v. 4, n. 96, 2024

International Journal of Health Science

Acceptance date: 05/11/2024

ANALYSIS AND COMPARISON OF THE ROMA INDEX WITH HEMOGRAM DATA AND CLINICAL SIGNS AND SYMPTOMS IN BRAZILIAN WOMEN WITH OVARIAN CANCER

Paula Santos
Responsible researcher, Beevi, Ribeirão Preto, São Paulo, Brazil

Luciana Chain Veronez
Research and Development, Beevi, Ribeirão Preto, São Paulo, Brazil

Alef Janguas
Bioinformatics, Beevi, Ribeirão Preto, São Paulo, Brazil

Abstract: One of the main discussions about women's health is the efficiency and effectiveness of using simplified methodologies, such as isolated biomarkers, to support differential diagnosis. Based on this premise, we used a database to identify the specificity and sensitivity of the ROMA index in detecting ovarian cancer compared to an intelligent computational model. This model was trained with anamnesis and blood count data to differentiate subtle nuances between patients with a potential diagnosis of ovarian cancer. The premise is that this type of model can provide better groupings for risk factors, allowing for more accurate data classifications that can better describe early patterns of silent diseases, such as ovarian cancer, in gynecological clinics, or at least in the early stages, to reduce mortality. Our goal is to identify clusters within data on signs and symptoms, combined with data on the immune and endocrine systems, that can group characteristics for better clinical outcomes in ovarian cancer. Our methodology was based on multicriteria analysis methods, achieving a sensitivity of 94.6% and a specificity of 97%, as measured by the kappa index. These results indicate the potential of our methodology to support physicians in the gynecological monitoring of women in Basic Health Units.
Keywords: artificial intelligence, ovarian cancer, early diagnosis

INTRODUCTION

Menstrual disorders pose a formidable challenge to global women's health initiatives. They affect the lives of countless women, causing physical discomfort and psychological disturbance, and can, in certain cases, negatively affect fertility. The imperative of early detection and effective intervention is critical to improving the well-being of affected women[1,2].

Menstruation constitutes a fundamental physiological process in the female reproductive life cycle, necessarily a crucial parameter for reproductive health, marked by consistent cyclical patterns, duration and sufficient blood discharge. However, the configurations of a "typical menstrual cycle" are subject to individual variations and evolve throughout the different stages of a woman's life. Faced with this reality, countless women around the world face menstrual irregularities, which can significantly reduce their quality of life and general health[2,3].

Therefore, our efforts are in-depth in the analysis of blood count data and hormone levels to identify changes and risks for menstrual disorders and neoplasms of the female reproductive system. Studies of this magnitude aim to identify risk factors during routine assessments in clinics and basic health units during the gynecological routine by health professionals such as doctors and nurses, as well as specialists in the field of oncology[4].

Faced with this growing disruptive scenario in the health sector, the development of platforms to support the tracking and monitoring of neoplasms such as ovarian cancer, using laboratory test data to observe small physiological changes that allow the identification of different nuances between the health-disease process, significantly benefiting early diagnosis, in addition to progressively monitoring changes, especially benign gynecological conditions, which can affect fertility, quality of life and because they present potential characteristics that favor the appearance of neoplasms. Such analysis can also promote greater public health policies for women's health and maternal health[5].

Data from the World Health Organization indicate that 314,000 women were diagnosed with ovarian cancer worldwide and approximately 69% die from ovarian cancer, which means that out of every 10 women, 7 are diagnosed. In the United States, there are an average of 22,440 new cases of ovarian can-

cer, a disease that kills about 14,000 women annually. In Brazil, the risk for each year from 2023 to 2025 is 7,310 cases, corresponding to an estimated risk of 6.62 new cases for every 100,000 women. Although ovarian cancer has a lower prevalence compared to breast cancer, it is three times more lethal, and the mortality rate from this type of cancer is expected to increase significantly by the year 2040[1-4]. This is explained by the fact that ovarian cancer presents risk factors associated with age, diet, personal and family history of ovarian cancer, breast cancer or cancer, reproductive history, hormonal therapy after menopause, genetic changes such as and/or mutations. Furthermore, studies show that the CA-125 tumor marker has low specificity for early diagnosis, as do ultrasound and the detection of the BRCA1 and BRCA2 genes. Data on ovarian cancer show that the disease has heterogeneous characteristics, being made up of more than 15 distinct types with a high mortality rate of 80% among women aged 45 to 65 years[3-6]. The proliferation of this neoplasm is autonomous, abnormal and uncontrolled, mainly in the lining epithelial tissue, a type of tissue derived from germ cells. Thus, there are several types with specific risk factors, clinical, molecular and pathological characteristics. We can also observe influences on germ cells, such as environment, age, reproductive and hormonal factors involved in the disease process. Indications for ovarian cancer mainly involve simplified CA-125 analysis and ultrasound evaluation to classify and identify the ovarian system[4-6]. However, these early diagnosis methods lack better defined parameters or the use of artificial intelligence to increase the risk of ovarian cancer. Another important aspect is that many of these platforms still lack analysis regarding accuracy, severity levels and study of what may be involved in ovarian cancer. Furthermore, studies of this magnitude allow the description of the baseline of a pathology in the country or regions. Furthermore, it can indicate to healthcare professionals the appropriate treatment measures for each stage of the disease, as the tool helps in understanding the disease. In conclusion, a platform with more accessible high-tech artificial intelligence models will significantly improve research and education. Furthermore, it will improve the quality of care, tracking and ovarian diagnosis, ensuring more accurate and timely interventions.

METHODS

Dataset: For the project, 2500 blood counts were acquired from women with: hormonal laboratory tests and two markers CA 125 and HE4, which did not contain the patients' personal data, only the age group that was from 35 years old and female biological sex. Regarding signs and symptoms, we collected general data on the patient's complaints, and signs evaluated during the physical examination.

Control Group:

- Group I: Women with no history of any gynecological alterations, with clinical exams within normal limits, and no history of sexually transmitted infections.
- Group II: Women with a history of treatment over 12 months ago, excluding cases of HPV and syphilis, who no longer exhibit signs and symptoms.
- Group III: Women diagnosed with ovarian cancer, post-surgical treatment.

GROUP WITH CONFIRMED DIAGNOSIS:

In the patient's electronic medical record, we will classify the diagnostic hypotheses as follows: Malignant neoplasm of the ovary and Benign neoplasm of the ovary. In this way, we will collect blood count tests, with hormonal levels and the molecular marker CA-125, for grouping: Stage 1 ovarian cancer; Stage 2 ovarian cancer; Stage 3 ovarian cancer; Stage 4 ovarian cancer.

International Journal of Health Science ISSN 2764-0159 DOI https://doi.org/10.22533/at.ed.15949624291010

Marie.IA PLATFORM[8,9,10]:

This study utilizes an intelligent computational model inspired by the work of Goodfellow et al. (2014), known for introducing Generative Adversarial Networks (GANs). The proposed methodology involves simulating 15,000 cases of ovarian cancer based on data extracted from scientific articles. The model uses this data to generate a wide variety of ovarian cancer cases, enabling an in-depth analysis of the characteristics associated with the disease.

After the simulation, the resulting data is used to test the efficacy of complete blood counts and specific biomarkers related to the pathophysiology of ovarian cancer. The process involves analyzing patterns in blood tests and correlating these patterns with the presence of biological markers that may indicate the presence or progression of the disease. The methodology aims to identify combinations of biomarkers and hematological parameters that could improve early detection and monitoring of ovarian cancer, providing a foundation for future clinical research.

Data Extraction: Specific clinical details from the anamnesis included age, signs and symptoms, hereditary tumors, diet as determinants for diagnosis, surgical procedures, prescribed medications, and etiology of diseases. Information extraction techniques such as Learning Patterns and Relationship Analysis were utilized to identify and structure this information in a structured format.

Validation and Evaluation: The evaluation metrics used were precision, recall, F1-score, ROC curve, confusion matrix, sensitivity, specificity, and accuracy.Patient symptom data. Data Analysis of Complete Blood Count, Hormone Levels, HE4 and CA-125: In this stage, the relationships between the following parameters were analyzed: red blood cells, hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (HCM), mean corpuscular hemoglobin concentration (MCHC), red blood cell distribution width (RDW), leukocyte count, total leukocytes, neutrophils, segmented neutrophils, lymphocytes, monocytes, eosinophils, basophils, platelets, growth hormone (GH), follicle-stimulating hormone (FSH), luteinizing hormone (LH), insulin, thyroid-stimulating hormone (TSH), testosterone, estradiol, parathyroid hormone (PTH), albumin, HDL cholesterol, low-density lipoprotein (LDL), creatinine, ferritin, and CA-125. Additionally, the data present in the blood count, including vitamins, minerals, and other elements, were evaluated to identify any patterns.

Data Mining: The data was organized and transformed from raw data into a dataset without empty fields and with all mandatory information. We analyzed the dataset for outliers (using oversampling and undersampling) and transformed variables to ensure they met the assumptions of the analysis method.

Feature Engineering: The goal of this step was to improve the performance of machine learning algorithms by reducing noise in the data, increasing the accuracy of predictions, and making the model more interpretable. Techniques used included normalization and standardization, time series analysis, and dimensionality reduction.

Model Training: Supervised machine learning methods used included decision trees, Random Forest, Gradient Boosting, and Support Vector Machine. Non-parametric discriminant analysis methods, such as k-Nearest Neighbors and Support Vector Machines, were applied based on the proximity of observations in feature spaces.**Statistical Analysis**: validation of diagnostic performance test.

RESULTS

The image shows a heatmap representing the correlations between various variables related to risk and biological markers in a dataset. Each cell in the heatmap indicates the correlation coefficient between two variables, with colors ranging from blue (negative correlation) to red (positive correlation). The intensity of the color reflects the strength of the correlation, with the gradient on the right sidebar indicating the correlation value, which ranges from -1 to 1. The variables include risk factors, levels of CA125, HE4, and other relevant laboratory metrics (Figure 1).

The image shows a line graph representing the performance of different classifiers based on the number of features used. The horizontal axis indicates the number of features, while the vertical axis displays the accuracy of the models. The RadialSVM classifier (blue line) demonstrates inconsistent performance, with a sharp drop in accuracy when more than 10 features are utilized. Its accuracy remains relatively low and stable between 0.70 and 0.75 as the number of features increases (Figure 2). In contrast, the RandomForest classifier (orange line) maintains consistently high accuracy, close to 1.0, regardless of the number of features, indicating its robustness concerning the number of variables selected.

AdaBoost (green line) generally shows high performance with slight variations as the number of features changes, consistently maintaining accuracy above 0.95. The DecisionTree classifier (red line) exhibits significant fluctuations in accuracy, though it generally remains around 0.95, with occasional drops and peaks depending on the number of features. The KNeighbors classifier (purple line) experiences a notable decline in accuracy as more features are included, particularly after 20 features, where accuracy drops to around 0.85 and continues to decrease. GradientBoosting (brown line) displays consistent and high performance, with accuracy close to 1.0, similar to RandomForest.

LinearSVM (pink line) shows fluctuating accuracy, but it tends to remain around 0.95, with a slight decrease as the number of features increases. Finally, the Logistic classifier (gray line) also exhibits strong performance, with accuracy varying little and staying close to 1.0 most of the time.

Based on the heatmap provided, the Logistic Regression classifier demonstrates consistently high performance across all thresholds, particularly from a threshold of 0.6 onwards, where it achieves an accuracy of 1.0000. This indicates that Logistic Regression has a strong ability to correctly classify data, making it a reliable choice for testing data. The other algorithms show more variability and do not consistently achieve perfect scores, especially at lower thresholds, reinforcing the decision to choose Logistic Regression for its stability and high accuracy (Table 1).

The comparison between the ROMA index used in ovarian cancer diagnosis and the Logistic Regression algorithm can be made based on the provided performance metrics for both.

For the ROMA index, the accuracy is 0.67619, sensitivity is 0.844828, and specificity is 0.468085. In contrast, the Logistic Regression algorithm shows an accuracy of 0.99, precision of 0.98 for class 0 and 1.00 for class 1, recall of 1.00 for class 0 and 0.97 for class 1, and an F1-score of 0.98 for class 0 and 0.99 for class 1. Additionally, the Logistic Regression algorithm has a train accuracy of 0.9749, an F1 Score of 0.987, and an accuracy of 0.9282 in 5-fold cross-validation.

When comparing these metrics, the Logistic Regression algorithm demonstrates significantly higher accuracy (0.99) compared to the ROMA index (0.67619), indicating that Logistic Regression is much more effective in correctly classifying cases. The sensitivity of the ROMA index (0.844828) is good but still

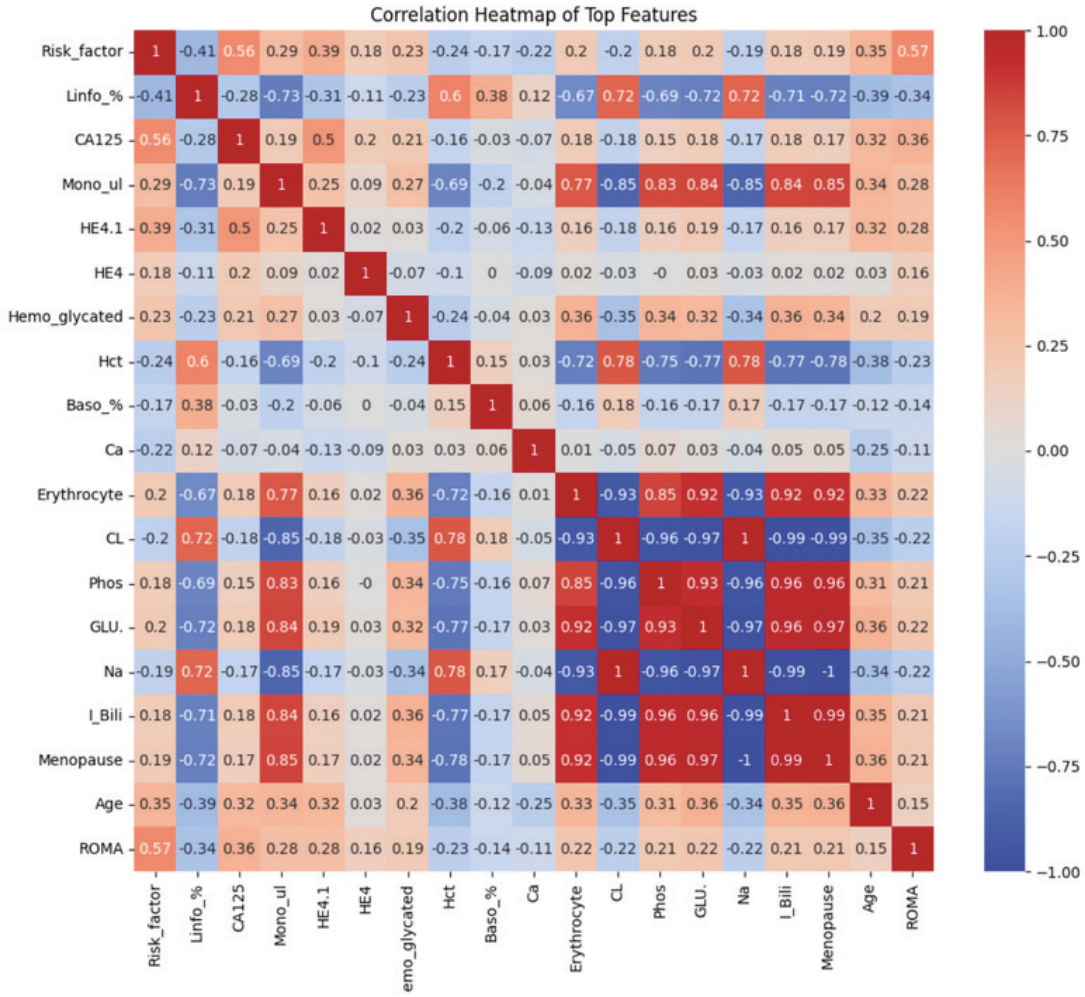

Figure 1: Correlation Heatmap of Top Features

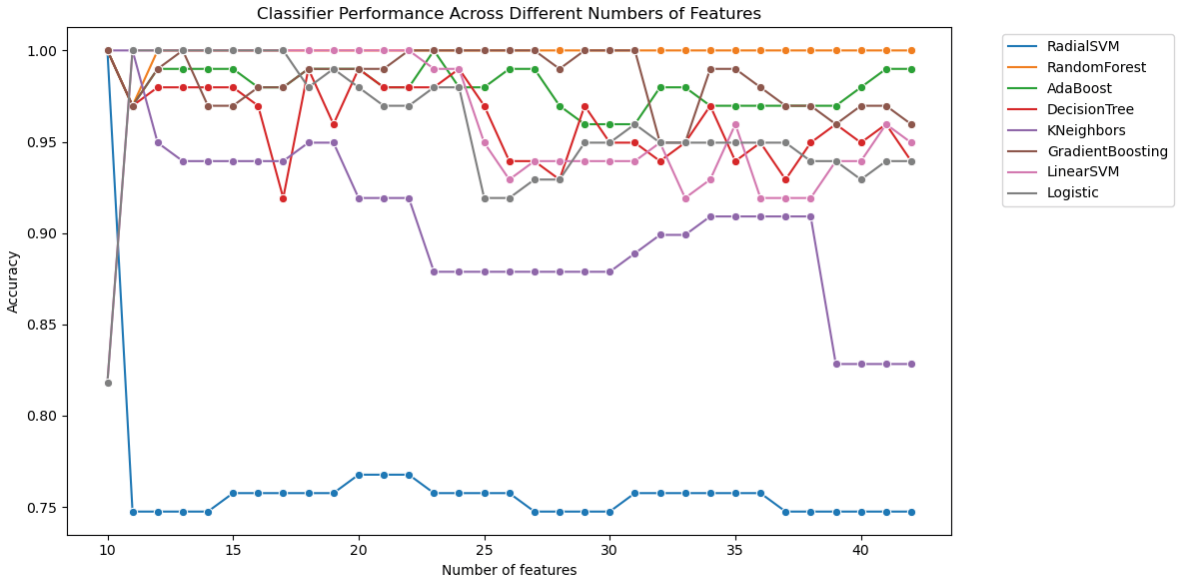

Figure 2: Classifier performance across different numbers of features

Classifier	0.4	0.5	0.6	0.8	0.9	0.95	1.0
RandomForest	0,6477270	0,8295450	1,0000000	1,0000000	0,9886360	0,9772730	0,9545450
GradientBoosting	0,6704550	0,6704550	0,7613640	0,7613640	0,7613640	0,7613640	0,7613640
AdaBoost	0,6363640	0,7840910	0,9772730	0,9886360	0,9659090	0,9772730	0,9772730
Logistic	0,6931820	0,9886360	1,0000000	1,0000000	1,0000000	1,0000000	1,0000000
DecisionTree	0,6931820	1,0000000	0,9886360	0,9886360	0,9886360	1,0000000	0,9886360
LinearSVM	0,6818180	0,9772730	0,9659090	0,9659090	0,9772730	0,9772730	0,9772730
KNeighbors	0,6363640	0,6477270	0,8750000	0,8295450	0,8295450	0,8409090	0,8295450
RadialSVM	0,7272730	0,9886360	1,0000000	1,0000000	1,0000000	0,9886360	1,0000000

Table 1: Classifier Performance Across Different Thresholds

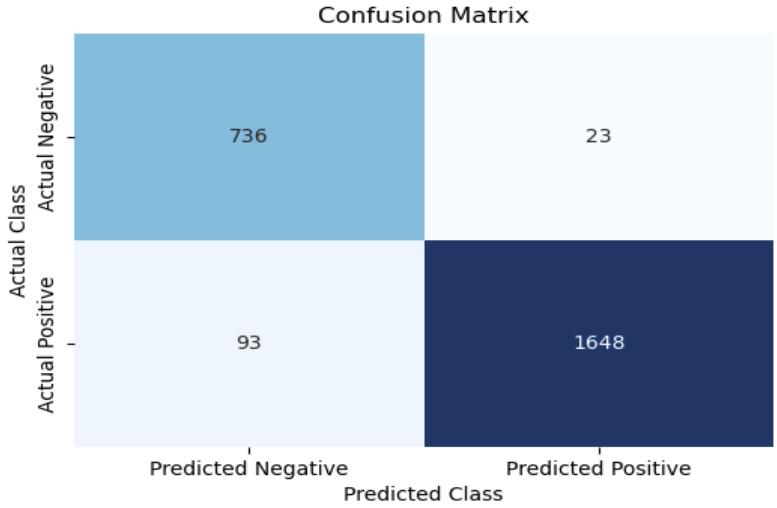

Figure 3: Confusion Matrix

lower than the sensitivity observed in Logistic Regression (1.00 for class 0 and 0.97 for class 1). Logistic Regression captures almost all positive cases, showing higher sensitivity. The specificity of the ROMA index (0.468085) is considerably lower than the performance of Logistic Regression, which shows high precision for both classes, indicating that Logistic Regression has a much better ability to avoid false positives. The F1 Score of Logistic Regression is nearly perfect (0.99), reflecting the balance between precision and recall. This balance indicates the superiority of Logistic Regression compared to the ROMA index.

In conclusion, the Logistic Regression algorithm demonstrates significantly superior performance compared to the ROMA index in terms of accuracy, sensitivity, specificity, and F1 Score. This suggests that Logistic Regression is a more robust and effective tool for classification and diagnosis in ovarian cancer cases compared to the ROMA index.

The confusion matrix (Figure 3) displayed shows the classification results of a diagnostic test. It consists of four quadrants: True Negatives (736), False Positives (23), False Negatives (93), and True Positives (1648). True Negatives represent the number of cases where the actual class was negative, and the test correctly predicted it as negative. False Positives indicate the cases where the actual class was negative, but the test incorrectly predicted a positive result. False Negatives are the cases where the actual class was positive, but the test failed to detect it, predicting a negative result instead. Finally, True Positives are the cases where the actual class was positive, and the test correctly identified them as such. The Y-axis represents the actual class labels, with the top being "Actual Negative" and the bottom being "Actual Positive," while the X-axis represents the predicted class labels, with the left being "Predicted Negative" and the right being "Predicted Positive".

DISCUSSION

Our results indicate that the risk factors for abdominal pain predominantly on the left side and sporadic, abdominal distension outside of menstrual periods and constipation, and the moderate positive correlation with CA125 (0.56) and albumin (0.57), indicated by lighter shades of red in the heat map, may be related to the pathophysiology of ovarian cancer. This relationship presents a differential for the clinical and physical evaluation of gynecological patients as a risk factor to stratify the population of women at risk for ovarian cancer. However, pain in the lower abdominal region is also present in diverticulosis and diverticulitis, thus, an investigation is necessary to identify and differentiate these pathologies, since there is a relationship between diverticulitis and diverticulosis with CA 125. Studies indicate that CA125 has been used to predict complications in acute colonic diverticulitis, showing potential as a marker for disease severity. Therefore, symptoms such as lower abdominal pain, whether on the right or left side, should be investigated as this region is related to other pathologies.[7,11].

Regarding CA125, data show that particularly in advanced stages of ovarian cancer, since high levels of CA125 are associated with the presence of tumor cells in the peritoneum and during inflammatory processes, the use of CA125 alone to identify early stages of ovarian cancer is therefore inefficient. The positive correlation between the risk factor variable and CA125 suggests that these reported symptoms may be directly related to increased tumor activity and associated inflammation, common characteristics in ovarian cancer. The increase in CA125 may be a reflection of the extent of the disease or the inflammatory response to the tumor. Thus, the observation of symptoms such as lower abdominal pain, constipation and abdominal distension outside may be symptoms that point to a potential triad of the pathophysiology of ovarian cancer.[7,11,12,13].

Albumina is a plasma protein essential for maintaining oncotic pressure and transporting various substances in the blood. In the context of cancer, low albumin levels may indicate malnutrition or chronic inflammation, both of which are common in patients with advanced cancer. The moderate positive correlation between risk factor variable and albumina may suggest that, although albumin is generally reduced in advanced disease states, its relationship with symptoms of abdominal pain and bloating outside of menstrual periods may reflect the complex interaction between the patient's nutritional status and tumor burden. Patients with more pronounced symptoms might paradoxically have less altered albumin levels, possibly due to compensatory mechanisms or individual variability in response to the tumor[14,15].

These moderate correlations between risk factor variables, CA125, and albumina may therefore provide useful information for understanding the underlying mechanisms of ovarian cancer and for the clinical assessment of symptoms reported by patients. Detecting these patterns can assist in early diagnosis, monitoring disease progression, and personalizing treatment[15,16].

In the context of ovarian cancer pathophysiology, the relationship between electrolytes such as chloride (Cl-) and sodium (Na), glucose receptors (GLU), phosphorus (Phos), and indirect bilirubin (I_Bili) can provide insights into metabolic and biochemical alterations associated with the development and progression of the disease. Chloride and sodium are essential electrolytes that play fundamental roles in maintaining osmotic balance and the electrical potential of cells. In cancerous cells, such as those in ovarian cancer, changes in the concentration of these ions may reflect alterations in the tumor microenvironment, which is often characterized by hypoxia, aci-

International Journal of Health Science ISSN 2764-0159 DOI https://doi.org/10.22533/at.ed.15949624291010

dosis, and electrolyte imbalance. Alterations in Cl- and Na- levels can influence tumor invasion and metastasis, as the acidic microenvironment and osmotic stress can promote cell proliferation and resistance to apoptosis (programmed cell death)[16,17,18].

Ovarian cancer, like many other types of cancer, may exhibit increased glucose uptake, known as the Warburg effect. This effect is characterized by the anaerobic metabolism of glucose to generate energy, even in the presence of oxygen. Glucose receptors (GLUTs) are often overexpressed in tumor cells to meet the high energy demand. The overexpression of glucose receptors is associated with tumor aggressiveness and a poorer prognosis, as the elevated glucose uptake sustains the growth and survival of cancer cells[18].

Phosphorus is an essential component of various biological molecules, including ATP (adenosine triphosphate), which is the primary source of cellular energy. Alterations in phosphorus levels may be associated with the uncontrolled cell proliferation observed in tumors. Additionally, altered phosphorus metabolism can influence protein phosphorylation, which is a critical process in regulating cell signaling, tumor growth, and treatment resistance[17,18].

Indirect bilirubin is a byproduct of hemoglobin breakdown and is often elevated in conditions involving cell destruction or liver dysfunction. In advanced cancers, liver dysfunction may occur due to metastasis or tumor pressure on the liver, leading to increased bilirubin levels. Elevated levels of indirect bilirubin may be associated with oxidative stress and inflammation, which are characteristics of the tumor microenvironment and may contribute to the progression of ovarian cancer[17-19].

In conclusion, the relationship between Cl-, Na-, glucose receptors, phosphorus, and indirect bilirubin in ovarian cancer may reflect the complex biochemical and metabolic changes that occur during the development and progression of the disease. These alterations not only sustain tumor growth but also may serve as potential biomarkers for diagnosis, prognosis, and therapeutic targets.

The Logistic Regression algorithm's superior performance compared to the ROMA index for diagnosing ovarian cancer, as previously discussed, is further emphasized when considering the correlations between various risk factors, biomarkers, and biochemical parameters associated with ovarian cancer.

The ROMA index, which includes CA125 as a primary component, showed moderate accuracy and specificity, indicating its limitations in comprehensive diagnostic performance. This is particularly evident when comparing it to the Logistic Regression algorithm, which demonstrated near-perfect accuracy, precision, and recall. The ROMA index's moderate correlation with CA125 (0.56) reflects its reliance on this biomarker to assess ovarian cancer risk, which is often influenced by tumor activity and inflammation, manifesting as symptoms like abdominal pain and bloating. However, the Logistic Regression algorithm, by integrating multiple features and potentially capturing non-linear relationships, surpasses ROMA in effectively predicting the presence of ovarian cancer[18-19].

Additionally, the correlation between albumin levels and the risk factor variables provides insights into the complex interaction between a patient's nutritional status and tumor burden. While the ROMA index focuses primarily on CA125, the Logistic Regression model likely benefits from incorporating a broader range of variables, such as albumin, to achieve a more holistic assessment of disease status. The moderate positive correlation between albumin and the risk factors suggests that the Logistic Regression algorithm might better account for the nuanced interactions between inflammation, nutritional status, and tumor progression, which are critical in ovarian cancer pathophysiology[19,20].

International Journal of Health Science ISSN 2764-0159 DOI https://doi.org/10.22533/at.ed.15949624291010

The relationships between electrolytes (Cl- and Na-), glucose receptors, phosphorus, and indirect bilirubin also underscore the metabolic and biochemical complexities of ovarian cancer that the Logistic Regression model could capture more effectively than the ROMA index. For instance, alterations in electrolyte levels, which are crucial for maintaining osmotic balance and cellular functions, could be indirectly associated with tumor microenvironment changes, such as hypoxia and acidosis, which the Logistic Regression model might identify as significant predictors of cancer progression.

Moreover, the Warburg effect, characterized by increased glucose uptake in cancer cells, and the overexpression of glucose receptors, could be better represented in the Logistic Regression model by including variables related to glucose metabolism, thus providing a more accurate prediction of tumor aggressiveness and prognosis compared to the ROMA index. The inclusion of phosphorus and indirect bilirubin levels in the Logistic Regression model could further enhance its predictive power by reflecting the tumor's metabolic demands and the systemic effects of advanced disease, such as liver dysfunction or oxidative stress.

In conclusion, the Logistic Regression model's ability to integrate a diverse range of variables and capture complex interactions makes it a more robust and effective tool for ovarian cancer diagnosis and prognosis compared to the ROMA index. The correlation between risk factors, biomarkers like CA125 and albumin, and metabolic indicators such as Cl-, Na-, glucose receptors, phosphorus, and indirect bilirubin provides a deeper understanding of the underlying mechanisms of ovarian cancer, which the Logistic Regression model can leverage to improve clinical outcomes. This comprehensive approach allows for better early diagnosis, monitoring of disease progression, and personalization of treatment strategies in ovarian cancer patients.

The provided confusion matrix can be interpreted as follows: the model correctly identified 1,648 cases of ovarian cancer (true positives) and 736 cases that were not ovarian cancer (true negatives). However, it made some errors, incorrectly classifying 23 cases that were not cancer as positive (false positives) and failing to identify 93 cases of ovarian cancer (false negatives).

With this information, we can calculate the model's sensitivity and specificity. Sensitivity, which measures the proportion of ovarian cancer cases correctly identified out of the total number of cancer cases, was approximately 94.6%. On the other hand, specificity, which measures the proportion of cases correctly identified as non-cancer out of the total number of negative cases, was about 97.0%.

This suggests that these 50 cases are part of the 93 false negatives, which may indicate that the model has a different performance when identifying early-stage ovarian cancer. The remaining 43 cases among the 93 false negatives likely refer to more advanced or intermediate stages of the disease that were also not identified by the model.

In conclusion, the model showed an overall sensitivity of 94.6% and a specificity of 97.0%. However, it failed to identify 50 cases of early-stage ovarian cancer, which points to a possible limitation in its ability to detect the disease in its early stages.

FUNDING

This work was supported by FAPESP - São Paulo Research Foundation, under grant number 24/01118-0. Additionally, we would like to acknowledge the financial support from FAPESP through the Marie.IA project: platform for screening ovarian neoplasms. The funders had no influence on the conduct of the study, collection, analysis or interpretation of the data, or on the writing of the manuscript.

REFERENCES

1. World Ovarian Cancer Coalition Atlas. World Ovarian Cancer Coalition. Disponível em: [https://worldovariancancercoalition.org](https://worldovariancancercoalition.org/wp-content/uploads/2023/04/World-Ovarian-Cancer-Coalition-Atlas-2023-FINAL.pdf). Acesso em: 3 jul. 2024.

2. Cancer Facts & Figures 2023. American Cancer Society. Disponível em: [https://www.cancer.org](https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2023.html). Acesso em: 3 jul. 2024.

3. Ovarian Cancer Statistics. Cancer Research UK. Disponível em: [https://www.cancerresearchuk.org](https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer). Acesso em: 3 jul. 2024.

4. FERLAY, J. et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer, v. 136, n. 5, p. E359-E386, 2015. DOI: 10.1002/ijc.29210.

5. SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer statistics, 2020. CA: A Cancer Journal for Clinicians, v. 70, n. 1, p. 7-30, 2020. DOI: 10.3322/caac.21590.

6. MCGUIRE, A.; BROWN, J. A.; MALIK, I. A.; et al. Effects of age on the detection and treatment of breast cancer. Cancers, v. 7, n. 2, p. 908-929, 2015. DOI: 10.3390/cancers7020819.

7. LIBERTI, Matthew V.; LOCASALE, Jason W. The Warburg Effect: How Does it Benefit Cancer Cells? Trends in Biochemical Sciences, v. 41, n. 3, p. 211-218, 2016. DOI: 10.1016/j.tibs.2015.12.001.

8. SANTOS, P. (2021). Intelligent computational model for the classification of Covid-19 with chest radiography compared to other respiratory diseases. *arXiv preprint arXiv:2108.05536*.

9. BENTO, Valdirene; SALAROLI, Bruno Frederico; SANTOS, Paula. MARIE: VALIDATION OF THE ARTIFICIAL INTELLIGENCE MODEL FOR COVID-19 DETECTION.

10. DEON, Reges Antonio; KORB, Arnildo; SANTOS, Paula. Artificial intelligence model in Body Scan images for monitoring tuberculosis in a prison complex. 2022.

11. Zager Y, Khalilieh S, Mansour A, Cohen K, Nadler R, Anteby R, Ram E, Horesh N, Nachmany I, Gutman M, Berger Y. The value of CA125 in predicting acute complicated colonic diverticulitis. Int J Colorectal Dis. 2023 Jun 30;38(1):182. doi: 10.1007/s00384-023-04478-7. Erratum in: Int J Colorectal Dis. 2023 Aug 9;38(1):209. doi: 10.1007/s00384-023-04486-7. PMID: 37389666.

12. GANAPATHY-KANNIAPPAN, Shanmugasundaram; GESCHWIND, Jean-Francois H. Tumor glycolysis as a target for cancer therapy: Progress and prospects. Molecular Cancer, v. 12, n. 1, p. 152, 2013. DOI: 10.1186/1476-4598-12-152.

13. MONTAGNANI, A.; DI FILIPPO, C.; CECCHI, E.; GRAZIANI, F. Electrolyte imbalances in cancer patients: A systematic review. Annals of Oncology, v. 24, supl. 4, p. iv44-iv48, 2013. DOI: 10.1093/annonc/mdt281.

14. HUANG, Chun; ZHA, Jing. The Impact of Electrolyte Imbalances on Tumor Biology. Journal of Cancer Research and Clinical Oncology, v. 146, n. 5, p. 1199-1210, 2020. DOI: 10.1007/s00432-020-03183-8.

15. BECK, Stacey E.; WEISS, Louis M. Phosphate metabolism in cancer. Expert Review of Anticancer Therapy, v. 12, n. 5, p. 631-637, 2012. DOI: 10.1586/era.12.29.

16. KWEE, S. A.; LIM, J.; KO, J. P.; ZHU, L. Phosphate metabolism in cancer: A focus on phosphate cytotoxicity in human colon cancer. Scientific Reports, v. 2, p. 588, 2012. DOI: 10.1038/srep00588.

17. VÍTEK, Libor. Bilirubin as a diagnostic marker and the role of bilirubin in disease pathophysiology. Current Pharmaceutical Design, v. 26, n. 35, p. 4308-4317, 2020. DOI: 10.2174/1381612826666200326161544.

18. TAPAN, Seyma; TANSEL, Ayfer; KOCER, Burak; KOSEOGLU, Hakan. The relation between serum bilirubin levels and other tumor markers in patients with different types of cancers. Turkish Journal of Gastroenterology, v. 21, n. 1, p. 5-10, 2010. DOI: 10.4318/tjg.2010.0030.

19. COLEMAN, R. L.; MONK, B. J.; SOOD, A. K.; HERZOG, T. J. Latest research and treatment of advanced-stage epithelial ovarian cancer. Nature Reviews Clinical Oncology, v. 10, n. 4, p. 211-224, 2013.

20. BAST, R. C. et al. A radioimmunoassay using monoclonal antibody to monitor the course of epithelial ovarian cancer. New England Journal of Medicine, v. 309, n. 15, p. 883-887, 1983.

21. JACOBS, I. J. et al. Multimodal approach to screening for ovarian cancer. The Lancet, v. 340, n. 8821, p. 268-271, 1992.