

Biomarkers in Breast Cancer Diagnosis: an Exploratory Data Analysis

Saulo Mendes de Melo
Applied Computational Intelligence
PPGETI
Fortaleza, Brazil
saulomelo96@hotmail.com

Abstract—This document is a model and instructions for \LaTeX . This and the `IEEEtran.cls` file define the components of your paper [title, text, heads, etc.]. ***CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

Index Terms—Breast cancer, PCA, biomarkers, Data analysis

I. INTRODUCTION

Cancer is among the leading causes of death worldwide and every year its various types claim a great number of lives. Amongst women victims, breast cancer (BC) is the leading type, with as much as 322.000 deaths in the year 1990 [1]. A study published in 2019 analyzing the US population has found that approximately 13% of women will be diagnosed with BC in their lifetime, and as much as 1 in 39 women will eventually succumb to it [2].

The incidence of BC among women increases with age. The probability of a diagnose for woman in the age of 20 is of 0.1%, and goes to 3.0% by the age of 80. The rates of incidence and mortality also hold some relationship with ethnicity. The incidence rate is higher among whites (130.8 per 100.000), but for blacks, while incidence is lower the mortality is up to 40% higher (28.4 per 100.00) [2]. Other risk factors associated are late first birth (post 30), nulliparity, use of oral contraceptives and having first and/or second-degree relatives diagnosed [3].

Although non-clinical factors provide a valuable insight, other studies have focused on the analysis of biological indicators more tipically related to medical assessment. Dalamaga [4] has analyzed Resistin as biomarker, and it's links to obesity and cancer. Crisóstomo et al. [5] provided a study of biomarkers in the context of BC, where groups were separated for both obese and non-obese control and patients, and an extensive statistical analysis was made with indicators such as Glucose, Insulin, Resistin, among others. A set metabolic characteristics was found in obese women with BC, which includes glucoes, insulin disorders and other anomalies.

The present work aims to provide an analysis of biomarkers on a clinical dataset, evaluate how correlated they might be with BC detection in patients and the possibilities of dimensionality reduction through the use of Principal Component Analysis.

REFERENCES

- [1] C. J. Murray and A. D. Lopez, "Mortality by cause for eight regions of the world: Global burden of disease study," *The lancet*, vol. 349, no. 9061, pp. 1269–1276, 1997.
- [2] C. E. DeSantis, J. Ma, M. M. Gaudet, L. A. Newman, K. D. Miller, A. Goding Sauer, A. Jemal, and R. L. Siegel, "Breast cancer statistics, 2019," *CA: a cancer journal for clinicians*, vol. 69, no. 6, pp. 438–451, 2019.
- [3] V. G. Vogel, "Epidemiology of breast cancer," in *The breast*. Elsevier, 2018, pp. 207–218.
- [4] M. Dalamaga, "Resistin as a biomarker linking obesity and inflammation to cancer: potential clinical perspectives," *Biomarkers in medicine*, vol. 8, no. 1, pp. 107–118, 2014.
- [5] J. Crisostomo, P. Matafome, D. Santos-Silva, A. L. Gomes, M. Gomes, M. Patrício, L. Letra, A. B. Sarmento-Ribeiro, L. Santos, and R. Seïça, "Hyperresistinemia and metabolic dysregulation: a risky crosstalk in obese breast cancer," *Endocrine*, vol. 53, no. 2, pp. 433–442, 2016.