 40 lobules containing 10 to 100 alveoli. Each breast lobe is generally considered to exist as a single entity. The ductal system is formed by several small ducts that drain the alveoli merging to culminate in one major duct that dilates in to a lactiferous sinus and then narrows and open through a constricted orifice onto the nipple. Dilated ducts in the non-lactating breast identified by ultrasound imaging are generally associated with pathologies such as breast benign diseases (ductal ectasia, fibrocystic disease and intra-ductal adenoma), or malignancy (PANDYA e MOORE, 2011; HASSIOTOU e GEDDES, 2013)

The ducts and lobules are composed by two layers of epithelial cells, luminal epithelial and basal myoepithelial. The luminal epithelial layer is the inner layer that encapsulates the ductal lumen, and which contains cuboidal epithelial cells, some of which have the potential to further differentiate into milk secretory cells (lactocytes) during lactation. The basal myoepithelial layer is the outer layer of contractile myoepithelial cells that tightly surround the luminal layer and have properties of smooth muscle cells. This entire structure is surrounded by the basement membrane and is thought to contain bi-potent mammary stem cell populations (HASSIOTOU e GEDDES, 2013).

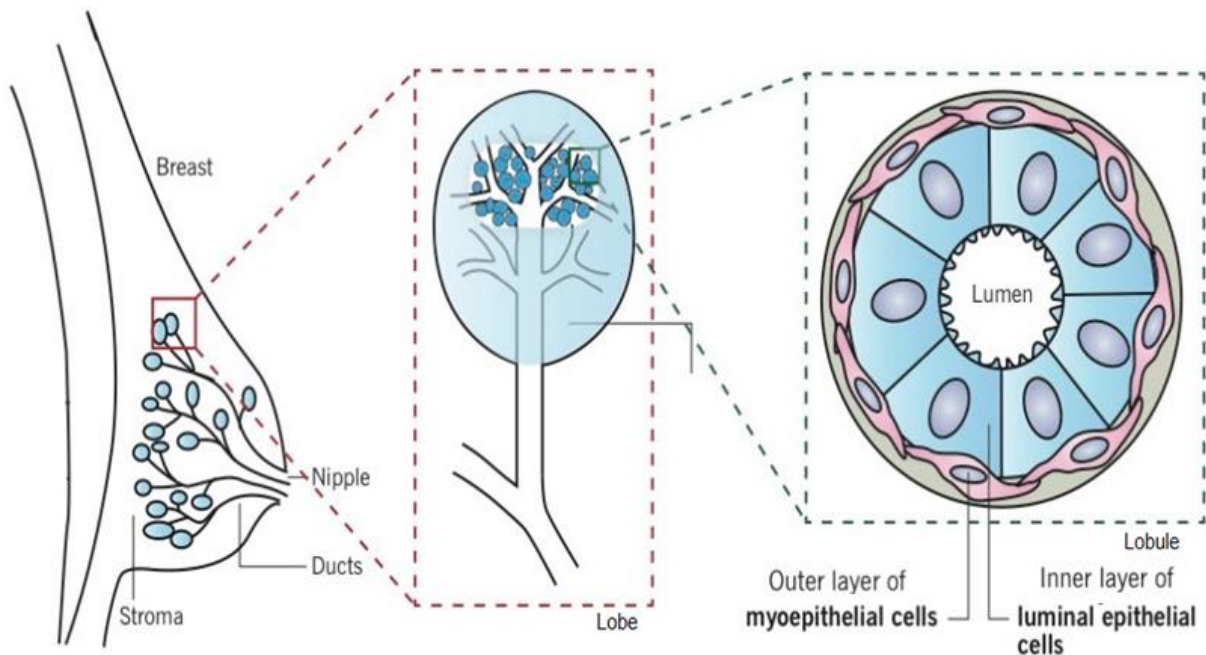


Figure 1 Breast anatomy and histology. Modified from WONG (2012) that based his figure on Pandya e Moore, 2008.

Cancer cells can exceed the duct and lobular basal membrane and invade the surrounding fatty and connective tissues of the breast, characterizing an invasive breast cancer. The invasive cancer can spread (metastatize) or not to the lymph nodes or other organs. On the other hand, when cancer cells don't exceed the basal membrane and don't invade the surrounding tissues, is called non-invasive breast cancer. Relative to origin, breast cancer can originate in the ducts (ductal carcinomas) or in the lobules (lobular carcinomas) (SHARMA *et al.*, 2010).

The invasive ductal carcinoma (IDC) is the most common type of breast cancer (80%) and the ductal carcinoma in situ (DCIS) is the most common type of non-invasive breast cancer (90%). The term "in situ" refers to cancer that has not spread past the area where it initially developed. A rare, but of good prognosis, form of breast cancer is the mucinous (colloid) carcinoma, which is formed by the mucus-producing cancer cells (SHARMA *et al.*, 2010).

The worldwide basis for breast cancer staging is the TNM (primary tumor [T], regional lymph nodes [N], distant metastases [M]) staging system, which began in 1959 as a product of the American Joint Committee for Cancer (AJCC) staging and

results reporting. This system is constantly changing through breast cancer experts and AJCC representatives reviews, and is already in the eighth edition. Based on TNM, the stage groups are 0, IA, IB, IIA, IIB, IIIA, IIIB, IIIC and IV (Attachments A, B, C, D, E) (GIULIANO *et al.*, 2017).

Briefly, pathological (p) cancer definition about the T comprises: TX, primary tumor cannot be assessed; T0, no evidence of primary tumor; Tis (DCIS), ductal carcinoma in situ (DCIS); Tis (Paget), Paget disease of the nipple not associated with invasive carcinoma and/or DCIS in the underlying breast parenchyma; T1, tumor ≤ 20 mm in greatest dimension; T2, tumor > 20 mm but ≤ 50 mm in greatest dimension; T3, tumor > 50 mm in greatest dimension; and T4, tumor of any size with direct extension to the chest wall and/or to the skin (ulceration or macroscopic nodules)(GIULIANO *et al.*, 2017).

Relative to N, there is: NX; regional lymph nodes cannot be assessed (eg, not removed for pathological study or previously removed); N0, no regional lymph node metastasis identified or isolated tumor cells only; N1, micrometastases or metastases in 1-3 axillary lymph nodes, and/or clinically negative internal mammary lymph nodes with micrometastases or macrometastases by sentinel lymph node biopsy; N2; metastases in 4-9 axillary lymph nodes, or positive ipsilateral internal mammary lymph nodes by imaging in the absence of axillary lymph node metastases; N3, metastases in 10 or more axillary lymph nodes, or in infraclavicular (level III axillary) lymph nodes, or positive ipsilateral internal mammary lymph nodes by imaging in the presence of one or more positive level I and II axillary lymph nodes, or in more than 3 axillary lymph nodes and micrometastases or macrometastases by sentinel lymph node biopsy in clinically negative ipsilateral internal mammary lymph nodes, or in ipsilateral supraclavicular lymph nodes (GIULIANO *et al.*, 2017).

Categories of M include: M0, no clinical or radiographic evidence of distant metastases; M1, distant metastases detected by clinical and radiographic means and/or histologically proven metastases larger than 0.2mm (GIULIANO *et al.*, 2017).

While anatomic TNM classifications remain the basis for the stage groups, histological tumor grade and the status of some molecular markers, like hormone receptors and human epidermal growth factor receptor 2 (HER2) status, are relevant

additional determinants of outcome and prognostic, and are now included into the staging system (GIULIANO *et al.*, 2017).

The grading system most widely used is based on that of Bloom and Richardson, as modified by Elston and Ellis, and evaluates three factors: proportion of gland or tubule formation, the degree of nuclear pleomorphism and the mitotic activity (dividing cells). Each of the 3 factors is given a score from 1 to 3 (with 1 being the closest to normal). According to the combined tumor score, the tumor is included in grade 1 or well differentiated (score 3 to 5), grade 2 (scores 6 or 7), and grade 3 (score 8 or 9), or poorly differentiated. High-grade breast cancers tend to recurrence and early metastasis while patients with low-grade tumors generally have a very good clinical outcome (VUONG *et al.*, 2014; GIULIANO *et al.*, 2017).

A molecular marker is defined as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention. Among the most well-established breast cancer molecular markers with prognostic and/or therapeutic value are hormone receptors (HR), HER2 oncogene, Ki-67 antigen and p53 proteins, and the genes for hereditary breast cancer. The main HR in breast cancer are estrogen receptor (ER) and progesterone receptor (PR), which are expressed proteins both in the epithelium and in breast stroma that bind to circulating hormones. Risk factors are closely associated with breast tumors ER+ and PR+ and may involve mechanisms related to exposure to the hormones estrogen and progesterone. HER2 is a transmembrane tyrosine kinase receptor belonging to a family of epidermal growth factor receptors structurally related to epidermal growth factor receptor (EGFR). It is encoded by ERBB2/HER2 oncogene that shows amplification in 20 to 30% of breast cancers and is considered a marker of poor prognosis (BANIN HIRATA *et al.*, 2014; VUONG *et al.*, 2014).

The Ki-67 antigen is non-histone nuclear protein that is linked to the cell cycle and is expressed in mid-G1, S, G2, and M phases of proliferating cells. Ki-67 score is the most often measured on histological sections by IHC methodology and is defined as the percentage of stained invasive carcinoma cells. The tumor protein p53 is involved in several critical pathways that are essential for genome integrity

maintenance and normal cellular homeostasis. Mutations in TP53 gene result in accumulation of altered p53 protein in nucleus, which is detected by IHC method and is an indicator of a poor clinical outcome (BANIN HIRATA *et al.*, 2014; VUONG *et al.*, 2014).

The expression of the markers ER, PR, HER2 and Ki-67, generally evaluate by immunohistochemistry (IHC), is used to classify breast cancer into molecular subtypes: luminal A (ER- and/or PR-positive / HER2-negative / Ki-67 < 14%), luminal B (ER- and/or PR-positive / HER2-negative / Ki-67 $\geq$ 14% ; ER- and/or PR-positive / HER2-positive / any Ki-67), HER2-positive (ER- and PR-negative / HER2-positive) and triple negative or basal-like (ER- and PR-negative / HER2-negative) (LI *et al.*, 2015; TOSS e CRISTOFANILLI, 2015). This classification can be used to identify high-risk phenotype and to select the most efficient therapies, since these subtypes have different clinical and pathological characteristics that are reflection of the specific molecular characteristics. The clinical behavior of triple negative is classically more aggressive than other types, like luminal A and B molecular subtypes, which are considered of best and intermediate prognosis, respectively (BANIN HIRATA *et al.*, 2014).

Then, the choice of treatment for breast cancer depends on its histological type and staging, the patient's general health and hormonal status, as well as their age and medical history. In resume, commonly therapies include surgery, radiation therapy, chemotherapy and endocrine treatments (DEPCIUCH *et al.*, 2017).

The main purpose of breast cancer screening tests is to establish early diagnostics and to apply proper treatment. Diagnostic techniques for breast cancer detection include: histopathological techniques; physical techniques (e.g., mammography, ultrasonography, magnetic resonance imaging [MRI], elastography, positron emission tomography [PET]); biological techniques; and optical techniques (e.g., photo acoustic imaging, fluorescence tomography). However, in general none of these techniques provides unique or revealing answers and have their limitations related to efficacy and production of false positive or false negative results. Basically, breast cancer diagnostic comprises four conventional techniques: histopathology,

mammography, ultrasonography and MRI (TRIA TIRONA, 2013; DEPCIUCH, KAZNOWSKA, *et al.*, 2016).

Histopathology is the current gold standard for cancer diagnosis and is the most precise diagnostic procedure for breast cancer, providing an accurate analysis of the morphological characteristics of the affected cells and tissues in tissue sections by light microscopy. However it is highly subjective, since requires the judgment of pathologists; invasive, because requires tissue samples obtained from biopsy or surgery; and time consuming (SIMONOVA e KARAMANCHEVA, 2013; BUNACIU *et al.*, 2015).

Mammography is the most widely used and studied technique for breast cancer diagnostics, it is fast and non-invasive method that involves imaging with X-rays. Nonetheless, it is associated with large error margins (can present false positive or negative results), low sensitivity and specificity and results without 100% certainty. Furthermore, it can be hazardous because it involves radiation and uncomfortable for some women (CHENG *et al.*, 2015; DEPCIUCH, KAZNOWSKA, *et al.*, 2016; PEAIRS *et al.*, 2017).

Ultrasonography is typically used as a supplement to mammography for women with dense breasts. Although it is non-invasive and fast, it has low resolution and large margin of error, and does not allow for the unique and precise localization of cancerous changes and does not provide information about types of change or how advanced cancer has become (DEPCIUCH, KAZNOWSKA, *et al.*, 2016; PEAIRS *et al.*, 2017).

MRI is currently used for screening high-risk patients in conjunction with mammography. It is non-invasive, fast and the patient is not subjected to high levels of radiation. However, it has high costs, low imaging resolution, impossibility to perform in patients with endoprosthesis, besides it is associated with the use of a contrast agent (DEPCIUCH, KAZNOWSKA, *et al.*, 2016; PEAIRS *et al.*, 2017).

Therefore, the increasing worldwide incidence of breast cancer and the absence of sufficient reliable methods for early detection requires a search for new and more effective techniques. A screening test for BC would ideally be minimally or non-invasive; simple; non-subjective; faster; with high sensitivity, specificity and

accuracy; cost-effective; high-throughput; and able to identify biochemical changes as biomarkers for cancer detection. Furthermore, the key objectives for new diagnostic systems include improving patient outcome through the identification of earlier stages, monitoring treatment and drug resistance, identifying high-risk populations for tumor progression, and subsequently reducing mortality (SIMONOVA e KARAMANCHEVA, 2013; CHENG et al., 2015; HUGHES et al., 2016).

One such candidate method is the Fourier transform infrared (FTIR) spectroscopy, a vibrational spectroscopy technique that can effectively provide, in a nondestructive and label-free manner, information concerning the structure and chemical composition of biological materials at the molecular level. The generation and progression of malignancy manifest themselves at the molecular level before morphologic changes take place. FTIR is sensitive to these changes in molecular compositions and structures according to diseased state, providing biochemical signatures of biological samples, like tissues, cells and bio fluids. FTIR spectroscopy has been used to detect carcinoma of several types of organs (EIKJE et al., 2005; ZHANG et al., 2011; BUNACIU et al., 2015).

1.2 Fourier transform infrared (FTIR) spectroscopy: technical principles, advantages and applications

Infrared (IR) spectroscopy is one variant of vibrational spectroscopy based on the vibrations of the atoms of a molecule. According to Stuart (2005) an infrared spectrum is obtained by passing infrared radiation through a sample and determining what fraction is absorbed at a particular energy. The energy at which any peak in an absorption spectrum appears corresponding to the frequency of a vibration of a part of a sample molecule.

IR radiation causes vibrations of bonds of molecules within the sample that absorbs it. The wavelength of the incident IR radiation absorbed depends on the atoms involved and the strength of any intermolecular interactions. For a molecule to show infrared absorptions its electric dipole moment must change during the vibration, molecule becomes infrared-active. These vibrational modes are quan-

titatively measurable by IR spectroscopy and then correlated directly to biochemical composition. The resultant infrared absorption spectrum is molecule specific and can be described as an infrared 'fingerprint' characteristic of the sample. For practical reasons the x-axis of infrared spectra is normally expressed in wavenumbers (cm^{-1}), the inverse of the wavelength of the infrared radiation ($\nu = 1/\lambda$) (SCHULTZ, 2002; ELLIS e GOODACRE, 2006; MOVASAGHI *et al.*, 2008; CLEMENS *et al.*, 2014).

The functional groups within the sample that absorb the infrared radiation can vibrate in one of a number of ways, either stretching, bending, deformation or combination vibrations. Vibrations can involve mainly a change in bond length (stretching) or bond angle (bending). Some bonds can stretch or bend in-phase (symmetrical stretching) or out-of-phase (asymmetric stretching) Bending vibrations include scissoring, rocking, wagging and twisting (STUART, 2005; ELLIS e GOODACRE, 2006).

The IR spectral region is adjacent to the visible spectral region on the electromagnetic spectrum (**Figure 2**), ranges from the red end of the visible spectrum at about 13000cm^{-1} to the onset of the microwave region at a wavelength of 10 cm^{-1} . The spectrum of the IR region is conventionally divided into three parts: near-infrared (NIR) region from 13000 to 4000 cm^{-1} , the mid-infrared (MIR) region from 4000 to 400 cm^{-1} and the far-infrared (FIR) region approximately from 400 to 10 cm^{-1} (STUART, 2005; BARTH, 2007; BUNACIU *et al.*, 2015)

IR applications employ mainly the MIR region, because it is the most informative spectrum for bio samples since it contains many absorptions corresponding to fundamental vibrations of the molecular species. The NIR and FIR regions have also provided some benefits and important information about certain materials (STUART, 2005; BARTH, 2007; BUNACIU *et al.*, 2015).

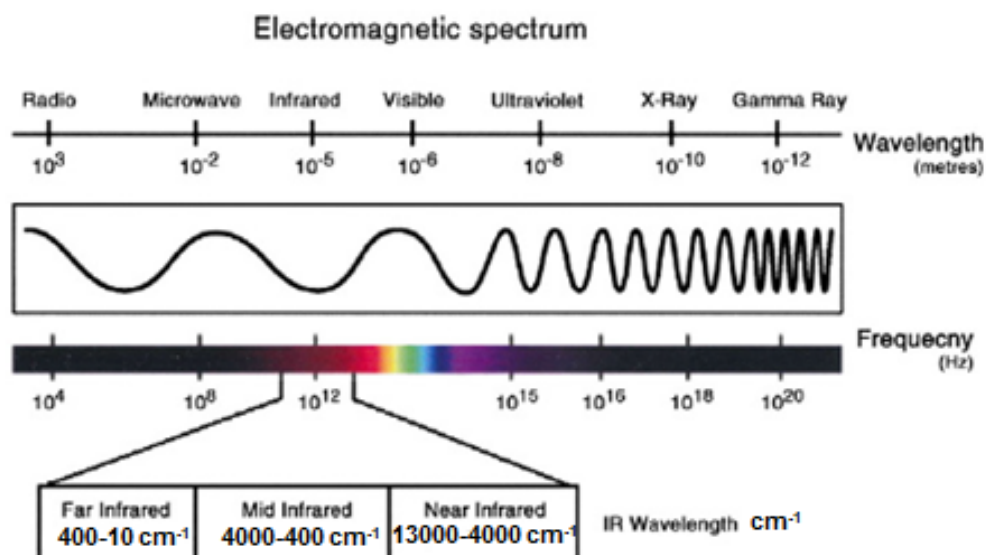


Figure 2 The electromagnetic spectrum. Reproduced with modifications from Noreen et al., 2012.

The MIR region ($4000\text{--}400\text{ cm}^{-1}$) can be approximately divided into four regions (**Figure 3**): the X–H stretching region ($4000\text{--}2500\text{ cm}^{-1}$), the triple-bond region ($2500\text{--}2000\text{ cm}^{-1}$), the double-bond region ($2000\text{--}1500\text{ cm}^{-1}$) and the fingerprint region ($1500\text{--}600\text{ cm}^{-1}$). The fundamental vibrations in the $4000\text{--}2500\text{ cm}^{-1}$ region are generally due to O–H, C–H and N–H stretching. Triple-bond stretching absorptions of $\text{C}\equiv\text{C}$ and $\text{C}\equiv\text{N}$ fall in the $2500\text{--}2000\text{ cm}^{-1}$ region due to the high force constants of the bonds. The principal bands in the $2000\text{--}1500\text{ cm}^{-1}$ region are due to $\text{C}=\text{C}$, $\text{C}=\text{O}$ and $\text{C}=\text{N}$ stretching. The fingerprint region $1500\text{--}650\text{ cm}^{-1}$ corresponds to absorption of bending and skeletal vibrations, this region contains the fundamental vibrational modes of key chemical bonds. The nature of a group frequency may generally be determined by the region in which it is located (STUART, 2005).

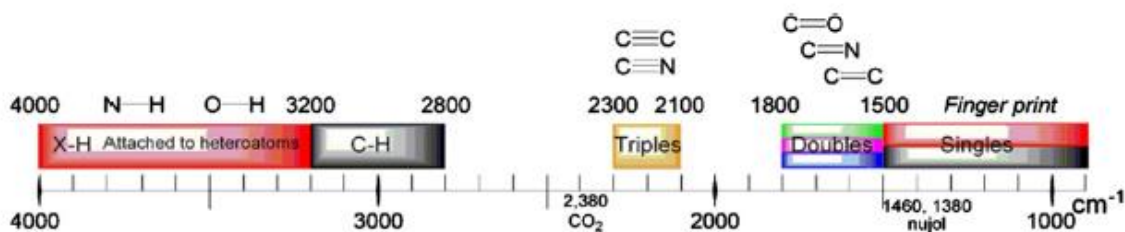


Figure 3 Major absorptions of bonds in organic molecules in the MIR region ($4000\text{--}400\text{ cm}^{-1}$) (NOREEN *et al.*, 2012)

FTIR spectroscopy is a type of infrared spectroscopic, so a vibrational technique, which refers to the study of the absorption of electromagnetic waves in MIR region ($4000\text{--}400\text{ cm}^{-1}$) (ANDREW CHAN e KAZARIAN, 2016). FTIR spectroscopy is based on the idea of the interference of radiation between two beams to yield an interferogram that is a signal produced as a function of the change of path length between the two beams. The basic components of the most common interferometer used in FTIR spectrometry, the Michelson interferometer, are shown schematically in **Figure 4**. The radiation emerging from the source is passed through an interferometer, which has a fixed and a movable mirror, to the sample before reaching a detector. Upon amplification of the signal, in which high-frequency contributions have been eliminated by a filter, the data are converted to digital form by an analog-to-digital converter and transferred to the computer for Fourier-transformation (STUART, 2005).

The Fourier transformation can be considered simply as a mathematical means of extracting the individual frequencies from the interferogram for final representation in an IR spectrum. Spectrometer performs two Fourier transformations, one by the interferometer and one by the computer (BARTH, 2007; BEEKES *et al.*, 2007).

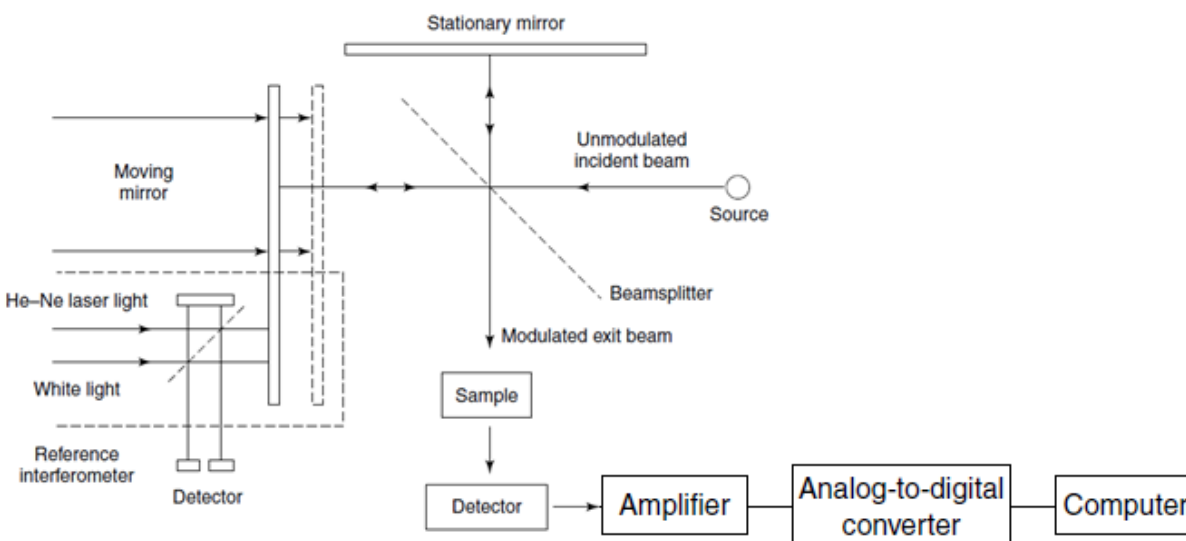


Figure 4 Schematic representation of a Michelson interferometer. Modified from Stuart (2005)

FTIR spectra can be acquired mainly in three different experimental configurations: transmission, transflection or attenuated total reflection (ATR). The first one operates by transmitting IR radiation through sample-substrate before the resulting radiation is detected. Although transmission mode experiments are the most common, the spectra are subject to a variety of physical effects occurring when measuring the sample. Transflection mode experiments detect the absorbed IR radiation after it is transmitted through the sample, reflected off by the substrate, and transmitted back through the sample. In transflection, there is formation of a standing wave perpendicular to sample surface that leads to spectral changes not related to the biochemistry of sample (BUNACIU *et al.*, 2015; LIMA, C. A. *et al.*, 2015).

Finally, the ATR mode operates on the principles of total internal reflection (**Figure 5**). The sample is placed into direct contact with a crystal (KRS-5, ZnSe, diamond, Si or Ge) with refractive index higher than sample, inducing total internal reflection of incident radiation, which is attenuated and penetrates into the sample as an evanescent wave. This evanescent wave is then altered and passed back to the detector on the FTIR spectrometer, providing a single spectrum, which represents the average signal from the sample area that light passed through. The penetration depth can be controlled and allows measurement from aqueous body fluids. ATR-FTIR provides to be a simple, reagent-free and powerful tool for analyze biological fluids and dry films samples with little or no preparation method. It enables a sample to be examined directly in the solid, liquid, or gas state without further preparation simply by transmitting the sample with infrared radiation. Furthermore, ATR-FTIR presents high signal-to-noise ratio (SNR) and does not present unwanted spectral contributions compared to transflection and transmission configurations (ELLIS e GOODACRE, 2006; BARTH, 2007; LANE e SEE, 2012; CLEMENS *et al.*, 2014; BUNACIU *et al.*, 2015; LIMA, C. A. *et al.*, 2015).

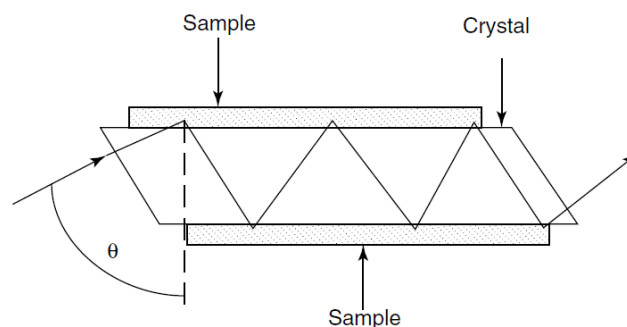


Figure 5 Scheme of a typical attenuated total reflectance cell. θ =the angle of incident radiation. Reproduced from Stuart (2005)

The raw spectra obtained from FTIR analysis can be manipulated of various ways to carry out quantitative analysis, for instance baseline correction and differentiation (derivatives). It is usual to use a baseline joining the points of lowest absorbance on a peak, preferably in reproducibly flat parts of the absorption line. Then it is used the absorbance difference between the baseline and the top of the band (STUART, 2005). Spectra may also be differentiated, for example in its first and second derivative (**Figure 6**), especially the last one gives a negative peak for each band and shoulder in the absorption spectrum. The second-derivative accentuates the bands, resolving broad and overlapped bands into individual, reduces the background interference, increases the specificity of absorption peaks for certain molecules and thus increases the accuracy of analysis by revealing the genuine biochemical characteristics (LEWIS et al., 2010; RIEPPO et al., 2012; ZELIG et al., 2015). Savitzky-Golay is an example of smoothing method used in differentiation that removes the background noise, preserving shapes of peaks. This method performs a local polynomial regression around each point, and creates a new, smoothed value for each data point. The parameters which may be adjusted include the type of smoothing function (polynomial order), and the width of the smoothing interval (number of points used in the regression) (ENKE e NIEMAN, 1976).

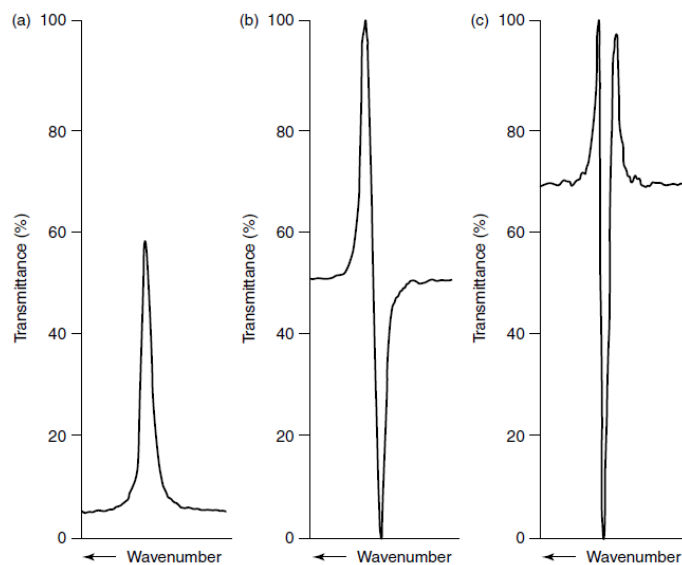


Figure 6 Differentiation of spectra: (a) single absorption peak; (b) first derivative; (c) second derivative. Reproduced from Stuart (2005)

Some advantages of FTIR spectroscopic are: (1) it is fast, reproducible and non-destructive; (2) it is relatively simple to operate, requires little technical expertise to run the instrument; (3) it requires only a small amount of sample for analysis; (4) when necessary, the sample preparation is very easy and quick; (5) it is label and reagent free technique, no consumable costs or chemical reagents are required; (6) it allows rapid and non-invasive detection of biochemical changes at the molecular level; (7) it is specific in differentiating biological materials; (8) it is a computational method which allows automated and repetitive analyses, leading to rapid and objective evaluation of the sample; (9) it is an inexpensive technique, with good cost-effective; (10) It allows for investigations in vivo, avoiding lengthy periods, stages of sampling and painful biopsies; (11) It can support surgeons in order to reduce the waiting time for pathological results; (12) it can be used during or before surgical operation; (13) it can be used for diagnostics of conditions, even in the very early stages of the disease; (14) it can be an excellent tool for the monitoring of the disease and its treatment; (15) it has many applications fields(ZHANG *et al.*, 2010; ELKINS, 2011; SIMONOVA e KARAMANCHEVA, 2013; DEPCIUCH, KAZNOWSKA, *et al.*, 2016).

The common types of molecules that may be studied using infrared spectroscopy are organic, inorganic, polymeric and biological. Briefly, pharmaceutical, food, agricultural, pulp and paper, paint, environmental, forensic and biomedical fields have exploited this technique. In pharmaceutical field, IR spectroscopy is important to evaluate the raw materials used in production, the active ingredients and the excipients. In food field, both MIR and NIR techniques can be used to obtain qualitative and quantitative information about food samples that are complex mixtures, with the major components being water, carbohydrates, proteins and fats. An example of agricultural application is the analysis of commercial grains by NIR spectroscopy, it is important to quantitate the composition that is mainly water, carbohydrates, protein, minerals, oil and fiber (STUART, 2005).

In the pulp and paper industries, IR spectroscopy it is important in quality control, because additives may be identified in paper in the MIR region. In paint industry, the quality control, failure analysis and product improvement are the purposes of IR using. Environmental problems relative to air, water and soil can be analyzed by IR spectroscopy (STUART, 2005). In forensic science, IR is used for analysis of samples recovered by investigators at crime scenes, including biological samples as blood, earwax, feces, fingernails, fingerprints, tears, hair, nasal, mucus, vaginal mucus, saliva, semen, and urine; and other evidences such as alcohol, drugs, fibers, paints, industrial chemicals, common household and laboratory chemicals, solvents and explosives (ELKINS, 2011). The biology and medicine field comprises many applications mainly in MIR region that are the focus of this study.

The application of IR spectroscopy in biological field dates back to 1950, with analysis of the conformational structure of polypeptides and proteins, and it was gradually also extended to the analysis of nucleic acids, lipids and carbohydrates. IR spectroscopy analysis of biological samples are more complex due to the superposition of all infrared-active vibrational modes of the various molecules present. Therefore, IR spectroscopy provides information about characteristic frequencies, intensities, and bandwidths in infrared spectra that allow the identification of functional groups and molecular structures of biological molecules,

and then the characterization of proteins, lipids, nucleic acids and carbohydrates of the material (BEEKES *et al.*, 2007).

For biological samples, the most important spectral regions measured are typically the fingerprint region, from 600 to 1450 cm^{-1} and the amide I/amide II region from 1500 to 1700 cm^{-1} . The higher-wavenumber region, 2550 to 3500 cm^{-1} is associated with stretching vibrations such as S-H, C-H, N-H and O-H, whereas the lower-wavenumber regions typically correspond to bending and carbon skeleton vibrations as C-O, C-N and C-C. The lipid, protein, carbohydrates and nucleic acids contents and their chemical structure can be found in specific peaks or regions along the MIR spectra. For instance, the wavenumber region 2800–3050 cm^{-1} is related to CH_2 and CH_3 stretching vibrations from fatty acid and the region 1500–1750 cm^{-1} (the amide I and II bands) are ascribable to C=O, N-H and C-N from proteins and peptides. (ELLIS e GOODACRE, 2006; MOVASAGHI *et al.*, 2008; BAKER *et al.*, 2014)

FTIR spectroscopy not only differentiates biological samples based on their characteristic spectral properties reflecting the chemical composition and structure, but informs about pathological biochemical alterations resulting from diseases. From a diagnostic perspective, FTIR spectroscopic fingerprints can be used to discriminate between different types of cells, tissues and body fluids, as well as detecting and discriminating different diseases or disease progression (BEEKES *et al.*, 2007; CLEMENS *et al.*, 2014).

A significant number of studies has used FTIR spectroscopy as a potential biochemical and metabolic fingerprinting tool for the rapid detection and diagnosis of several diseases or dysfunctions, such as arthritis (ELLIS e GOODACRE, 2006), cancer, prion diseases, bone diseases, atherosclerosis, kidney stones and gallstones, diabetes, osteoarthritis (KRAFFT *et al.*, 2009) and depressive disorders (DEPCIUCH, SOWA-KUCMA, *et al.*, 2016).

It has already been shown that FTIR spectroscopy can be applied to a variety of biological samples, including tissue (formalin-fixed, paraffin-embedded or fresh), cells (fixed or live), microorganisms, blood, saliva, urine, tears, earwax, feces, fingernails,

hair, nasal mucus, vaginal mucus, breast milk and semen (ELKINS, 2011; TREVISAN et al., 2014).

Since the early 1990s several researchers have used FTIR spectroscopy to distinguish between normal and neoplastic tissues. The constant development of new technology and analytical methods enables moves toward diagnosing neoplastic change at earlier stages and potential to do in vivo (KENDALL *et al.*, 2009).

The FTIR spectroscopy is capable to show the structural changes of the cells at the molecular level in several human cancers. These changes are caused by different vibrational modes in the molecules of the cells and tissues as result of carcinogenesis. Then, FTIR spectra could show normal or malignant cells according to their spectral characteristic appearance, since the unique vibrational modes of major functional groups are characterized by the changes in the spectra (BUNACIU *et al.*, 2015). The most significant differences in the spectrums of normal and cancerous tissues and cells have been observed in the wavenumber region between 400 and 4000 cm^{-1} . Thus, the mid- infrared region was reported as a major cancer diagnosis indicator (SIMONOVA e KARAMANCHEVA, 2013).

FTIR spectroscopy has much more potential than only discriminate normal or malignant samples, FTIR micro-spectroscopic imaging with ability to image tissues and cells without necessity of sample staining or fixation, could help to guide biopsy procedures, to characterize cancers as those likely to progress, and to guide surgical resection by detection of tumor margins (MACKANOS e CONTAG, 2010).

A wide range of biological studies of cancer have used FTIR spectroscopy, which shown promise as a sensitive diagnostic tool to detect and discriminate different types of cancer, such as: breast (BACKHAUS *et al.*, 2010; KHANMOHAMMADI *et al.*, 2010; CHEN *et al.*, 2014; ZELIG *et al.*, 2015; DEPCIUCH, KAZNOWSKA, *et al.*, 2016), ovarian (GAJJAR *et al.*, 2013; LIMA, K. M. *et al.*, 2015), lung (YANO *et al.*, 2000; LEWIS *et al.*, 2010; SUN *et al.*, 2013), gastric (COLAGAR *et al.*, 2011; SHENG, WU, *et al.*, 2013), head and neck (oral, oropharyngeal, and laryngeal) (MENZIES *et al.*, 2014), skin (EIKJE *et al.*, 2005), thyroid (ZHANG *et al.*, 2011), colon (NALLALA *et al.*, 2012), colorectal (DONG *et al.*, 2014), brain (NOREEN *et al.*, 2012; NOREEN *et al.*, 2013), cervical (EL-TAWIL *et al.*, 2008), hepatocellular

(ZHANG *et al.*, 2013), leukemia (SHENG, LIU, *et al.*, 2013), prostate (PEZZEI *et al.*, 2010; HUGHES *et al.*, 2014), brain metastasis (KRAFFT *et al.*, 2006; NOREEN *et al.*, 2011) and esophagus (WANG *et al.*, 2003; MAZIAK *et al.*, 2007).

Specifically for breast cancer, FTIR spectroscopy has been used for various purposes: detection (BACKHAUS *et al.*, 2010; KHANMOHAMMADI *et al.*, 2010; CHEN *et al.*, 2014; ZELIG *et al.*, 2015; DEPCIUCH, KAZNOWSKA, *et al.*, 2016); monitoring treatment (DEPCIUCH *et al.*, 2017); identification of brain metastases (KRAFFT *et al.*, 2006); intraoperative detection of sentinel lymph node metastases (TIAN *et al.*, 2015); characterization of breast cancer cells as well as the tumor microenvironment (BENARD *et al.*, 2014); grade discrimination in DCIS (ductal carcinoma in situ) and IDC (invasive ductal carcinoma) (REHMAN *et al.*, 2010); determination of premalignant cancer-like phenotype in normal women (MALINS *et al.*, 1995); evaluation of multidrug resistance (KRISHNA *et al.*, 2006); evaluation of anticancer drugs effects on breast cancer cell lines (MDA-MB-231, MCF-7, SK-BR-3 and HBL-100) (MIGNOLET e GOORMAGHTIGH, 2015); analysis of extracellular matrix by FTIR imaging on histopathological specimens (KUMAR *et al.*, 2013) and monitoring chemotherapy effects by FTIR micro-spectroscopic (ZAWLIK *et al.*, 2016).

The most FTIR spectroscopy studies that analyzes breast cancer has used as biological sample, tissues (FABIAN *et al.*, 2006; MEHROTRA *et al.*, 2007; DEPCIUCH, KAZNOWSKA, *et al.*, 2016; VERDONCK *et al.*, 2016), cell lines (LANE e SEE, 2012; WU *et al.*, 2015; GAVGIOTAKI *et al.*, 2016) and blood (BACKHAUS *et al.*, 2010; ZELIG *et al.*, 2015). There are no works demonstrating the use of FTIR spectroscopy for breast cancer diagnosis and prognosis, with saliva as the biological sample.

A FTIR spectroscopy study compared normal non-cancerous breast tissue, breast cancer tissues and normal breast tissues around the cancerous breast region. Spectra collected from breast cancer patients shows changes in carotenoids, fats, carbohydrate and protein levels (e.g., lack of amino acids, changes in the concentration of amino acids, structural changes) in comparison with normal breast tissues (DEPCIUCH, KAZNOWSKA, *et al.*, 2016).

Zelig et al. (2015) studied the detection of breast cancer by analyzing plasma and peripheral blood mononuclear cells (PBMCs) using FTIR micro-spectroscopy. Several bands in the FTIR spectra of both blood components significantly distinguished patients with and without breast cancer, with a sensitivity of ~90% and a specificity of ~80% for breast cancer detection. They also observed in cancer group an influence of several clinical parameters, such as the involvement of lymph nodes, on the infrared spectra.

An ATR-FTIR and gold nanotechnology study analyzed structural differences between cancerous breast cells (MCF-7 line) and normal breast cells (MCF-12F line). They found shifts of wavenumber and higher peak intensities in spectra between the normal cells and cancerous cells, with significant changes in the spectrum range from 2854–2956 cm^{-1} (LANE e SEE, 2012).

1.3 Saliva: a promising biological fluid

Saliva is a complex and dynamic biological fluid that, as mentioned, is also used in FTIR spectroscopy. Detection and quantitative analysis of biochemical characteristics of saliva in MIR region allows identify components with highly specific bands at a particular set of wavenumbers depending on the molecular composition and structure (KHAUSTOVA *et al.*, 2009).

Saliva is composed by 98 % water and 2 % of other important compounds, such as electrolytes (Na, K, Ca, Mg, hydrogen carbonates, and phosphates), mucus (mucopolysaccharides and glycoproteins), antiseptic substances (hydrogen peroxide, IgA), and several enzymes (α -amylase, lysozymes, lingual lipase). There are two main types of salivary glands in the mouth: minor glands (approximately 600) positioned throughout the oral cavity, and the major glands (submandibular, sublingual and parotid) located in and around the mouth and throat. The type of saliva that each gland produces varies between them. For example, the saliva produced by the sublingual, submandibular, and minor mucosal glands is rich in mucins (MUC5B and MUC7) and contains only a small amount of amylase. In contrast, saliva from the parotid gland is rich in amylase (20%), proline-rich proteins (60%), and

phosphoproteins - statherin (7%), but without representation of mucins (PINK *et al.*, 2009; PFAFFE *et al.*, 2011).

This biological fluid has many functions, it participates in smooth digestion and the ingestion of food; mediates taste sensations; and cooperates in repairing soft tissue. Most importantly, a whole range of immune and defensive processes take place via the salivary proteins, since it has antimicrobial, immunomodulatory and anti-inflammatory properties, as well as several other relevant features (PINK *et al.*, 2009; ABRAO *et al.*, 2016).

Saliva harbors a wide spectrum of proteins/peptides, nucleic acids, electrolytes, and hormones that originate from multiple local and systemic sources. Most of the organic compounds in saliva are produced locally in the salivary glands, but some molecules pass into saliva from blood by transcellular (e.g., passive diffusion of lipophilic molecules such as steroid hormones and active transport of proteins via ligand-receptor binding) or paracellular (e.g., extracellular ultrafiltration) means (PFAFFE *et al.*, 2011; WANG *et al.*, 2016).

It has been well recognized that saliva reflects the physiological state of the body, including the emotional condition, and endocrine, nutritional and metabolic changes. Circulating biomolecules that originate from a diseased process may eventually be transported into the salivary glands, which will then consequently modify the composition of saliva. Then, salivary biomarkers can be exploited for the early diagnosis of some oral and systemic diseases, such as: caries; periodontal diseases; oral cancer; diabetes; cardiovascular, autoimmune and renal diseases; pancreatic, breast, lung and prostate cancers, among others (MALAMUD, 2011; ABRAO *et al.*, 2016; WANG *et al.*, 2016; ZHANG *et al.*, 2016).

The main goal of saliva analysis to diagnose systemic malignancies has been the discovery, verification, and validation of a panel of biomarkers that can be used in early detection of cancer (WANG *et al.*, 2016). For example in breast cancer researches using saliva, Bigler *et al.*, 2009 suggests that the protein expression of the receptor tyrosine kinase HER-2 in saliva can be helpful to measure patient response to chemotherapy, and Zhang *et al.*, 2010 reports the discovery and validation of one

protein biomarker (carbonic anhydrase 6 (CA6) protein) and eight mRNA biomarkers for the noninvasive detection.

Therefore, the association of physiological illness with physiological activity of the salivary glands suggests the possibility of using the saliva as a diagnostic medium, which possesses a number of biochemical and logistical advantages over analysis of other biological fluids. Saliva is simple, fast and safe to collect; is easy to store; is non-invasive; may be collected repeatedly without discomfort, risk and pain to the patient, is simple to prepare, involving centrifugation before storage and the addition of a cocktail of protease inhibitors (AGHA-HOSSEINI et al., 2009; BIGLER et al., 2009; KHAUSTOVA et al., 2009; ABRAO et al., 2016).

Taking into account the alarming epidemiological data on breast cancer and the problems of current diagnostic methods; the potential of FTIR as a non-invasive, non-destructive, inexpensive and label-free technique; and the many biochemical and logistical advantages of saliva as biological fluid for the diagnosis of various diseases, studies based on these pillars are of extreme potential and importance. The utility of FTIR spectroscopy in the investigation of saliva from breast cancer patients could be since early diagnosis, until monitoring of the disease and its treatment by analysis of the entire biochemical signature of the sample, including proteins, lipids, nucleic acids and carbohydrates.

2 OBJECTIVES

2.1 General Objectives

The aim of the study was to investigate the utility of ATR-FTIR spectroscopy as a diagnostic and prognostic tool for breast cancer using saliva.

2.2 Specific Objectives

- Identify specific infrared wavenumbers or intervals in the spectra that significantly discriminate breast cancer saliva from benign breast disease and control saliva;
- Determine whether there are significant biochemical changes in saliva composition by clinical parameters within the group of breast cancer patients;
- Describe the possible vibrational modes and molecules which may contribute to the spectral differences.

3 MATERIAL AND METHODS

3.1 Ethical aspects and study subjects

The study was conducted at a Brazilian university hospital (HC-UFU, Uberlandia, Minas Gerais, Brazil) under local Human Research Ethics Committee (protocol number 064/2008) (**Attachment E**) and based on the standards of the Declaration of Helsinki. All participants signed a free and informed consent form. The subjects were randomly selected from population before perform routine breast cancer screening and/or surgery. Exclusion criteria were age below 18 years, primary tumor site other than the breast, and physical and/or mental inability to respond the tools necessary to data collection. The study group included 30 subjects: 10 with confirmed breast cancer by clinical, histological, and pathologic examination; 10 with some benign breast disease, like fibroadenomas, atypical ductal hyperplasia, papilloma or other; and 10 without pathological findings, the control group. In this study was used the tumor–node–metastasis (TNM) cancer classification, which is according to the American Joint Committee on Cancer (AJCC) and the International Union for Cancer Control (UICC). This classification evaluate the extent of the primary tumor (T), regional lymph nodes (N), and distant metastases (M) and provides staging based on T, N, and M (GIULIANO *et al.*, 2017).

3.2 Sample collection and preparation

For each participant, saliva samples were collected before surgery in *Salivette*® tubes (Sarstedt, Germany), consisting of a neutral cotton swab and a conical tube. The patient chewed the swab for three minutes, which was then returned to the tube that was covered with a lid. Then the saliva from the swab was recovered by centrifugation for 2 minutes at 1000 x g and stored at -20°C. Then the saliva samples (200 µL) were lyophilized overnight. This freeze-drying of the samples removes the strong water infrared light absorption from spectra which may mask the

signal from the sample and may reduce the intensity of the compounds under investigation (KHAUSTOVA *et al.*, 2010; ZELIG *et al.*, 2015).

3.3 ATR-FTIR spectroscopy

The spectra were measured in the 4000 to 400 cm^{-1} wavenumber region using a FTIR spectrometer VERTEX 70/70v (Bruker Corporation, Germany) coupled with Platinum diamond ATR, which consist of a diamond disc as an internal-reflection element. The lyophilized sample was placed on the ATR crystal and then the spectrum was recorded. The spectrum of air was used as a background before each sample analysis. Background and sample spectra were taken at a room with temperature around 21-23°C, at a spectral resolution of 4 cm^{-1} and to each measurement 32 scans were performed.

3.4 Spectral data preprocessing

The original FTIR spectra were normalized and the baseline was corrected using OPUS software. This software was also used to calculate absorbance of area under spectral regions that correspond to specific saliva components, applying parameters already described (KHAUSTOVA *et al.*, 2010).

Second differentiation spectra from the original were carried out using Savitzky-Golay method in Origin 9.1 software in order to accentuates the bands, resolve overlapped bands and increases the accuracy of analysis by revealing the genuine biochemical characteristics (LEWIS *et al.*, 2010; ZELIG *et al.*, 2015). In these smoothing pretreatment, the parameters of the Savitzky-Golay filter such as the polynomial order and points of window was chosen in order to find the relatively optimum smoothing effect. The parameters were set as 2 for polynomial order and 20 for points of window.

3.5 Statistical analysis

After the spectral preprocessing, the original and derivative values were used on the statistical analysis. First, values of absorbance at specific wavenumbers and spectral regions was submitted to normality test. According to the results, parametric tests for variables with normal distribution, or non-parametric tests for variables without normal distribution were performed. The specific tests applied are indicated on the legend of the figures. It was assumed p values less than 0.05 as statistically significant and confidence intervals (CI) of 0.95. Statistical analysis were carried out using GraphPad Prism versions 5.00 and 7.03 (GraphPad Software, USA).

It is relevant to inform about a specific test performed in Figure 13 and 14. We applied an unpaired multiple t-test to all wavenumbers of the second derivative spectra comparing the means between breast cancer patients from a clinical subgroup and breast cancer patients from other clinical subgroup. This multiple test applies an unpaired t-test for each line between the two subgroups (each line represents each wavenumber of the spectrum) and provides the p-value for each compared line (wavenumber). Then, the statistically significant p-values along the spectra were placed in a ranking of the most statistically significant (lowest p-value, that is, the number 1 in rank), to the least significant (higher p-value, that is, the number 25 in rank for example). Thus, figures 13A, 13C, 14A and 14C describe all ranked p-values resulting from the multiple t-test (Y axis) which were significant at each wavenumber of the spectrum (X axis).

4 RESULTS

4.1 Study subjects characterization

The main characteristics of the study subjects are shown in **Table 1**, which describes basic demography characteristics of study groups and in **Table 2**, that report clinical, hormonal, diagnostic and therapy characteristics of patients with breast cancer.

The groups of breast cancer, benign breast disease and control patients consisted of 10 women each one, with a mean age of 53.3 ± 11.2 , 41.5 ± 4.2 and 43.2 ± 16.0 years, respectively. History of smoking was similar between the groups of patients, history of alcoholism was found only in benign and control patients, and family history of breast cancer was reported only on the group of cancer patients.

Table 1 Demography characteristics of breast cancer, benign breast disease and control patients

	Breast Cancer n=10	Benign n=10	Control n=10
Age (years)			
Range	42.0 – 75.0	33.0 – 49.0	22.0 – 63.0
Average $\pm$ SD	53.3 ± 11.2	41.5 ± 4.2	43.2 ± 16.0
History of Smoking (%)	30%	40%	30%
History of Alcoholism (%)	0	40%	10%
Family History of Breast Cancer (%)	40%	0	0

Table 2 Clinical, hormonal, diagnostic and therapy characteristics of breast cancer patients

Variable	Patients (n=10)	
	n	%
Histological subtype		
Invasive ductal carcinoma	6	60
In situ ductal carcinoma	3	30
Mucinous carcinoma	1	10
Histological grade		
G2	5	50
G3	2	20
NR	3	30
Primary tumor		
pTX	1	10
pTis	3	30
pT1	4	40
pT2	2	20
Regional lymph nodes		
pNX	2	20
pN0	5	50
pN1	1	10
pN2	1	10
NR	1	10
Distant metastases		
pM0	7	70
NR	3	30
TNM Staging		
0	2	20
I	1	10
II	2	20
NR	5	50
Status ER		
Positive	8	80
NR	2	20
Status PR		
Positive	8	80
NR	2	20
Status HER2		
Positive	2	20
Negative	6	60
NR	2	20
p53		
Positive	8	80
NR	2	20
Ki67		
≤ 14%	5	50
> 14%	3	30
NR	2	20
Molecular phenotype		
Luminal A	4	40
Luminal B	4	40
NR	2	20
Therapy		
Surgery (S)	1	10
S + Radiotherapy (RT)	1	10
S + RT + Hormone therapy (HT)	3	30
S + RT + HT + Chemotherapy (CT)	5	50

Abbreviations: G1, grade 1, G2, grade 2, G3, grade 3, NR, not reported; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth; p53, tumor protein p53; ki67, antigen ki67

4.2 FTIR analysis of saliva spectra between breast cancer, benign and control patients

Firstly, the spectral absorptions of saliva among breast cancer, benign and control patients were analyzed. The averages of the infrared original spectra of saliva for each group of patients are presented in **Figure 7A**, showing the region 1800-800 cm^{-1} that comprises the main biochemical data. The second-derivative spectra were also analyzed, since additional vibrational modes can be detected in this analysis that provide more detailed information. Averages of the second-derivative infrared spectra of saliva for each group of patients are presented in **Figure 7B**, also showing the region 1800-800 cm^{-1} .

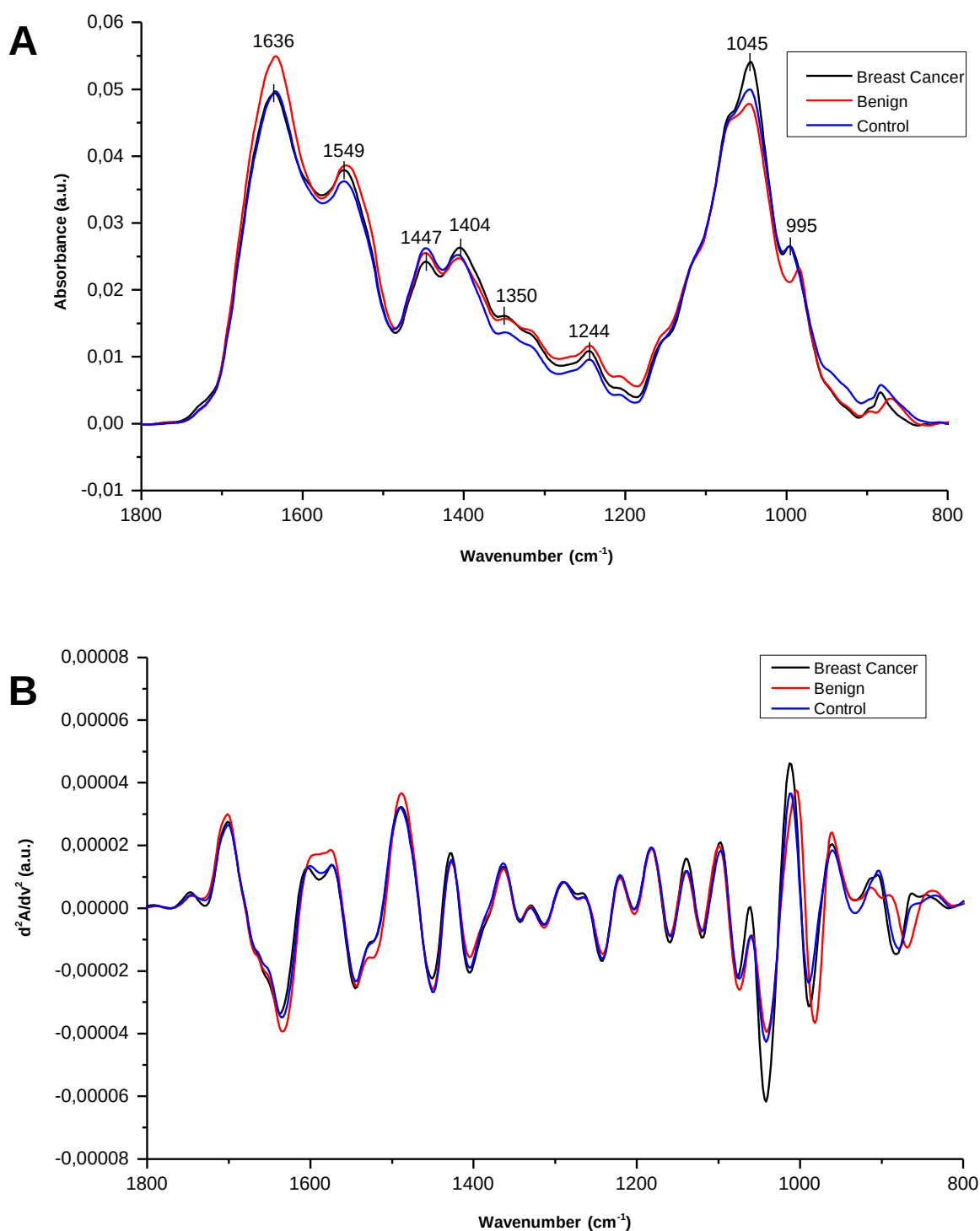


Figure 7 FTIR spectra for breast cancer, benign breast disease and control saliva. (A) Average original spectra and (B) average second derivative spectra between wavenumbers 1800 cm^{-1} and 800 cm^{-1} for breast cancer (black line), benign breast disease (red line) and control saliva (blue line). The absorbance bands of the major functional groups in biomolecules are indicated in A and detailed in Table 3.

A resume of the assignments of main wavenumbers and their respective vibrational modes and molecular sources composing saliva original spectra shown in **Table 3**. In general it was possible to verify that the averages of the infrared original spectra exhibit absorption bands associated with proteins, nucleic acids, lipids and carbohydrates. The protein content is mainly attributed to wavenumbers at 1636 cm^{-1} and 1549 cm^{-1} that corresponds to amide I and amide II, respectively. Asymmetric bending of methyl groups at 1447 cm^{-1} ; $\nu_{\text{sym}}\text{COO}^-$ of amino acids and $\delta_{\text{sym}}\text{CH}_3$ of methyl groups at 1404 cm^{-1} ; and $\nu_{\text{as}}\text{PO}_2^-$ of phosphorylated protein and $\nu\text{C-N}$ of amide III at 1244 cm^{-1} are also associated to protein content. Wavenumbers at 1244 cm^{-1} , 1045 cm^{-1} and 995 cm^{-1} are due to vibrations of functional groups such as $\nu_{\text{as}}\text{PO}_2^-$, C-O/C-C and uracil ring present in nucleic acids. Vibrational modes of CH_2 and COO^- of fatty acids, at 1447 cm^{-1} and 1404 cm^{-1} respectively, and $\nu_{\text{as}}\text{PO}_2^-$ of phospholipid correspond to lipid content. Carbohydrates content is related to $\nu\text{OH}/\delta\text{OH}$ and $\nu\text{C-O}$, $\delta\text{C-O}$, both at 1045 cm^{-1} .

Table 3 Assignments of main wavenumbers indicated in the average original saliva ATR-FTIR spectra of the Fig. 7A

Peak (cm ⁻¹)	Proposed vibrational mode	Molecular source
1636	Amide I (ν C=O, ν C–N, δ N–H)	Protein
1549	Amide II (δ N–H, ν C–N stretching)	Protein
1447	CH ₂ symmetric bending (δ_{sym} CH ₂)	Lipid
	CH ₃ asymmetric bending (δ_{as} CH ₃)	Protein (methyl groups)
1404	CH ₃ symmetric bending (δ_{sym} CH ₃)	Protein (methyl groups)
	COO ⁻ symmetric stretching (ν_{sym} COO ⁻)	Lipid (fatty acids)/Protein (amino acids)
1244	PO ₂ ⁻ asymmetric stretching (ν_{as} PO ₂ ⁻)	Nucleic acid(phosphodiester group)/Phospholipid/Phosphorylated protein
	Amide III (ν C–N)	Protein (e.g. collagen)
1045	PO ₂ ⁻ symmetric stretching (ν_{sym} PO ₂ ⁻)	Nucleic acid (RNA/DNA)
	O–H stretching, O–H bending (ν OH, δ OH)	Carbohydrates (glycogen)
	C–O stretching, C–O bending of the C–OH groups (ν C–O, δ C–O)	Carbohydrates (glucose, fructose, glycogen, etc.)
995	C–O ribose/C–C	Nucleic acid (RNA)
	RNA uracil ring stretching; uracil ring bending	Nucleic acid (RNA)

Assignments based on different references: (STUART, 2005; MOVASAGHI *et al.*, 2008; BELLISOLA e SORIO, 2012; ORPHANOU *et al.*, 2015) Abbreviations: ν = stretching vibrations, δ = bending vibrations, sym = symmetric vibrations and as = asymmetric vibrations.

The second derivative spectra were analyzed in details to identify wavenumbers which are important in the differentiation between the three groups of patients (**Figure 8**). The major wavenumbers were found at ~ 2929, 2696, 2659, 2322 (3000 cm⁻¹ - 2200 cm⁻¹ region, Figure 2A), 2059, 1728, 1450, 1404 (2200 cm⁻¹ - 1300 cm⁻¹ region, Figure 2B), 1159, 1120, 1041 and 613 cm⁻¹ (1300 cm⁻¹ - 600 cm⁻¹ region, Figure 2C). Among them, wavenumber around 1041 cm⁻¹ was the only statistically significant. In general, the most wavenumbers presented absorption similar between benign and control.

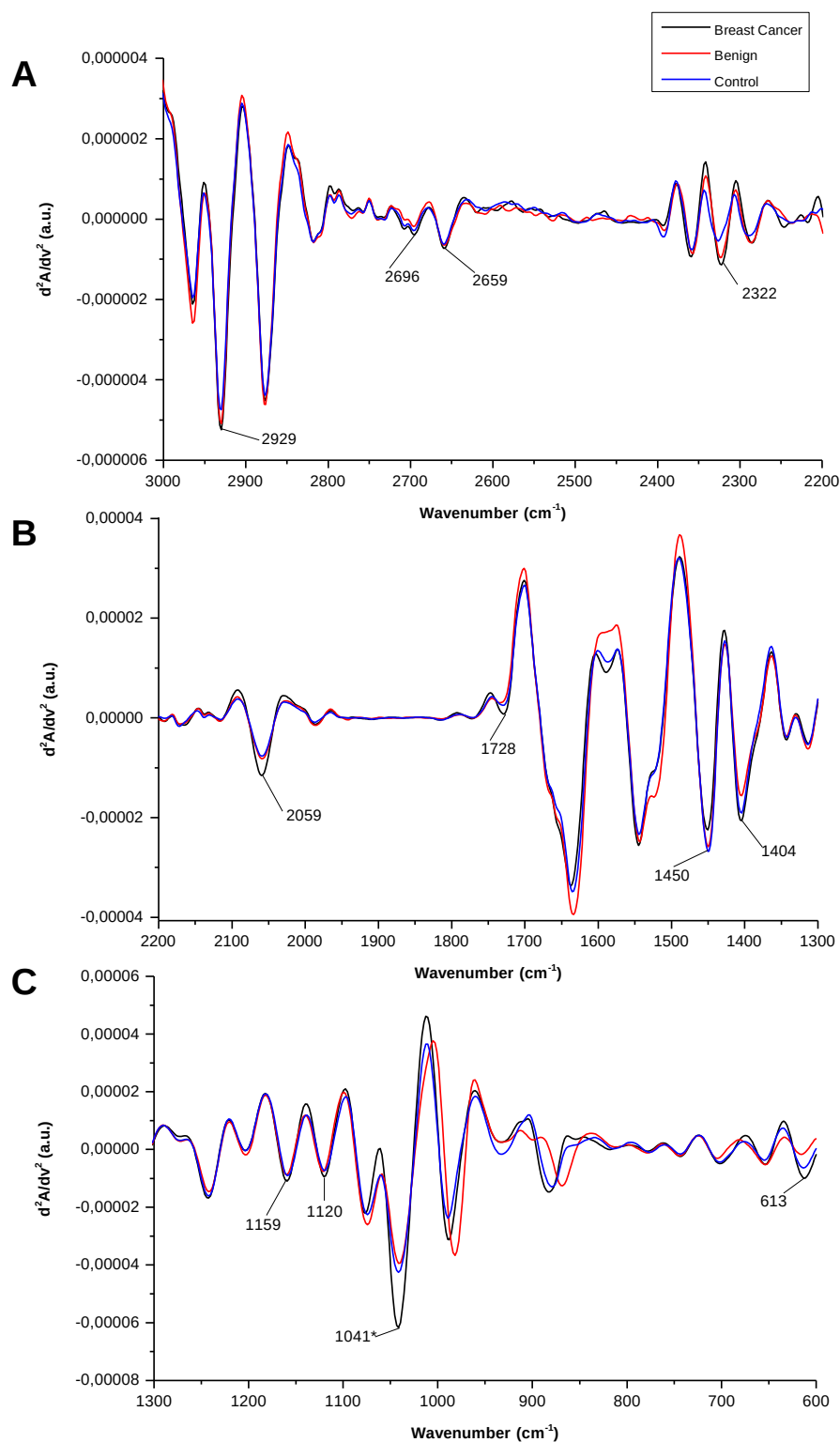


Figure 8 Detailed average second derivative spectra and significant wavenumbers. Average spectra between (A) 3000 cm^{-1} and 2200 cm^{-1} , (B) 2200 cm^{-1} and 1300 cm^{-1} and (C) 1300 cm^{-1} and 600 cm^{-1} for breast cancer (black line), benign breast disease (red line) and control saliva (blue line). The statistically significant wavenumber is indicated by * (* $p < 0.05$, comparison of groups via unpaired t-test with Welch's correction)

The proposed vibrational modes and molecular sources of the wavenumbers indicated on the second derivative spectra (Fig. 8) are presented in detail in **Table 4**. In summary, the absorption bands were associated with vibrational modes of proteins at 1450, 1404, 1159 and 1120 cm^{-1} ; nucleic acids at 2929, 1159, 1120 and 1041 cm^{-1} ; lipids at 2929, 1728, 1450, 1404, 1159 and 1120 cm^{-1} ; and carbohydrates at 2929, 1159 and 1120 cm^{-1} . Particularly the statistically significant peak 1041 cm^{-1} was associated to nucleic acid, due to vibrations of PO_2^- and C–O.

Table 4 Assignments of main FTIR peaks of the average second derivative spectra

2 nd derivative peak (cm^{-1})	Proposed vibrational mode	Molecular source
2929	C-H stretching ($\nu_{\text{C-H}}$)	Carbohydrates/Lipid
	CH_2 asymmetric stretching (ν_{asCH_2})	Nucleic acid/Lipid
1728	C=O stretching ($\nu_{\text{C=O}}$)	Lipid (fatty acid ester)
1450	CH_2 symmetric bending (δ_{symCH_2})	Lipid
	CH_3 asymmetric bending (δ_{asCH_3})	Protein (methyl groups)
1404	CH_3 symmetric bending (δ_{symCH_3})	Protein (methyl groups)
	COO^- symmetric stretching (ν_{symCOO^-})	Lipid (fatty acids)/Protein (amino acids)
1159	C-O stretching ($\nu_{\text{C-O}}$)	Protein/Carbohydrate
	CO–O–C asymmetric stretching ($\nu_{\text{asCO-O-C}}$)	Lipid
	RNA ribose C–O stretching	Nucleic acid (RNA)
1120	Phosphorylated saccharide residue	Carbohydrate
	Mannose-6-phosphate	Protein (glycoprotein)
	RNA ribose C–O stretching	Nucleic acid (RNA)
	P-O-C symmetric stretching ($\nu_{\text{sP-O-C}}$)	Phospholipid
1041*	PO_2^- symmetric stretching ($\nu_{\text{symPO}_2^-}$)	Nucleic acid (RNA/DNA)
	RNA ribose C–O stretching	Nucleic acid (RNA)

Assignments based on different references (STUART, 2005; 2006; MOVASAGHI *et al.*, 2008; BELLISOLA e SORIO, 2012; ORPHANOU *et al.*, 2015)

Abbreviations: ν = stretching vibrations, δ = bending vibrations, sym = symmetric vibrations and as = asymmetric vibrations.

The comparison of the second derivative absorbance of the statistically significant peak 1041 cm^{-1} between the three study groups is presented in **Figure 9**. Breast cancer patients showed higher absorption (lower value in second derivative) than the benign ($p=0.039$), and no significant difference compared to the control ($p=0.094$). Control and benign groups were similar ($p=0.740$).

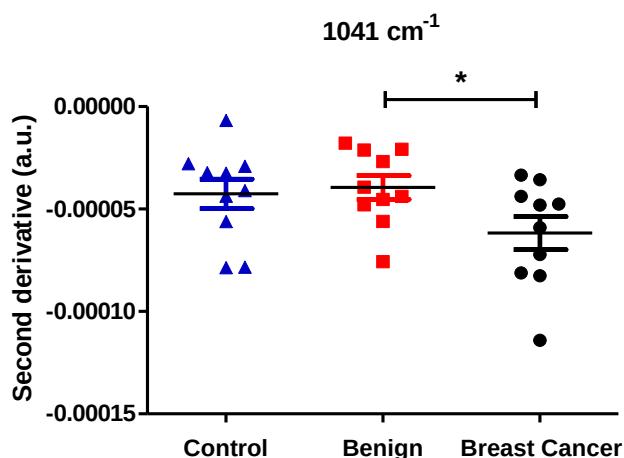
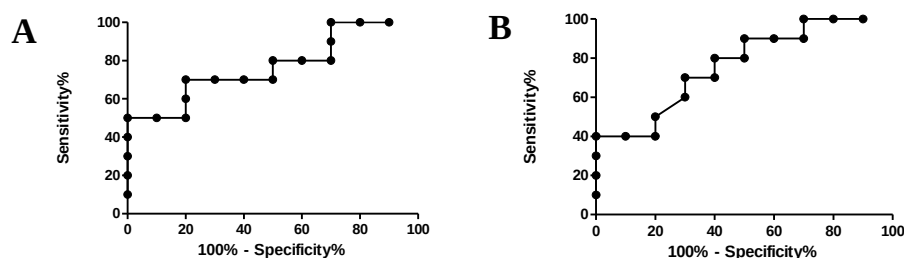


Figure 9 Scatter plot of the statistically significant wavenumber 1041 cm^{-1} for breast cancer (black), benign breast disease (red) and control saliva (blue). The line represents the mean and the error bars (whiskers) represent standard error of the mean (SEM) (* $p<0.05$, comparison of groups via unpaired t-test with Welch's correction).

Since the wavenumber 1041 cm^{-1} shown importance for the discrimination between breast cancer and benign, we evaluated ROC curve and calculated the area under the curve (A.U.C) which is presented in the **Figure 10**. The ROC curve analysis shows a reasonable accuracy of the FTIR tool to discriminate breast cancer from benign and control patients, with an A.U.C of 0.770 for breast cancer vs. control and 0.765 for breast cancer vs. benign. Using the ROC curve, it was possible to select the optimal cut-off that distinguished the groups of patients. This yielded a sensitivity of 70% and a specificity of 80% for breast cancer vs. control and a sensitivity of 80% and a specificity of 60% for breast cancer vs. benign.



	Breast Cancer vs Control	Breast Cancer vs Benign
A.U.C.	0.7700	0.7650
Std. Error	0.1081	0.1066
95% CI	0.5580-0.9820	0.5561-0.9739
P value	0.04130	0.04521
Cut off	$>-4.37510^{-5}$	$>-4.80.10^{-5}$
Sensitivity%	70.00	80.00
(95% CI)	(34.75-93.33%)	(44.39-97.48%)
Specificity%	80.00	60.00
(95% CI)	(44.39-97.48%)	(26.24-87.84%)

Figure 10 ROC curves made from the wavenumber 1041 cm^{-1} . (A) ROC curve for breast cancer vs. control and (B) breast cancer vs. benign breast disease. The table describes the ROC curves results about area under the curve, standard error, 95% confidence interval, p-value, cut off, sensitivity and specificity. Statistically significant differences are represented by * (* $p < 0.05$)

In addition to absorbance analysis of the wavenumbers along the spectrum, the absorbance of the areas of some bands of saliva components was evaluated. Among the evaluated areas, only the interval between 1433 cm^{-1} and 1302.9 cm^{-1} , which has already been associated with α -amylase (KHAUSTOVA *et al.*, 2010), presented statistically significant difference (**Figure 11**). This spectral region also comprises vibrational modes associated with lipids (C-H bending; $\nu_{\text{sym}}\text{COO}^-$ of fatty acids), carbohydrates (coupled C-O and C-C stretching and bending) and proteins (vibrations of amino acids, like O-H bending of serine and CO_2^- symmetric stretching of acid glutamic; $\nu_{\text{sym}}\text{COO}^-$ of amino acids; $\delta_{\text{sym}}\text{CH}_3$ of methyl groups) (STUART, 2005; BELLISOLA e SORIO, 2012). It can be seen that the average in the breast cancer patients is greater than average in the benign ($p=0.0451$) and control ($p=0.0123$) patients. It is important to note that as observed in the other results, the groups of

patients benign and control were similar, did not show significant difference ($p=0.5656$).

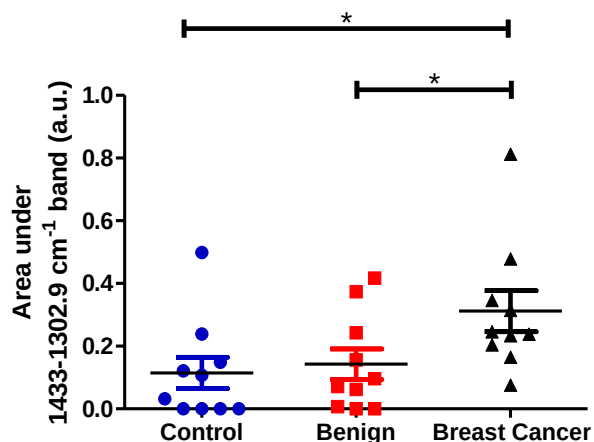
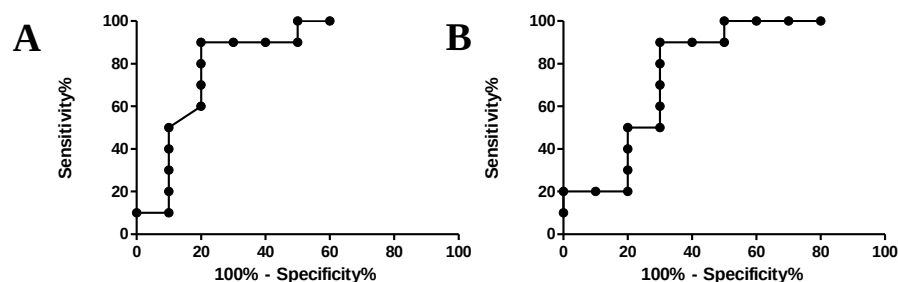


Figure 11 Area under the band between 1433cm^{-1} and 1302.9cm^{-1} of the breast cancer, benign breast disease and control saliva. The line represents the mean and the error bars (whiskers) represent standard error of the mean (SEM) (* $p<0.05$, pairwise comparison of groups via Mann-Whitney test).

Since the band in range 1433cm^{-1} and 1302.9cm^{-1} seems to be important for the discrimination between the patients, we gave special attention to it by evaluating ROC curve and calculating the A.U.C which is presented in the **Figure 12**. The ROC curve analysis shows a good accuracy of the FTIR tool to discriminate between breast cancer and the other groups of patients, with an A.U.C of 0.835 for breast cancer vs. control and 0.770 for breast cancer vs. benign. Using the ROC curve, it was possible to select the optimal cut-off that distinguished the groups of patients. This yielded a sensitivity of 90% and a specificity of 80% for breast cancer vs. control and a sensitivity of 90% and a specificity of 70% for breast cancer vs. benign.



	Breast Cancer vs Control	Breast Cancer vs Benign
A.U.C.	0.8350	0.7700
Std. Error	0.09883	0.1130
95% CI	0.6413-1.029	0.5485-0.9915
P value	0.01136	0.04130
Cut off	> 0.1570	> 0.1605
Sensitivity% (95% CI)	90.00 (55.50-99.75%)	90.00 (55.50-99.75%)
Specificity% (95% CI)	80.00 (44.39-97.48%)	70.00 (34.75-93.33%)

Figure 12 ROC curves made from area under the band between 1433cm^{-1} and 1302.9cm^{-1} . (A) ROC curve for breast cancer vs. control and (B) breast cancer vs. benign breast disease. The table describes the ROC curves results about area under the curve, standard error, 95% confidence interval, p-value, cut off, sensitivity and specificity. Statistically significant differences are represented by * (* $p < 0.05$)

4. 3 FTIR analysis of saliva spectra within the group of breast cancer patients

The analysis of changes in saliva biochemical composition by clinical characteristics within the group of breast cancer patients is shown in **Figure 13 and 14**. We applied a multiple t-test to all wavenumbers of the second derivative spectra comparing the means between patients from two different clinical subgroups. Statistically significant wavenumbers along the spectra are ranked according to p-value where the lowest p-value is indicated by the highest rank of 1.

Figures 13A and 13B show that analysis by histological subtype (invasive carcinoma vs. in situ) had a significant effect on absorption of nucleic acid at 1066-cm^{-1} .

1031 cm^{-1} and 997-989 cm^{-1} , mainly at $\sim 1040 \text{ cm}^{-1}$, due to $\nu_{\text{sym}}\text{PO}_2^-$ and RNA ribose C–O stretching, and at $\sim 990 \text{ cm}^{-1}$ due to C–O ribose/C–C, respectively. As shown in Figures 13C and 13D, significant differences in absorption were located at two main regions between patients with tumor size up to 20 mm (T1) vs. greater than 20 mm and smaller than 50mm (T2): mainly at 1047-1039 cm^{-1} and at 1448-1444 cm^{-1} . The significant peaks were around 1450 cm^{-1} , that corresponds to $\delta_{\text{as}}\text{CH}_3$ of proteins and to $\delta_{\text{sym}}\text{CH}_2$ of lipids, and 1040 cm^{-1} .

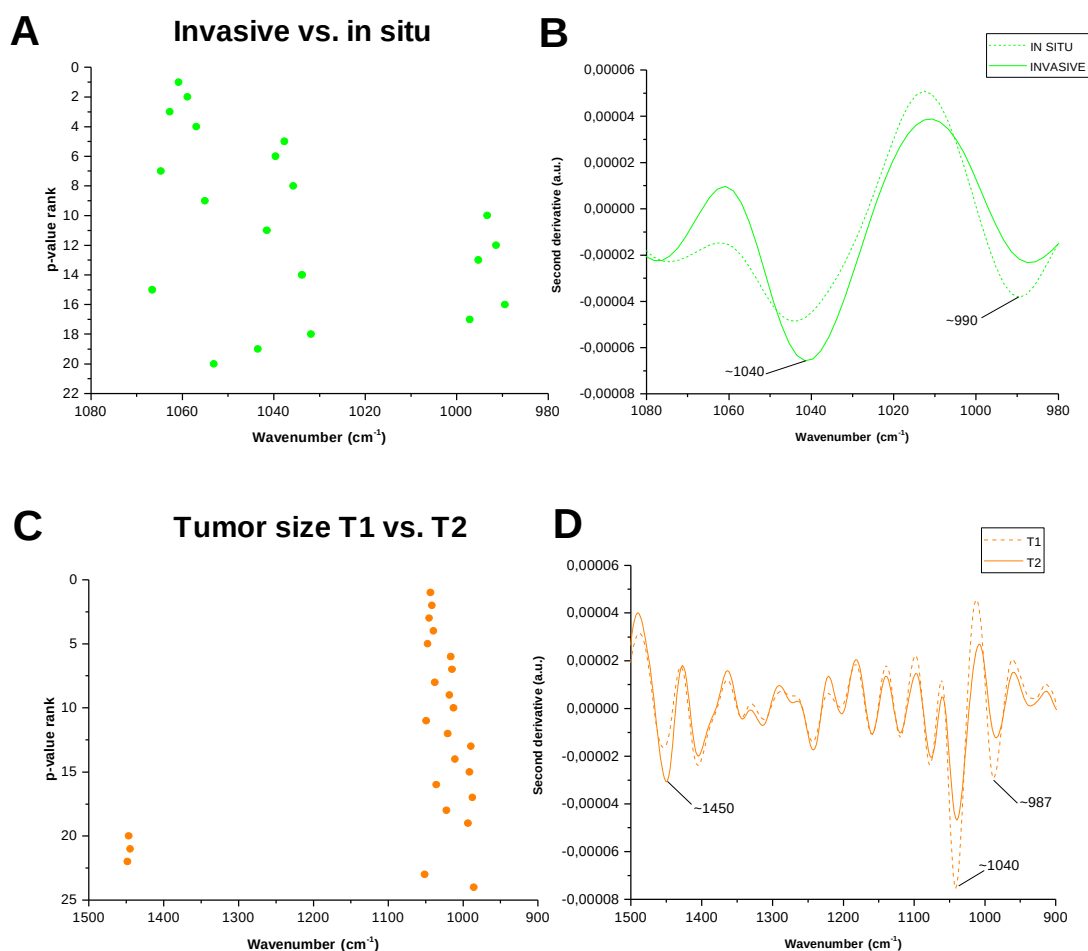


Figure 13 T-test analysis of the second derivative spectra of the group of breast cancer patients and significant wavenumbers. Comparison between the following clinical characteristics: (A,B) Histological subtype, invasive carcinoma vs. in situ and (C,D) Tumor size, T1 (up to 20 mm) vs. T2 (greater than 20 mm and smaller than 50mm). Statistically significant wavenumbers are ranked according to p-value where the lowest p-value is indicated by the highest rank of 1. Comparisons by multiple unpaired t-test with Holm-Šidák correction.

On analysis by cancer staging, stage I or II vs. 0 (Fig.14A and 14B), significant effect on absorption were observed in several wavenumbers, most of the changes in

absorption were located at 1649-1543 cm^{-1} , 1066-985 cm^{-1} and 530-522 cm^{-1} regions. The most significant bands were found at $\sim 1635 \text{ cm}^{-1}$, due to amide I, 1543 cm^{-1} , due to amide II, 1040 cm^{-1} and 990 cm^{-1} . The first two peaks are related to proteins and the last two correspond to nucleic acids. Analysis by molecular subtype, luminal A vs. luminal B, yielded a significant difference in absorption at three regions of the spectra: 1058-1049 cm^{-1} , 1028-987 cm^{-1} and 887-877 cm^{-1} , mainly around 990 cm^{-1} and 882 cm^{-1} that corresponds to C-O deoxyribose/C-C of nucleic acids (Fig.14C and 14D). Lymph node involvement (present vs. absent) and histological grade (G2 vs. G3) were not associated with any significant change in absorption (data not shown).

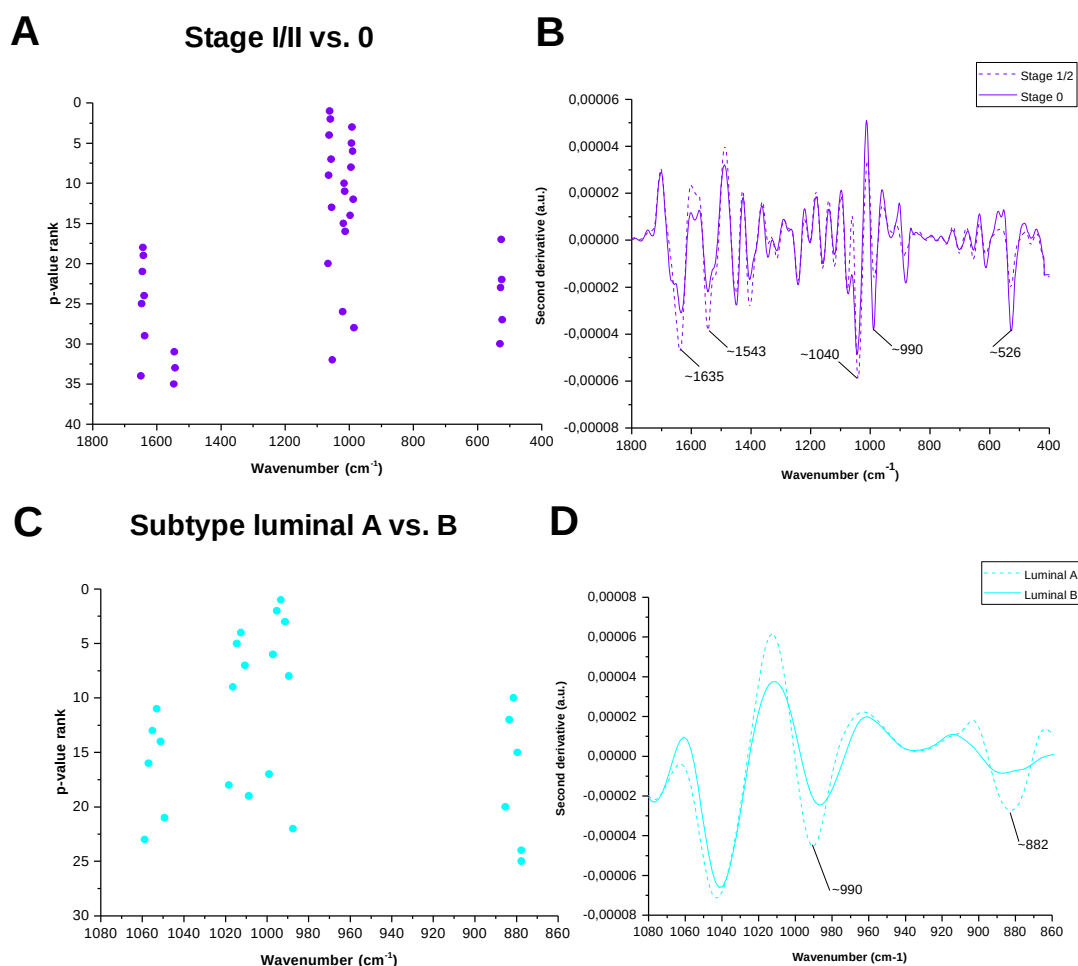


Figure 14 T-test analysis of the second derivative spectra of the group of breast cancer patients and significant wavenumbers.. Comparison between the following clinical characteristics: (A,B) Cancer staging, stage I or II vs. 0 and (C,D) Molecular subtype, luminal A vs. luminal B. Statistically significant wavenumbers are ranked according to p-value where the lowest p-value is indicated by the highest rank of 1. Comparisons by multiple unpaired t-test with Holm-Šídák correction.

Table 5 summarizes the data about analysis within the group of breast cancer patients, as statistically significant wavenumbers, proposed vibrational mode, molecular source and clinical characteristics with higher absorption for each peak. As can be seen, there were difference in absorption of proteins, carbohydrates, lipids and nucleic acids between the clinical parameters analyzed, especially on the fingerprint region that comprises 1500 to 600 cm^{-1} . Some bands like 987 and 882 cm^{-1} weren't shown because they weren't assigned to any vibrational mode of biomolecules.

Table 5 Assignments of statistically significant wavenumbers within the group of breast cancer patients

2nd derivative peak (cm⁻¹)	Proposed vibrational mode	Molecular source	Clinical characteristics with higher absorption at these peaks
1635	β -sheet structure of amide I	Protein	Stage I/II
1543	Amide II (δ N-H, ν C-N)	Protein	Stage I/II
1450	CH ₂ symmetric bending (δ_{sym} CH ₂)	Lipid	Tumor size T2
	CH ₃ asymmetric bending (δ_{as} CH ₃)	Protein (methyl groups)	
1040	PO ₂ ⁻ symmetric stretching (ν_{sym} PO ₂ ⁻)	Nucleic acid (RNA/DNA)	Invasive carcinoma, Tumor size T1,
	RNA ribose C-O stretching	Nucleic acid (RNA)	Stage I/II
990	C-O ribose/C-C	Nucleic acid (RNA)	In situ carcinoma, Stage 0, Subtype luminal A
882	C-O deoxyribose/C-C	Nucleic acid (DNA)	Subtype luminal A

Assignments based on different references: (STUART, 2005; 2006; MOVASAGHI *et al.*, 2008; BELLISOLA e SORIO, 2012; ORPHANOU *et al.*, 2015)

Abbreviations: ν = stretching vibrations, δ = bending vibrations, sym = symmetric vibrations and as = asymmetric vibrations.

5 DISCUSSION

FTIR spectroscopy is a well-established analytical technique with potential for use in routine clinical analyzes of a wide range of sample types, allowing rapid, high-throughput and non-destructive screening (BELLISOLA e SORIO, 2012). As an important diagnostic body fluid, saliva contains many kinds of biomolecules that may be associated with disease transformation and can be very useful as tumor biomarkers for human disease diagnosis, mainly proteins and nucleic acids (FENG *et al.*, 2015). This study is the first to generate FTIR spectra from saliva and derive chemical fingerprints for the purpose of diagnosis and prognosis of breast cancer. It aimed to show the main biochemical differences between saliva from breast cancer, benign breast disease and control patients.

One of the objectives of the study was to determine which wavenumbers were significantly different between saliva of cancer, benign and control patients. Some biochemical components of saliva have highly specific bands in the infrared region, such as proteins (α -amylase; albumin; cystatins; mucins; proline-rich proteins; sIgA); hormones (cortisol, testosterone) and lipids (cholesterol and mono/diglycerides of fatty acids) (CAETANO JÚNIOR *et al.*, 2015). Using ATR-FTIR, Khaustova and colleagues (2009) also found total protein, glucose, secretory immunoglobulin A, urea, amylase, cortisol and inorganic phosphate with precision and reproducibility. In according to this, in the present work we could see along the original spectra many bands of absorption that corresponds to chemical groups related to proteins, lipids, nucleic acids and carbohydrates. Furthermore, many other molecules that have been accurately detected in saliva could probably be associated with chemical groups found in a saliva spectra, such as other hormones steroids (estriol, estrogen, progesterone, aldosterone); antibodies (IgG, IgA, sIgA, IgM); growth factors (EGF, NGF, VEGF, IGF); cytokines and chemokines (IL-1 beta, IL-8, IL-6, MCP-1, CX3CL1, GRO-1 alpha, troponin I, TNF alpha); and nucleic acids (human DNA, microbial DNA, mRNA, siRNA, micro RNA) (MALAMUD, 2011).

In order to deep the analysis, the second-derivative spectra was analyzed. The second-derivative accentuates the bands, resolving broad and overlapped bands into individual, reducing the background interference, and increasing the accuracy of analysis by revealing the genuine biochemical characteristics (LEWIS *et al.*, 2010; ZELIG *et al.*, 2015). The results obtained in this study corroborate this statement, because through derivative analysis it was possible to identify more clearly the bands of absorption.

Among the seven main wavenumbers observed in the second-derivative ATR-FTIR spectra, six of them had tendency of higher absorbance in breast cancer patients relative to the other patients, with a statistically significant difference only for wavenumber around 1041 cm^{-1} between breast cancer and benign patients. The ROC curve analysis of these wavenumber shown a reasonable accuracy with an A.U.C of 0.770, a sensitivity of 70% and a specificity of 80% for breast cancer vs. control, and A.U.C of 0.7650, sensitivity of 80% and specificity of 60% for breast cancer vs. benign. In addition to absorbance analysis of the wavenumbers along the spectra, the absorbance of the area between 1433 cm^{-1} and 1302.9 cm^{-1} was increased in breast cancer patients compared to benign and control. The ROC curve analysis of this region shown a good accuracy with an A.U.C of 0.835, a sensitivity of 90% and a specificity of 80% for breast cancer vs. control, and A.U.C of 0.770, sensitivity of 90% and specificity of 70% for breast cancer vs. benign. It is important to note that most results between benign and control patients was similar, this is in concordance with literature that have been report none or only small differences between them (ZELIG *et al.*, 2015)

These satisfactory results of sensitivity and specificity describe ATR-FTIR as a promising tool to distinguish patients with breast cancer from control and benign patients. It is known that the conventional breast screening methods have many limitations, such as low sensitivity and specificity. Various levels of sensitivity and specificity for detecting breast cancer have been published, for instance, sensitivity of 67,8%, 83%, 94,4% and specificity of 75%, 34%, 26,4% for mammography, ultrasound and MRI, respectively (WANG, 2017). .

It is known that increase in absorbance at a wavenumber in one sample relative to another can be due to different reasons, including the increase in the frequency of a bond vibration mode of specific biomolecules (LEWIS *et al.*, 2010). It is possible that increase in absorbance levels of breast cancer patients at the wavenumber 1041 cm^{-1} is due to increased levels of nucleic acids, and in the 1433-1302.9 cm^{-1} region is due to increased levels of protein, lipids and carbohydrates. Previous studies on cancer cells and tissues using FTIR spectroscopy also reported many changes in the phosphate region, which corresponds mainly to nucleic acids and carbohydrates, and significant increased ratio of CH_2/CH_3 in the higher region of lipids and protein absorptions (ZELIG *et al.*, 2015).

In this study, a possible explanation for the reason why breast cancer allows higher absorptions of biomolecules is that the saliva harbors a wide spectrum of molecules that originate from systemic sources, since they can pass from blood to saliva mainly by passive diffusion of lipophilic molecules (e.g. steroid hormones), or active transport of proteins via ligand-receptor binding (PFAFFE *et al.*, 2011). So, saliva presents biomarkers that reflect the physiological state of the body, such as, breast cancer. There are numerous putative salivary molecular biomarkers that are probably altered in the presence of breast cancer. Higher levels of some proteins, carbohydrates and nucleic acids have already been found in the saliva of breast cancer patients in comparison to normal controls, which corroborates with the results found in this study. In general, these biomarkers were evaluated by proteomic, immunological and biomolecular techniques.

Higher levels of many proteins were observed in the saliva of breast cancer patients, such as: a) endothelial growth factor (VEGF) and epidermal growth factor (EGF), which are potent angiogenic factors; b) carcinoembryonic antigen (CEA) that is a glycoprotein and well-established serum tumor marker for breast cancer (BROOKS *et al.*, 2008); c) soluble form of c-erbB-2 (HER2) protein, that is a receptor tyrosine kinase, product of c-erbB-2 oncogene and marker of poor prognosis (STRECKFUS, BIGLER, DELLINGER, *et al.*, 2000); d) p53 that is a tumor suppressor protein product of oncogene p53, it regulates target genes that induce cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism, and it is

indicator of poor clinical outcome (STRECKFUS, BIGLER, TUCCI, et al., 2000). Some works also found mRNA biomarkers (ZHANG *et al.*, 2010) and carbohydrate CA15-3, which is a tumor marker found on the surface of cancer cells and sheds into the blood stream, being used to monitor advanced and metastatic cases (STRECKFUS, BIGLER, DELLINGER, et al., 2000; AGHA-HOSSEINI et al., 2009).

Further analysis of the cancer group revealed an influence of several clinical parameters into the infrared spectra. It was found that clinical characteristics associated to a more advanced and aggressive phenotype of tumor, such as stage I or II, larger tumor size, and an invasive carcinoma, showed higher absorption in peaks related to protein and nucleic acid. This result corroborate with previous work in the literature, since some biomarkers found in saliva have been related to specific tumor behaviors and poor prognostic (BANIN HIRATA *et al.*, 2014). The important peaks were found mostly in the fingerprint region of MIR spectra that has the fundamental vibrational modes of key chemical bonds of biomolecules. The correlation of specific spectral changes with clinical parameters of cancer patients indicates for possible contribution of FTIR spectroscopy not only to diagnosis but also prognosis.

6 CONCLUSION

This study reports on the first application of ATR-FTIR spectroscopy to discriminate breast cancer saliva from benign breast disease and control for diagnosis and prognosis purposes. Comparing the three groups of patients, was found higher absorbance levels in breast cancer patients at wavenumber 1041 cm^{-1} , with reasonable accuracy, and in the area of $1433\text{-}1302.9\text{ cm}^{-1}$ region, with good accuracy due to high sensitivity and specificity. These increase in absorbance levels between breast cancer and the other two groups was associated to changes in vibrational modes of nucleic acids, protein, lipids and carbohydrates. It was also found that the changes in absorptions bands within the breast cancer group are dependent of the tumor phenotype determined by clinical parameters, and are related mainly to protein and nucleic acid.

Therefore the FTIR spectroscopy was capable to show biochemical changes of saliva components as result of breast carcinogenesis that cause different vibrational modes in the biomolecules. It is important to note that differently from other methods that search biomarkers in saliva, FTIR detect changes at a multi-molecular level, being a promising tool for early diagnosis and prognosis of breast cancer.

More extensive studies are needed to verify these preliminary results. One perspective is to increase the sample size used, that is possibly a limiting factor, with samples already collected and collecting additional material. Another perspective to validate our method is to perform the Multivariate Data Analysis, specifically the Principal Component Analysis (PCA), which will allow to identify with more precision the modifications between the study groups, evaluating the ability of the method in detecting and discriminating patients, and thus clustering them in the study groups.

REFERENCES

- ABRAO, A. L. et al. Salivary proteomics: A new adjuvant approach to the early diagnosis of familial juvenile systemic lupus erythematosus. **Med Hypotheses**, v. 89, p. 97-100, Apr 2016. <https://doi.org/10.1016/j.mehy.2016.02.010>
- AGHA-HOSSEINI, F.; MIRZAI-DIZGAH, I.; RAHIMI, A. Correlation of serum and salivary CA15-3 levels in patients with breast cancer. **Med Oral Patol Oral Cir Bucal**, v. 14, n. 10, p. e521-524, Oct 01 2009. <https://doi.org/10.4317/medoral.14.e521>
- ANDREW CHAN, K. L.; KAZARIAN, S. G. Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells. **Chem Soc Rev**, v. 45, n. 7, p. 1850-1864, Apr 07 2016. <https://doi.org/10.1039/C5CS00515A>
- BACKHAUS, J. et al. Diagnosis of breast cancer with infrared spectroscopy from serum samples. **Vibrational Spectroscopy**, v. 52, n. 2, p. 173-177, 2010/03/18/ 2010.
- BAKER, M. J. et al. Using Fourier transform IR spectroscopy to analyze biological materials. **Nat Protoc**, v. 9, n. 8, p. 1771-1791, Aug 2014. <https://doi.org/10.1038/nprot.2014.110>
- BANIN HIRATA, B. K. et al. Molecular markers for breast cancer: prediction on tumor behavior. **Dis Markers**, v. 2014, p. 513158, 2014. <https://doi.org/10.1155/2014/513158>
- BARTH, A. Infrared spectroscopy of proteins. **Biochim Biophys Acta**, v. 1767, n. 9, p. 1073-1101, Sep 2007. <https://doi.org/10.1016/j.bbabbio.2007.06.004>
- BEEKES, M.; LASCH, P.; NAUMANN, D. Analytical applications of Fourier transform-infrared (FT-IR) spectroscopy in microbiology and prion research. **Vet Microbiol**, v. 123, n. 4, p. 305-319, Aug 31 2007. <https://doi.org/10.1016/j.vetmic.2007.04.010>
- BELLISOLA, G.; SORIO, C. Infrared spectroscopy and microscopy in cancer research and diagnosis. **Am J Cancer Res**, v. 2, n. 1, p. 1-21, 2012.
- BENARD, A. et al. Infrared imaging in breast cancer: automated tissue component recognition and spectral characterization of breast cancer cells as well as the tumor microenvironment. **Analyst**, v. 139, n. 5, p. 1044-1056, Mar 07 2014. <https://doi.org/10.1039/c3an01454a>
- BIGLER, L. R.; STRECKFUS, C. F.; DUBINSKY, W. P. Salivary biomarkers for the detection of malignant tumors that are remote from the oral cavity. **Clin Lab Med**, v. 29, n. 1, p. 71-85, Mar 2009. <https://doi.org/10.1016/j.cll.2009.01.004>

BROOKS, M. N. et al. Salivary protein factors are elevated in breast cancer patients. **Mol Med Rep**, v. 1, n. 3, p. 375-378, May-Jun 2008. <https://doi.org/10.3892/mmr.1.3.375>

BUNACIU, A. A.; HOANG, V. D.; ABOUL-ENEIN, H. Y. Applications of FT-IR Spectrophotometry in Cancer Diagnostics. **Critical Reviews in Analytical Chemistry**, v. 45, n. 2, p. 156-165, 2015/04/03 2015.

CAETANO JÚNIOR, P. C.; STRIXINO, J. F.; RANIERO, L. Analysis of saliva by Fourier transform infrared spectroscopy for diagnosis of physiological stress in athletes. **Research on Biomedical Engineering**, v. 31, p. 116-124, 2015. <https://doi.org/10.1590/2446-4740.0664>

CHEN, S. H. et al. [Fourier transform infrared spectroscopy in detection of breast cancer]. **Zhejiang Da Xue Xue Bao Yi Xue Ban**, v. 43, n. 4, p. 494-500, Jul 2014.

CHENG, F. et al. Investigation of salivary free amino acid profile for early diagnosis of breast cancer with ultra performance liquid chromatography-mass spectrometry. **Clin Chim Acta**, v. 447, p. 23-31, Jul 20 2015. <https://doi.org/10.1016/j.cca.2015.05.008>

CLEMENS, G. et al. Vibrational spectroscopic methods for cytology and cellular research. **Analyst**, v. 139, n. 18, p. 4411-4444, Sep 21 2014. <https://doi.org/10.1039/C4AN00636D>

COLAGAR, A. H.; CHAICHI, M. J.; KHADJVAND, T. Fourier transform infrared microspectroscopy as a diagnostic tool for distinguishing between normal and malignant human gastric tissue. **J Biosci**, v. 36, n. 4, p. 669-677, Sep 2011. <https://doi.org/10.1007/s12038-011-9090-5>

DEPCIUCH, J. et al. Monitoring breast cancer treatment using a Fourier transform infrared spectroscopy-based computational model. **J Pharm Biomed Anal**, v. 143, p. 261-268, Jun 08 2017. <https://doi.org/10.1016/j.jpba.2017.04.039>

DEPCIUCH, J. et al. Application of Raman Spectroscopy and Infrared Spectroscopy in the Identification of Breast Cancer. **Appl Spectrosc**, v. 70, n. 2, p. 251-263, Feb 2016. <https://doi.org/10.1177/0003702815620127>

DEPCIUCH, J. et al. Phospholipid-protein balance in affective disorders: Analysis of human blood serum using Raman and FTIR spectroscopy. A pilot study. **J Pharm Biomed Anal**, v. 131, p. 287-296, Nov 30 2016. <https://doi.org/10.1016/j.jpba.2016.08.037>

DONG, L. et al. Evaluation of FTIR spectroscopy as diagnostic tool for colorectal cancer using spectral analysis. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 122, p. 288-294, 2014/03/25/ 2014.

EIKJE, N. S.; AIZAWA, K.; OZAKI, Y. Vibrational spectroscopy for molecular characterisation and diagnosis of benign, premalignant and malignant skin tumours. **Biotechnol Annu Rev**, v. 11, p. 191-225, 2005. [https://doi.org/10.1016/S1387-2656\(05\)11006-0](https://doi.org/10.1016/S1387-2656(05)11006-0)

EL-TAWIL, S. G. et al. Comparative study between Pap smear cytology and FTIR spectroscopy: a new tool for screening for cervical cancer. **Pathology**, v. 40, n. 6, p. 600-603, Oct 2008. <https://doi.org/10.1080/00313020802320622>

ELKINS, K. M. Rapid presumptive "fingerprinting" of body fluids and materials by ATR FT-IR spectroscopy. **J Forensic Sci**, v. 56, n. 6, p. 1580-1587, Nov 2011. <https://doi.org/10.1111/j.1556-4029.2011.01870.x>

ELLIS, D. I.; GOODACRE, R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. **Analyst**, v. 131, n. 8, p. 875-885, Aug 2006. <https://doi.org/10.1039/b602376m>

ENKE, C. G.; NIEMAN, T. A. Signal-to-Noise Ratio Enhancement by Least-Squares Polynomial Smoothing. **Analytical Chemistry**, v. 48, n. 8, p. 705A-712A, 1976/07/01 1976.

FABIAN, H. et al. Diagnosing benign and malignant lesions in breast tissue sections by using IR-microspectroscopy. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 1758, n. 7, p. 874-882, 2006/07/01/ 2006.

FENG, S. et al. Surface-enhanced Raman spectroscopy of saliva proteins for the noninvasive differentiation of benign and malignant breast tumors. **Int J Nanomedicine**, v. 10, p. 537-547, 2015. <https://doi.org/10.2147/IJN.S71811>

GAJJAR, K. et al. Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. **Analyst**, v. 138, n. 14, p. 3917-3926, Jul 21 2013. <https://doi.org/10.1039/c3an36654e>

GAVGIOTAKI, E. et al. Distinction between breast cancer cell subtypes using third harmonic generation microscopy. **Journal of Biophotonics**, p. n/a-n/a, 2016.

GHERSEVICH, S.; CEBALLOS, M. P. Mammaglobin A: review and clinical utility. **Adv Clin Chem**, v. 64, p. 241-268, 2014. <https://doi.org/10.1016/B978-0-12-800263-6.00006-9>

GIULIANO, A. E. et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. **CA Cancer J Clin**, Mar 14 2017.

HANAHAN, D. Rethinking the war on cancer. **Lancet**, v. 383, n. 9916, p. 558-563, Feb 8 2014.
[https://doi.org/10.1016/S0140-6736\(13\)62226-6](https://doi.org/10.1016/S0140-6736(13)62226-6)

HANAHAN, D.; WEINBERG, R. A. Hallmarks of cancer: the next generation. **Cell**, v. 144, n. 5, p. 646-674, Mar 4 2011. <https://doi.org/10.1016/j.cell.2011.02.013>

HASSIOTOU, F.; GEDDES, D. Anatomy of the human mammary gland: Current status of knowledge. **Clin Anat**, v. 26, n. 1, p. 29-48, Jan 2013. <https://doi.org/10.1002/ca.22165>

HUGHES, C. et al. Assessing the challenges of Fourier transform infrared spectroscopic analysis of blood serum. **J Biophotonics**, v. 7, n. 3-4, p. 180-188, Apr 2014.
<https://doi.org/10.1002/jbio.201300167>

HUGHES, C. et al. Introducing Discrete Frequency Infrared Technology for High-Throughput Biofluid Screening. **Sci Rep**, v. 6, p. 20173, Feb 04 2016. <https://doi.org/10.1038/srep20173>

INCA. **Estimativa 2016: Incidência de câncer no Brasil** SAÚDE, M. D. Brasil 2016.

KALIA, M. Biomarkers for personalized oncology: recent advances and future challenges. **Metabolism**, v. 64, n. 3 Suppl 1, p. S16-21, Mar 2015. <https://doi.org/10.1016/j.metabol.2014.10.027>

KEEN, J. C.; DAVIDSON, N. E. The biology of breast carcinoma. **Cancer**, v. 97, n. 3 Suppl, p. 825-833, Feb 1 2003. <https://doi.org/10.1002/cncr.11126>

KENDALL, C. et al. Vibrational spectroscopy: a clinical tool for cancer diagnostics. **Analyst**, v. 134, n. 6, p. 1029-1045, Jun 2009. <https://doi.org/10.1039/b822130h>

KHANMOHAMMADI, M. et al. Chemometrics assisted investigation of variations in infrared spectra of blood samples obtained from women with breast cancer: a new approach for cancer diagnosis. **Eur J Cancer Care (Engl)**, v. 19, n. 3, p. 352-359, May 2010. <https://doi.org/10.1111/j.1365-2354.2008.01062.x>

KHAUSTOVA, S. et al. Noninvasive biochemical monitoring of physiological stress by Fourier transform infrared saliva spectroscopy. **Analyst**, v. 135, n. 12, p. 3183-3192, Dec 2010.
<https://doi.org/10.1039/c0an00529k>

KHAUSTOVA, S. A. et al. Assessment of biochemical characteristics of the saliva using Fourier transform mid-infrared spectroscopy. **Bull Exp Biol Med**, v. 148, n. 5, p. 841-844, Nov 2009.
<https://doi.org/10.1007/s10517-010-0831-5>

KOH, E. H. et al. Upregulation of human mammaglobin reduces migration and invasion of breast cancer cells. **Cancer Invest**, v. 32, n. 1, p. 22-29, Jan 2014.
<https://doi.org/10.3109/07357907.2013.861473>

KRAFFT, C. et al. Identification of primary tumors of brain metastases by infrared spectroscopic imaging and linear discriminant analysis. **Technol Cancer Res Treat**, v. 5, n. 3, p. 291-298, Jun 2006.
<https://doi.org/10.1177/153303460600500311>

KRAFFT, C. et al. Disease recognition by infrared and Raman spectroscopy. **J Biophotonics**, v. 2, n. 1-2, p. 13-28, Feb 2009. <https://doi.org/10.1002/jbio.200810024>

KRISHNA, C. M. et al. Combined Fourier transform infrared and Raman spectroscopic approach for identification of multidrug resistance phenotype in cancer cell lines. **Biopolymers**, v. 82, n. 5, p. 462-470, 2006. <https://doi.org/10.1002/bip.20485>

KUMAR, S. et al. Change in the microenvironment of breast cancer studied by FTIR imaging. **Analyst**, v. 138, n. 14, p. 4058-4065, Jul 21 2013. <https://doi.org/10.1039/c3an00241a>

LANE, R.; SEE, S. S. Attenuated total reflectance Fourier transform infrared spectroscopy method to differentiate between normal and cancerous breast cells. **J Nanosci Nanotechnol**, v. 12, n. 9, p. 7395-7400, Sep 2012. <https://doi.org/10.1166/jnn.2012.6582>

LEWIS, P. D. et al. Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum. **BMC Cancer**, v. 10, n. 1, p. 640, 2010. <https://doi.org/10.1186/1471-2407-10-640>

LI, J. et al. Clinicopathological classification and traditional prognostic indicators of breast cancer. **Int J Clin Exp Pathol**, v. 8, n. 7, p. 8500-8505, 2015.

LIMA, C. A. et al. ATR-FTIR spectroscopy for the assessment of biochemical changes in skin due to cutaneous squamous cell carcinoma. **Int J Mol Sci**, v. 16, n. 4, p. 6621-6630, Mar 24 2015.
<https://doi.org/10.3390/ijms16046621>

LIMA, K. M. et al. Segregation of ovarian cancer stage exploiting spectral biomarkers derived from blood plasma or serum analysis: ATR-FTIR spectroscopy coupled with variable selection methods. **Biotechnol Prog**, v. 31, n. 3, p. 832-839, May-Jun 2015. <https://doi.org/10.1002/btpr.2084>

MACKANOS, M. A.; CONTAG, C. H. Fiber-optic probes enable cancer detection with FTIR spectroscopy. **Trends Biotechnol**, v. 28, n. 6, p. 317-323, Jun 2010. <https://doi.org/10.1016/j.tibtech.2010.04.001>

MALAMUD, D. Saliva as a diagnostic fluid. **Dent Clin North Am**, v. 55, n. 1, p. 159-178, Jan 2011.
<https://doi.org/10.1016/j.cden.2010.08.004>

MALINS, D. C. et al. The etiology and prediction of breast cancer. Fourier transform-infrared spectroscopy reveals progressive alterations in breast DNA leading to a cancer-like phenotype in a high proportion of normal women. **Cancer**, v. 75, n. 2, p. 503-517, Jan 15 1995.

[https://doi.org/10.1002/1097-0142\(19950115\)75:2<503::AID-CNCR2820750213>3.0.CO;2-0](https://doi.org/10.1002/1097-0142(19950115)75:2<503::AID-CNCR2820750213>3.0.CO;2-0)

MAZIAK, D. E. et al. Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: An exploratory study. **Cancer Detection and Prevention**, v. 31, n. 3, p. 244-253, 2007/01/01/ 2007.

MEHROTRA, R. et al. Infrared spectroscopic analysis of tumor pathology. **Indian J Exp Biol**, v. 45, n. 1, p. 71-76, Jan 2007.

MENZIES, G. E. et al. Fourier transform infrared for noninvasive optical diagnosis of oral, oropharyngeal, and laryngeal cancer. **Transl Res**, v. 163, n. 1, p. 19-26, Jan 2014.

<https://doi.org/10.1016/j.trsl.2013.09.006>

MIGNOLET, A.; GOORMAGHTIGH, E. High throughput absorbance spectra of cancerous cells: a microscopic investigation of spectral artifacts. **Analyst**, v. 140, n. 7, p. 2393-2401, Apr 07 2015.

<https://doi.org/10.1039/C4AN01834F>

MOVASAGHI, Z.; REHMAN, S.; UR REHMAN, D. I. Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues. **Applied Spectroscopy Reviews**, v. 43, n. 2, p. 134-179, 2008/02/01 2008.

NALLALA, J. et al. Infrared spectral imaging as a novel approach for histopathological recognition in colon cancer diagnosis. **J Biomed Opt**, v. 17, n. 11, p. 116013, Nov 2012.

<https://doi.org/10.1117/1.JBO.17.11.116013>

NOREEN, R. et al. FTIR spectro-imaging of collagen scaffold formation during glioma tumor development. **Anal Bioanal Chem**, v. 405, n. 27, p. 8729-8736, Nov 2013.

<https://doi.org/10.1007/s00216-013-7337-8>

NOREEN, R. et al. Detection of collagens in brain tumors based on FTIR imaging and chemometrics.

Anal Bioanal Chem, v. 401, n. 3, p. 845-852, Aug 2011. <https://doi.org/10.1007/s00216-011-4899-1>

NOREEN, R. et al. FTIR spectro-imaging of collagens for characterization and grading of gliomas. **Biotechnol Adv**, v. 30, n. 6, p. 1432-1446, Nov-Dec 2012.

<https://doi.org/10.1016/j.biotechadv.2012.03.009>

ORPHANO, C. M. et al. The detection and discrimination of human body fluids using ATR FT-IR spectroscopy. **Forensic Sci Int**, v. 252, p. e10-16, Jul 2015.

<https://doi.org/10.1016/j.forsciint.2015.04.020>

PANDYA, S.; MOORE, R. G. Breast development and anatomy. **Clin Obstet Gynecol**, v. 54, n. 1, p. 91-95, Mar 2011. <https://doi.org/10.1097/GRF.0b013e318207ffe9>

PEAIRS, K. S. et al. Screening for breast cancer. **Semin Oncol**, v. 44, n. 1, p. 60-72, Feb 2017. <https://doi.org/10.1053/j.seminoncol.2017.02.004>

PEZZEI, C. et al. Characterization of normal and malignant prostate tissue by Fourier transform infrared microspectroscopy. **Mol Biosyst**, v. 6, n. 11, p. 2287-2295, Nov 2010. <https://doi.org/10.1039/c0mb00041h>

PFATTE, T. et al. Diagnostic potential of saliva: current state and future applications. **Clin Chem**, v. 57, n. 5, p. 675-687, May 2011. <https://doi.org/10.1373/clinchem.2010.153767>

PINK, R. et al. Saliva as a diagnostic medium. **Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub**, v. 153, n. 2, p. 103-110, Jun 2009. <https://doi.org/10.5507/bp.2009.017>

REHMAN, S. et al. Fourier Transform Infrared Spectroscopic Analysis of Breast Cancer Tissues; Identifying Differences between Normal Breast, Invasive Ductal Carcinoma, and Ductal Carcinoma In Situ of the Breast. **Applied Spectroscopy Reviews**, v. 45, n. 5, p. 355-368, 2010/09/17 2010.

RIEPO, L. et al. Application of second derivative spectroscopy for increasing molecular specificity of Fourier transform infrared spectroscopic imaging of articular cartilage. **Osteoarthritis Cartilage**, v. 20, n. 5, p. 451-459, May 2012. <https://doi.org/10.1016/j.joca.2012.01.010>

SCHULTZ, C. P. The potential role of Fourier transform infrared spectroscopy and imaging in cancer diagnosis incorporating complex mathematical methods. **Technol Cancer Res Treat**, v. 1, n. 2, p. 95-104, Apr 2002. <https://doi.org/10.1177/153303460200100201>

SHARMA, G. et al. Various types and management of breast cancer: An overview. **Journal of Advanced Pharmaceutical Technology & Research**, v. 1, n. 2, p. 109-126, April 1, 2010 2010.

SHENG, D. et al. Distinction of leukemia patients' and healthy persons' serum using FTIR spectroscopy. **Spectrochim Acta A Mol Biomol Spectrosc**, v. 101, p. 228-232, Jan 15 2013. <https://doi.org/10.1016/j.saa.2012.09.072>

SHENG, D. et al. Comparison of serum from gastric cancer patients and from healthy persons using FTIR spectroscopy. **Spectrochim Acta A Mol Biomol Spectrosc**, v. 116, p. 365-369, Dec 2013. <https://doi.org/10.1016/j.saa.2013.07.055>

SIMONOVA, D.; KARAMANCHEVA, I. Application of Fourier Transform Infrared Spectroscopy for Tumor Diagnosis. **Biotechnology & Biotechnological Equipment**, v. 27, n. 6, p. 4200-4207, 2013/01/01 2013.

STEWART, B. W., WILD, C. P. **World Cancer Report 2014**. PUBLICATION-WHO, I. N.: 630 p. 2014.

STRECKFUS, C. et al. The presence of soluble c-erbB-2 in saliva and serum among women with breast carcinoma: a preliminary study. **Clin Cancer Res**, v. 6, n. 6, p. 2363-2370, Jun 2000.

STRECKFUS, C. et al. A preliminary study of CA15-3, c-erbB-2, epidermal growth factor receptor, cathepsin-D, and p53 in saliva among women with breast carcinoma. **Cancer Invest**, v. 18, n. 2, p. 101-109, 2000. <https://doi.org/10.3109/07357900009038240>

STUART, B. H. Biological Applications. In: (Ed.). **Infrared Spectroscopy: Fundamentals and Applications**: John Wiley & Sons, Ltd, 2005. p.137-165. ISBN 9780470011140.

_____. Infrared Spectroscopy of Biological Applications: An Overview. In: (Ed.). **Encyclopedia of Analytical Chemistry**: John Wiley & Sons, Ltd, 2006. ISBN 9780470027318.

SUN, X. et al. Detection of lung cancer tissue by attenuated total reflection-Fourier transform infrared spectroscopy-a pilot study of 60 samples. **J Surg Res**, v. 179, n. 1, p. 33-38, Jan 2013. <https://doi.org/10.1016/j.jss.2012.08.057>

TIAN, P. et al. Intraoperative detection of sentinel lymph node metastases in breast carcinoma by Fourier transform infrared spectroscopy. **Br J Surg**, v. 102, n. 11, p. 1372-1379, Oct 2015. <https://doi.org/10.1002/bjs.9882>

TOSS, A.; CRISTOFANILLI, M. Molecular characterization and targeted therapeutic approaches in breast cancer. **Breast Cancer Res**, v. 17, p. 60, Apr 23 2015. <https://doi.org/10.1186/s13058-015-0560-9>

TREVISAN, J. et al. Measuring similarity and improving stability in biomarker identification methods applied to Fourier-transform infrared (FTIR) spectroscopy. **J Biophotonics**, v. 7, n. 3-4, p. 254-265, Apr 2014. <https://doi.org/10.1002/jbio.201300190>

TRIA TIRONA, M. Breast cancer screening update. **Am Fam Physician**, v. 87, n. 4, p. 274-278, Feb 15 2013.

VERDONCK, M. et al. Characterization of human breast cancer tissues by infrared imaging. **Analyst**, v. 141, n. 2, p. 606-619, Jan 21 2016. <https://doi.org/10.1039/C5AN01512J>

VUONG, D. et al. Molecular classification of breast cancer. **Virchows Arch**, v. 465, n. 1, p. 1-14, Jul 2014. <https://doi.org/10.1007/s00428-014-1593-7>

WANG, A. et al. Oral Biofluid Biomarker Research: Current Status and Emerging Frontiers. **Diagnostics (Basel)**, v. 6, n. 4, Dec 17 2016.

WANG, J. S. et al. FT-IR spectroscopic analysis of normal and cancerous tissues of esophagus. **World J Gastroenterol**, v. 9, n. 9, p. 1897-1899, Sep 2003. <https://doi.org/10.3748/wjg.v9.i9.1897>

WANG, L. Early Diagnosis of Breast Cancer. **Sensors (Basel)**, v. 17, n. 7, Jul 05 2017. <https://doi.org/10.3390/s17071572>

WONG, E. Breast anatomy and histology. 2012. Disponível em: < www.pathophys.org/breast-cancer/breastnormal-copy/ >. Acesso em: 21/06.

WU, B. B. et al. Fourier transform infrared spectroscopy for the distinction of MCF-7 cells treated with different concentrations of 5-fluorouracil. **J Transl Med**, v. 13, p. 108, Apr 02 2015. <https://doi.org/10.1186/s12967-015-0468-2>

YANO, K. et al. Direct measurement of human lung cancerous and noncancerous tissues by fourier transform infrared microscopy: can an infrared microscope be used as a clinical tool? **Anal Biochem**, v. 287, n. 2, p. 218-225, Dec 15 2000. <https://doi.org/10.1006/abio.2000.4872>

ZAWLIK, I. et al. FPA-FTIR Microspectroscopy for Monitoring Chemotherapy Efficacy in Triple-Negative Breast Cancer. **Sci Rep**, v. 6, p. 37333, Nov 18 2016. <https://doi.org/10.1038/srep37333>

ZELIG, U. et al. Early detection of breast cancer using total biochemical analysis of peripheral blood components: a preliminary study. **BMC Cancer**, v. 15, p. 408, 2015. <https://doi.org/10.1186/s12885-015-1414-7>

ZHANG, C. Z. et al. Saliva in the diagnosis of diseases. **Int J Oral Sci**, v. 8, n. 3, p. 133-137, Sep 29 2016. <https://doi.org/10.1038/ijos.2016.38>

ZHANG, L. et al. Discovery and preclinical validation of salivary transcriptomic and proteomic biomarkers for the non-invasive detection of breast cancer. **PLoS One**, v. 5, n. 12, p. e15573, Dec 31 2010. <https://doi.org/10.1371/journal.pone.0015573>

ZHANG, X. et al. Profiling serologic biomarkers in cirrhotic patients via high-throughput Fourier transform infrared spectroscopy: toward a new diagnostic tool of hepatocellular carcinoma. **Transl Res**, v. 162, n. 5, p. 279-286, Nov 2013. <https://doi.org/10.1016/j.trsl.2013.07.007>

ZHANG, X. et al. Intraoperative detection of thyroid carcinoma by fourier transform infrared spectrometry. **J Surg Res**, v. 171, n. 2, p. 650-656, Dec 2011.
<https://doi.org/10.1016/j.jss.2010.05.031>

ATTACHMENTS

Attachment A – American Joint Committee on Cancer Definition of Primary Tumor (T)—Clinical (cT) and Pathological (pT)

T CATEGORY	T CRITERIA
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Tis (DCIS) a	Ductal carcinoma in situ (DCIS)
Tis (Paget)	Paget disease of the nipple NOT associated with invasive carcinoma and/or carcinoma in situ (DCIS) in the underlying breast parenchyma. Carcinomas in the breast parenchyma associated with Paget disease are categorized based on the size and characteristics of the parenchymal disease, although the presence of Paget disease should still be noted.
T1	Tumor ≤ 20 mm in greatest dimension
T1mi	Tumor ≤ 1 mm in greatest dimension
T1a	Tumor > 1 mm but ≤ 5 mm in greatest dimension (round any measurement from >1.0-1.9 mm to 2 mm)
T1b	Tumor > 5 mm but ≤ 10 mm in greatest dimension
T1c	Tumor > 10 mm but ≤ 20 mm in greatest dimension
T2	Tumor > 20 mm but ≤ 50 mm in greatest dimension
T3	Tumor > 50 mm in greatest dimension

T4	Tumor of any size with direct extension to the chest wall and/or to the skin (ulceration or macroscopic nodules); invasion of the dermis alone does not qualify as T4
T4a	Extension to the chest wall; invasion or adherence to pectoralis muscle in the absence of invasion of chest wall structures does not qualify as T4
T4b	Ulceration and/or ipsilateral macroscopic satellite nodules and/or edema (including peau d'orange) of the skin that does not meet the criteria for inflammatory carcinoma
T4c	Both T4a and T4b are present
T4d	Inflammatory carcinoma (see "Rules for Classification")

a Lobular carcinoma in situ is a benign entity and is removed from TNM staging in the American Joint Committee on Cancer (AJCC) Cancer Staging Manual, eighth edition.

Attachment B – American Joint Committee on Cancer Definition of Regional Lymph Nodes—Clinical (cN) and Pathological (pN)

N CATEGORY	N CRITERIA
cN ^a	
cNX ^b	Regional lymph nodes cannot be assessed (eg, previously removed)
cN0	No regional lymph node metastases (by imaging or clinical examination)
cN1	Metastases to movable ipsilateral level I and II axillary lymph node(s)
cN1mi ^c	Micrometastases (approximately 200 cells, larger than 0.2 mm, but none larger than 2.0 mm)
cN2	Metastases in ipsilateral level I and II axillary lymph nodes that are clinically fixed or matted; or in ipsilateral internal mammary lymph nodes in the absence of axillary lymph node metastases
cN2a	Metastases in ipsilateral level I and II axillary lymph nodes fixed to one another (matted) or to other structures
cN2b	Metastases only in ipsilateral internal mammary lymph nodes in the absence of axillary lymph node metastases
cN3	Metastases in ipsilateral infraclavicular (level III axillary) lymph node(s) with or without level I and II axillary lymph node involvement; or in ipsilateral internal mammary lymph node(s) with level I and II axillary lymph node metastases; or metastases in ipsilateral supraclavicular lymph node(s) with or without

N CATEGORY	N CRITERIA
	axillary or internal mammary lymph node involvement
cN3a	Metastases in ipsilateral infraclavicular lymph node(s)
cN3b	Metastases in ipsilateral internal mammary lymph node(s) and axillary lymph node(s)
cN3c	Metastases in ipsilateral supraclavicular lymph node(s)
pN <u>d</u>	
pNX	Regional lymph nodes cannot be assessed (eg, not removed for pathological study or previously removed)
pN0	No regional lymph node metastasis identified or ITCs only
pN0(i+)	ITCs only (malignant cell clusters no larger than 0.2 mm) in regional lymph node(s)
pN0(mol+)	Positive molecular findings by reverse transcriptase-polymerase chain reaction (RT-PCR); no ITCs detected
pN1	Micrometastases; or metastases in 1-3 axillary lymph nodes; and/or clinically negative internal mammary lymph nodes with micrometastases or macrometastases by sentinel lymph node biopsy
pN1mi	Micrometastases (approximately 200 cells, larger than 0.2 mm, but none larger than 2.0 mm)
pN1a	Metastases in 1-3 axillary lymph nodes, at least one metastasis larger than 2.0 mm
pN1b	Metastases in ipsilateral internal mammary sentinel lymph nodes, excluding ITCs

N CATEGORY	N CRITERIA
pN1c	pN1a and pN1b combined
pN2	Metastases in 4-9 axillary lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the absence of axillary lymph node metastases
pN2a	Metastases in 4-9 axillary lymph nodes (at least one tumor deposit larger than 2.0 mm)
pN2b	Metastases in clinically detected internal mammary lymph nodes with or without microscopic confirmation; with pathologically negative axillary lymph nodes
pN3	Metastases in 10 or more axillary lymph nodes; or in infraclavicular (level III axillary) lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the presence of one or more positive level I and II axillary lymph nodes; or in more than 3 axillary lymph nodes and micrometastases or macrometastases by sentinel lymph node biopsy in clinically negative ipsilateral internal mammary lymph nodes; or in ipsilateral supraclavicular lymph nodes
pN3a	Metastases in 10 or more axillary lymph nodes (at least one tumor deposit larger than 2.0 mm); or metastases to the infraclavicular (level III axillary lymph) nodes
pN3b	pN1a or pN2a in the presence of cN2b (positive internal mammary lymph nodes by imaging); or pN2a in the presence of pN1b

N CATEGORY	N CRITERIA
pN3c	Metastases in ipsilateral supraclavicular lymph nodes

Abbreviation: ITCs, isolated tumor cells.

a The (sn) and (f) suffixes should be added to the N category to denote confirmation of metastasis by sentinel lymph node biopsy or fine-needle aspiration/core needle biopsy, respectively.

b The cNX category is used sparingly in patients with regional lymph nodes that were previously surgically removed or if there is no documentation of physical examination of the axilla.

c cN1mi is rarely used but may be appropriate in patients who undergo sentinel lymph node biopsy before tumor resection, which is most likely to occur in patients who receive neoadjuvant therapy.

d The (sn) and (f) suffixes should be added to the N category to denote confirmation of metastasis by sentinel lymph node biopsy or fine-needle aspiration/core needle biopsy, respectively, with NO further resection of lymph nodes.

Attachment C – American Joint Committee on Cancer Definition of Distant Metastasis (M)

CATEGORIES FOR DISTANT METASTASES—CLINICAL AND PATHOLOGICAL (CM0, CM1, PM1)	
M CATEGORY	M CRITERIA
M0	No clinical or radiographic evidence of distant metastases ^a
cM0(i+)	No clinical or radiographic evidence of distant metastases in the presence of tumor cells or and no deposits no greater than 0.2 mm detected microscopically or by using molecular techniques in circulating blood, bone marrow, or other nonregional lymph node tissue in a patient without symptoms or signs of metastases
M1	Distant metastases detected by clinical and radiographic means (cM) and/or histologically proven metastases larger than 0.2 mm (pM)

^a Note that imaging studies are not required to assign the cM0 category.

Attachment D – American Joint Commission on Cancer TNM Anatomic Stage Groups^a

WHEN T IS...	AND N IS...	AND M IS...	THEN THE STAGE GROUP IS... ^b
Tis	N0	M0	0
T1	N0	M0	IA
T0	N1mi	M0	IB
T1	N1mi	M0	IB
T0	N1	M0	IIA
T1	N1	M0	IIA
T2	N0	M0	IIA
T2	N1	M0	IIB
T3	N0	M0	IIB
T1	N2	M0	IIIA
T2	N2	M0	IIIA
T3	N1	M0	IIIA
T3	N2	M0	IIIA
T4	N0	M0	IIIB

WHEN T IS...	AND N IS...	AND M IS...	THEN THE STAGE GROUP IS... ^b
T4	N1	M0	IIIB
T4	N2	M0	IIIB
Any T	N3	M0	IIIC
Any T	Any N	M1	IV

a The Anatomic Stage Group table should only be used in global regions where biomarker tests are not routinely available. Cancer registries in the United States must use the Prognostic Stage Group table for case reporting.

b Notes for Anatomic Stage Grouping:

T1 includes micrometastases (T1mi).

T0 and T1 tumors with lymph node micrometastases only are excluded from stage IIA and are classified as stage IB.

M0 includes M0 with isolated tumor cells (i+).

The designation pM0 is not valid; any M0 is clinical.

If a patient presents with M1 disease before neoadjuvant systemic therapy, then the stage is stage IV and remains stage IV regardless of response to neoadjuvant therapy.

Stage designation may be changed if postsurgical imaging studies reveal the presence of distant metastases, provided the studies are performed within 4 months of diagnosis in the absence of disease progression and provided the patient has not received neoadjuvant therapy.

Staging after neoadjuvant therapy is denoted with a “yc” or “yp” prefix to the T and N classification. No stage group is assigned if there is a complete pathological response (pCR) to neoadjuvant therapy: for example, ypT0ypN0cM0.

Attachment E – Approval of the study by Human Research Ethics Committee



Universidade Federal de Uberlândia
Pró-Reitoria de Pesquisa e Pós-Graduação
COMITÊ DE ÉTICA EM PESQUISA - CEP
Avenida João Neves de Ávila, nº. 2121 – Bloco 1A – Sala 224 - Campus Santa Mônica - Uberlândia-MG –
CEP 38408-144 - FONE/FAX (34) 3239-4131; e-mail: cep@proppa.ufu.br; www.comissao.proppa.ufu.br

ANÁLISE FINAL Nº. 047/13 DO COMITÊ DE ÉTICA EM PESQUISA PARA O ADENDO AO PROTOCOLO REGISTRO
CEP/UFU 064/08

Projeto Pesquisa: "Seleção, caracterização e validação clínica de peptídeos miméticos e ligantes de proteínas associadas a tumores de mama obtidos por phage display".

Pesquisador: Luiz Ricardo Goulart Filho

De acordo com as atribuições definidas na Resolução CNS 196/96, o CEP manifesta-se pela aprovação do adendo ao protocolo de pesquisa proposto.
O adendo ao protocolo não apresenta problemas de ética nas condutas de pesquisa com seres humanos, nos limites da redação e da metodologia apresentadas.

O CEP/UFU lembra que:

- a- segundo a Resolução 196/96, o pesquisador deverá arquivar por 5 anos o relatório da pesquisa e os Termos de Consentimento Livre e Esclarecido, assinados pelo sujeito de pesquisa.
- b- poderá, por escolha aleatória, visitar o pesquisador para conferência do relatório e documentação pertinente ao projeto.
- c- a aprovação do protocolo de pesquisa pelo CEP/UFU dá-se em decorrência do atendimento a Resolução 196/96/CNS, não implicando na qualidade científica do mesmo.

Data de entrega do Relatório Final: *dezembro de 2017*

SITUAÇÃO: ADENDO AO PROTOCOLO APROVADO

ADENDO:

- 1) Alteração do Cronograma de Execução do Projeto.

OBS.: O CEP/UFU LEMBRA QUE QUALQUER MUDANÇA NO PROTOCOLO DEVE SER INFORMADA IMEDIATAMENTE AO CEP PARA FINS DE ANÁLISE E APROVAÇÃO DA MESMA.

Uberlândia, 26 de agosto de 2013.

Prof. Dra. Sandra Terezinha de Farias Furtado
Coordenadora do CEP/UFU

Orientações ao pesquisador

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 196/96 - Item IV.1.f) e deve receber uma cópia do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado (Item IV.2.d).
- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprova (Res. CNS Item III.3.z), aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa (Item V.3) que requeiram ação imediata.
- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo (Res. CNS Item V.4). É papel de o pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.
- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprobatório do CEP, para serem juntadas ao protocolo inicial (Res.251/97, Item III.2.e). O prazo para entrega do relatório é de 120 dias após o término da execução prevista no cronograma do projeto, conforme norma.