Identification of BCC skin cancer using Fourier Transform Infrared Spectroscopy and Machine Learning

Daniella Lúmara Peres[1]*; Sajid Farooq[1]; Rocío Raffaeli[2];
Maria Virginia Croce[2]; Adela Croce[2]; Denise Maria Zezell[1]**.

[1]Nuclear and Energy Research Institute, IPEN-CNEN, São Paulo, Brazil
[2]Universidad Nacional de La Plata, UNLP, La Plata, Argentina

*daniellalumara@usp.br
**zezell@usp.br

INTRODUCTION

Incidence according gender in Brazil - 2022

Non-melanoma skin 31%
Other cancers 27%
Breast 10%
Prostate 10%
Colo rectum 7%
Lung 5%
Stomach 3%
Cervix 3%
LNR 2%
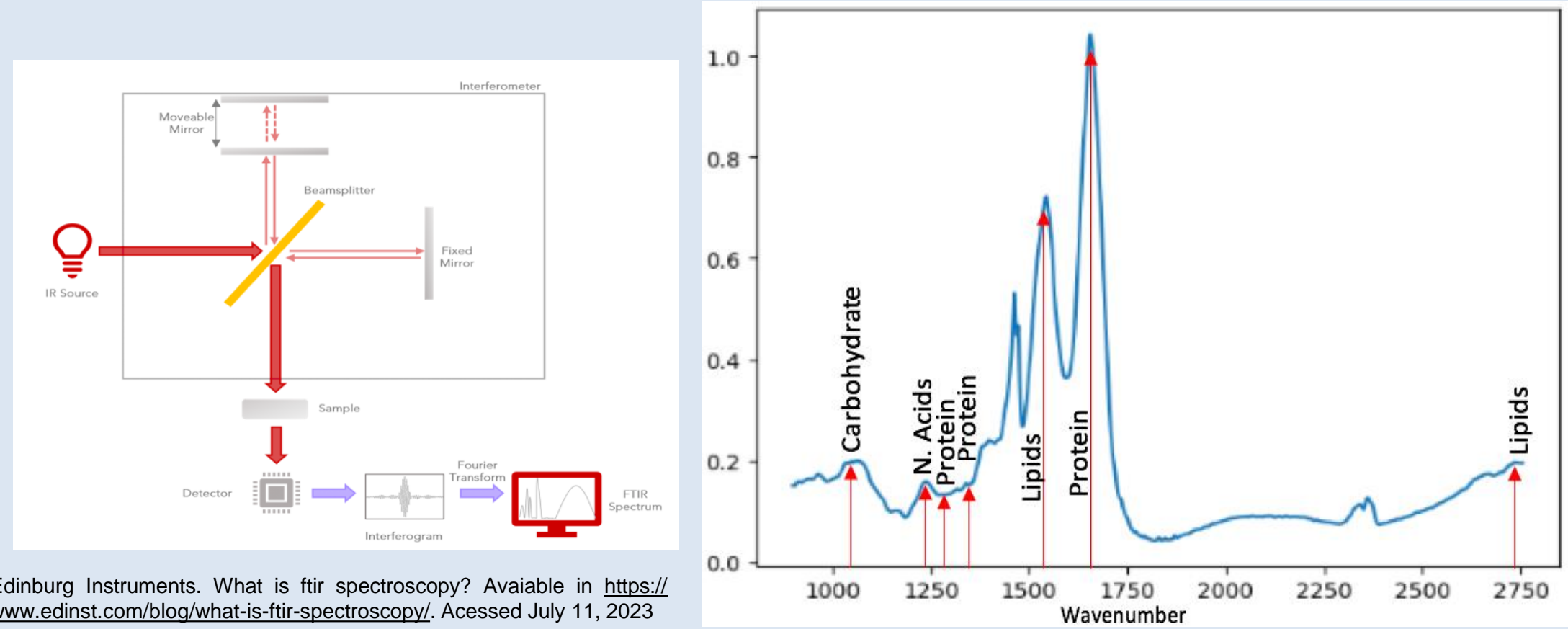Thyroid 2%
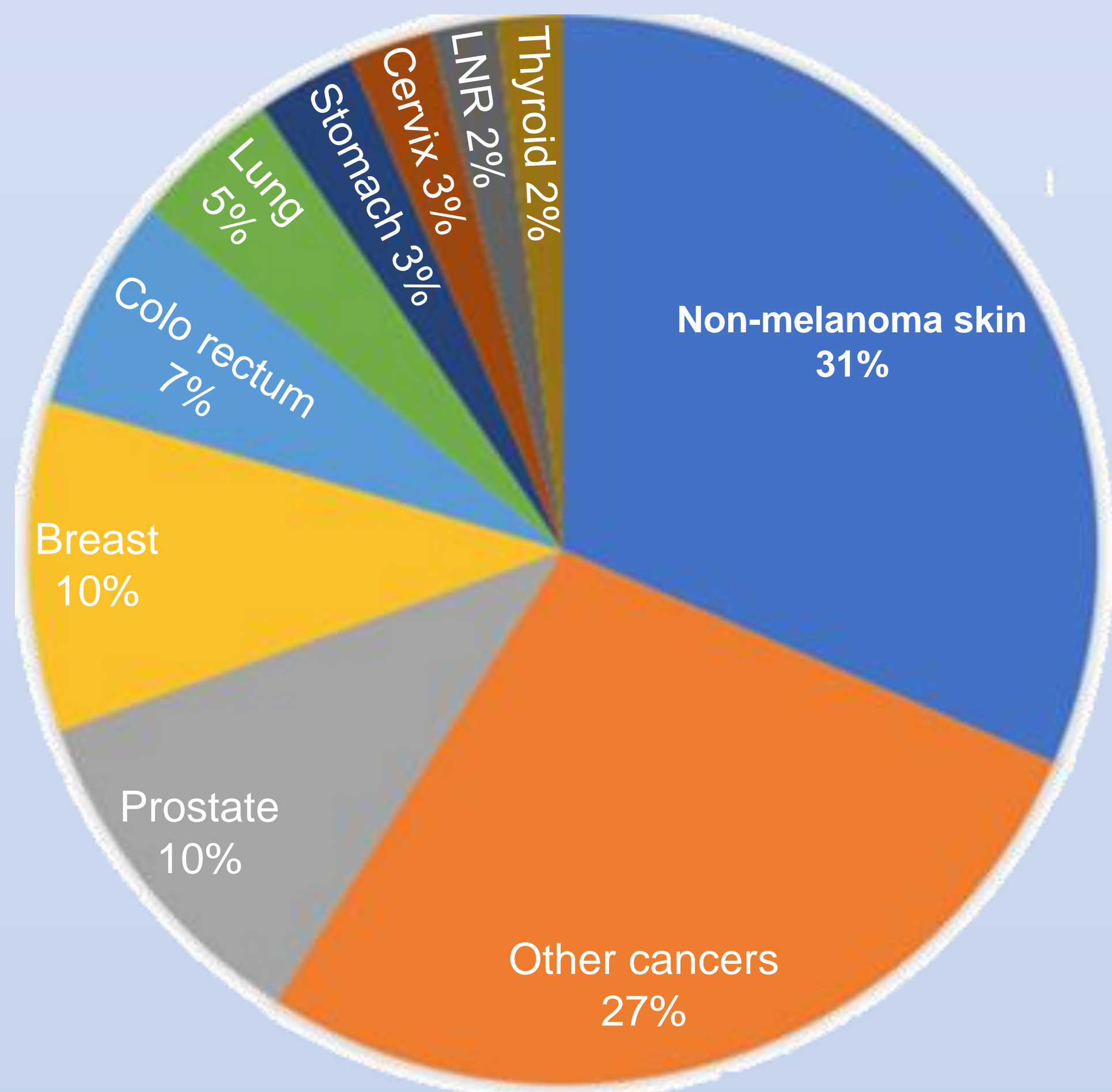
- Most of all cases of malignant neoplasms of the skin are NMSC;

- **Basal cell carcinoma (BCC)** is the most common (70%) type of NMSC.
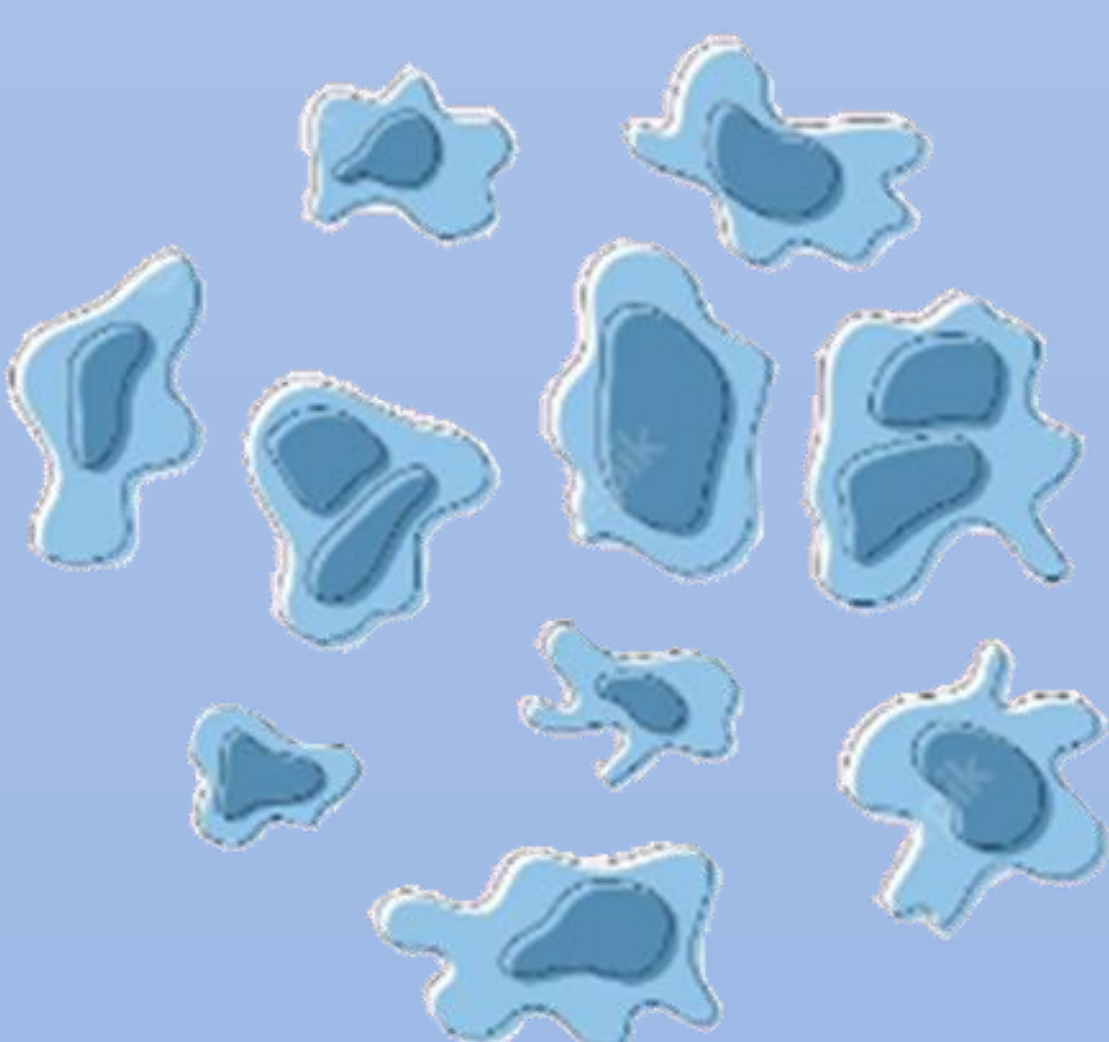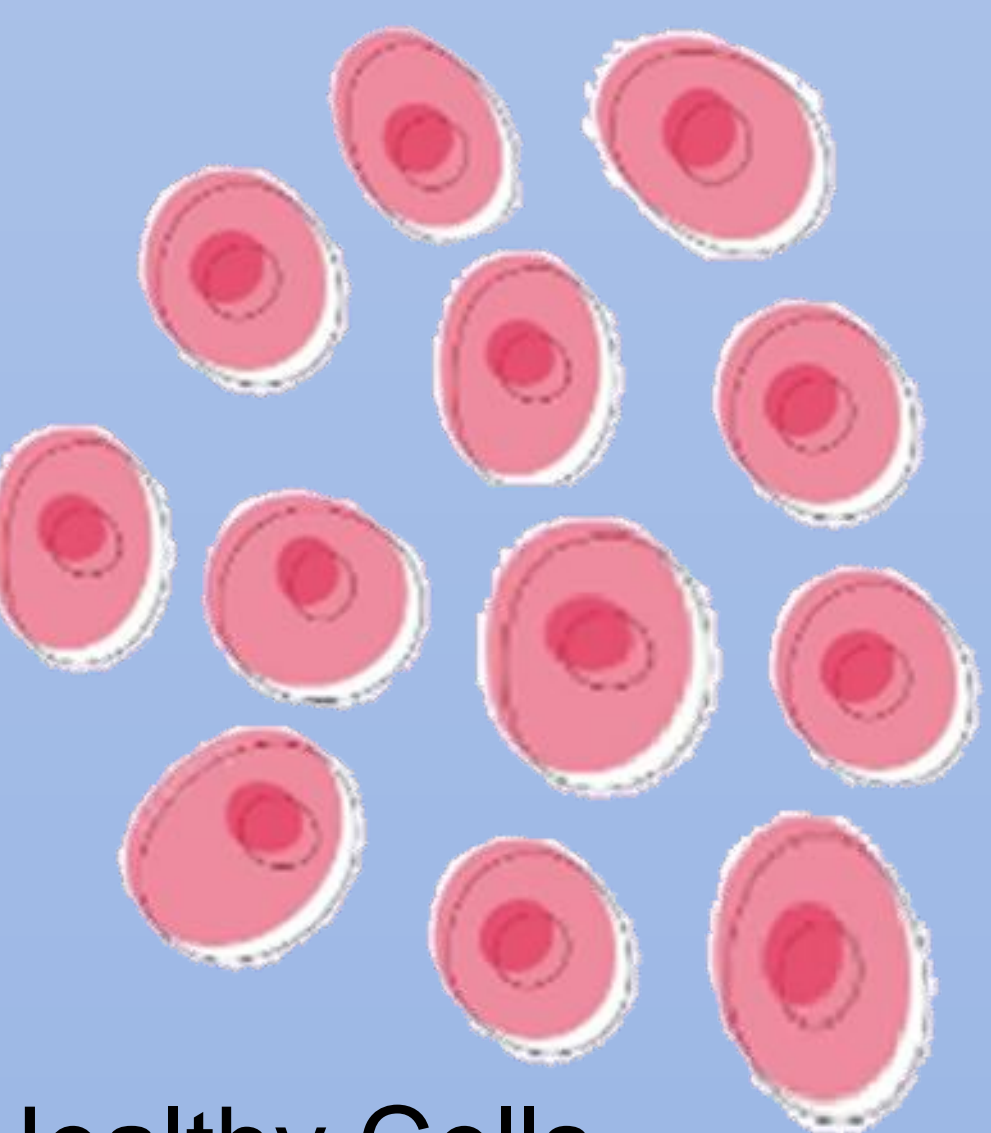
ATR-FTIR Spectroscopy
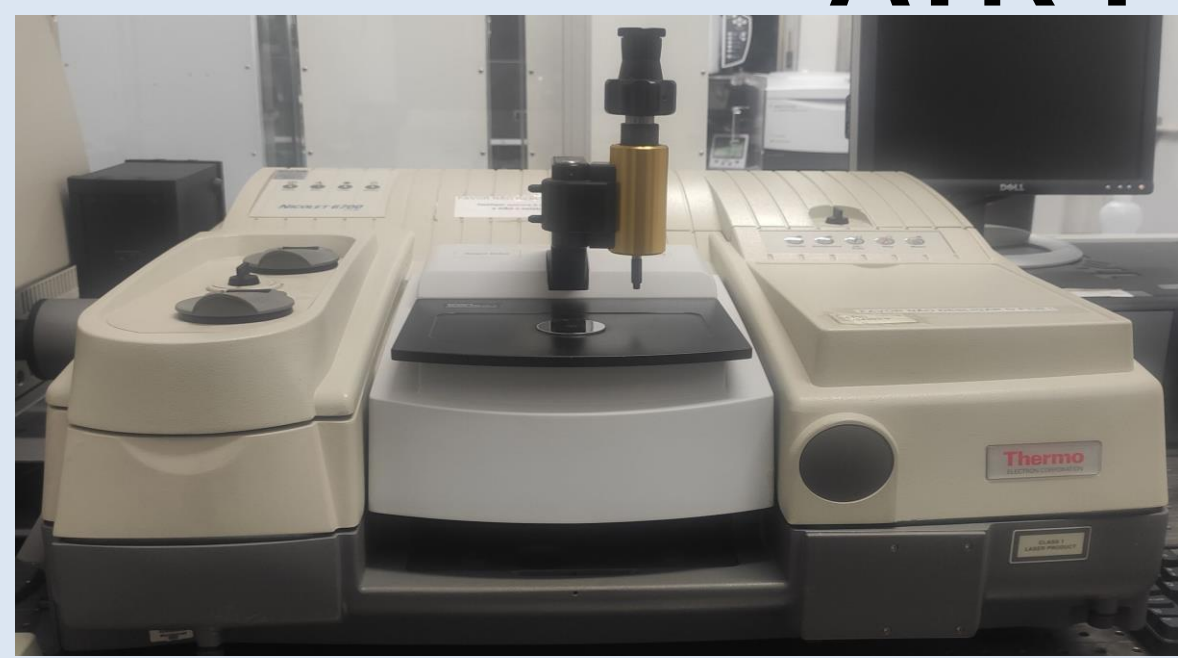
Minimal sample preparation

Edinburg Instruments. What is ftir spectroscopy? Avaiable in https://www.edinst.com/blog/what-is-ftir-spectroscopy/. Acessed July 11, 2023.
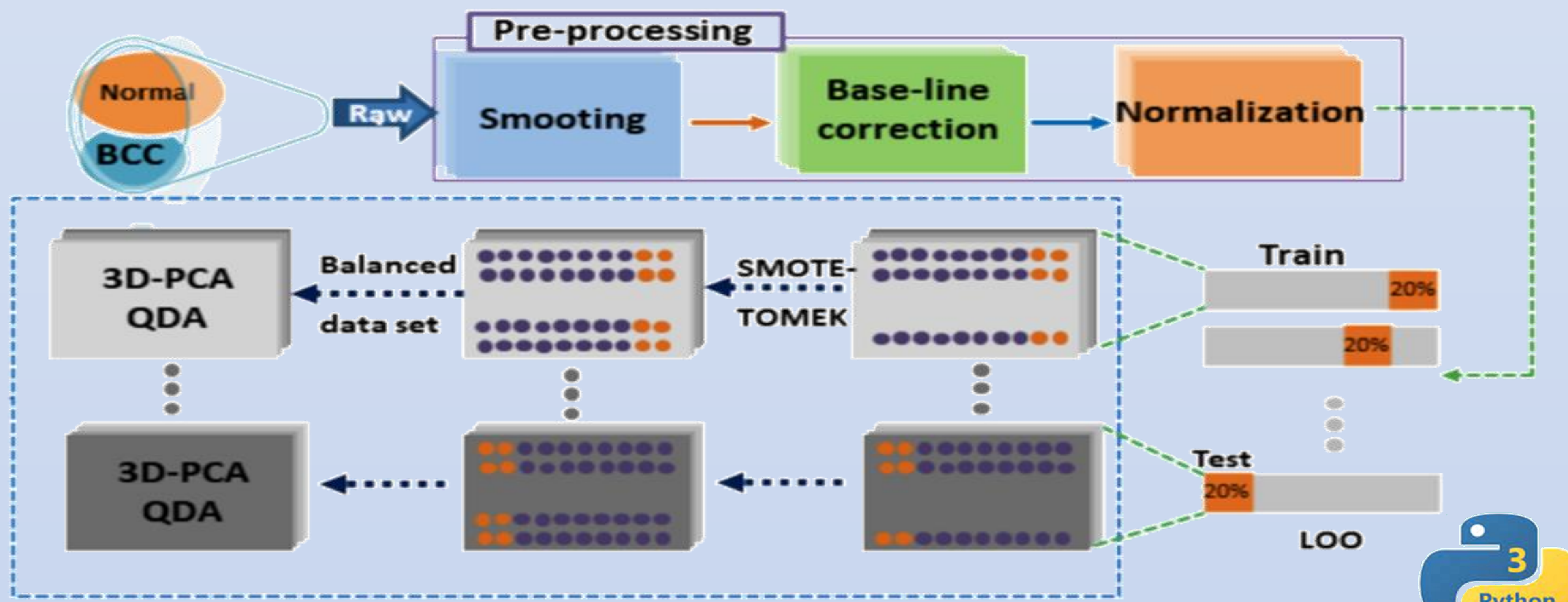
OBJECTIVES

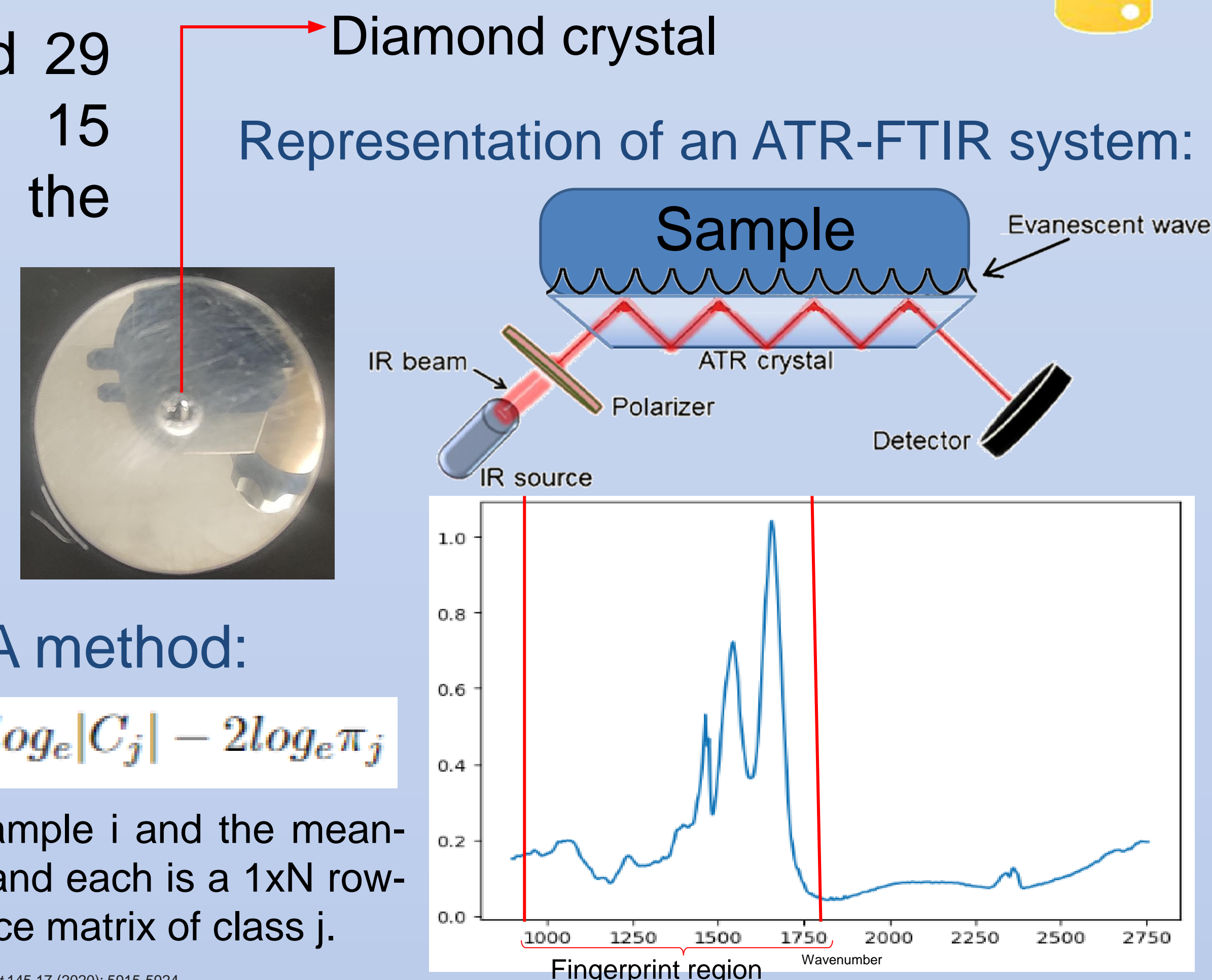3D-discriminant analysis + ATR-FTIR spectroscopy

Healthy Cells

BCC cells

MATERIAL AND METHODS

For this study, we used 29 basal cell carcinoma and 15 healthy skin samples. All the samples were parafinized.

- Spectral resolution of 4cm^{-1};
- 100 scans by spectrum.

Diamond crystal

Representation of an ATR-FTIR system:

Fingerprint region

Scores of 3D-PCA-QDA method:

$$Q_{ij} = (x_i - \overline{x_j})^T C_j^{-1} (x_i - \overline{x_j}) + log_e|C_j| - 2log_e\pi_j$$

x_i and x_j are the mean-scores of T for sample i and the mean-scores of class j for their respective PCs and each is a 1xN row-vector. C_j represent the variance-covariance matrix of class j.

*Morais, Camilo LM, et al. "A three-dimensional discriminant analysis approach for hyperspectral images." Analyst 145.17 (2020): 5915-5924.

RESULTS AND DISCUSSION

Average spectra

DF X number of samples

ROC curve
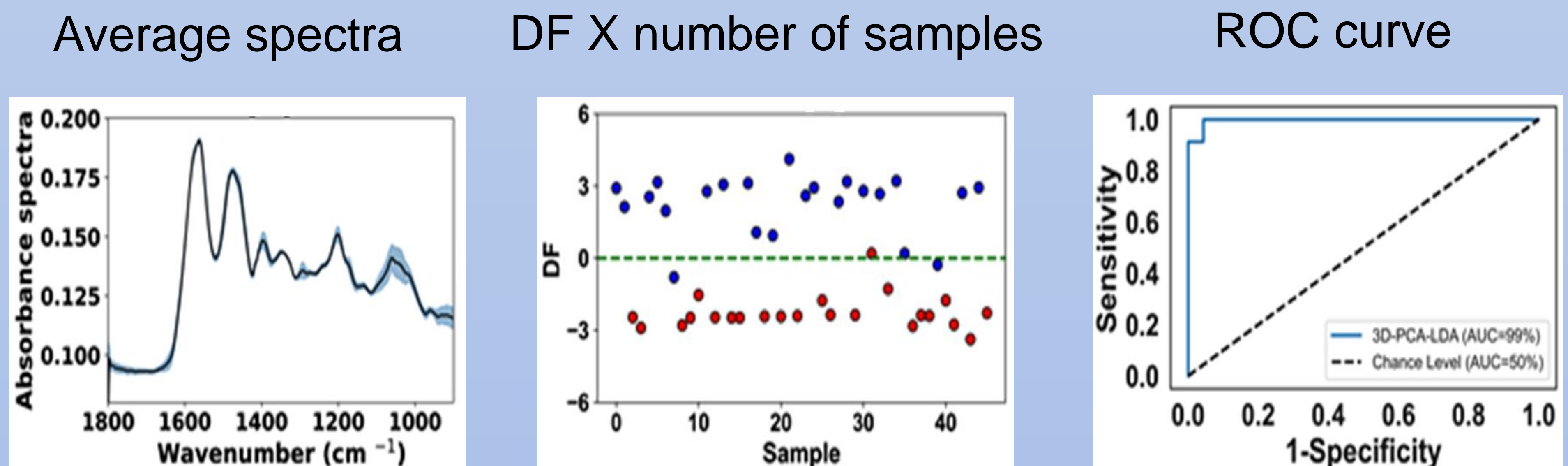
CONCLUSION

With a computational approach based on 3D-discriminant analysis (3D-PCA-QDA) of ATR-FTIR spectra, we were able to differentiate healthy and unhealthy human tissue samples. Our results demonstrated that the computational model achieved accuracy up to 99% and depicted interest to use it for basal cell carcinoma.

ACKNOWLEDGMENTS

GOVERNO DO ESTADO DE SÃO PAULO

ipen instituto de pesquisas energéticas e nucleares

cnen Comissão Nacional de Energia Nuclear

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES