

Article

Extracting Knowledge from Machine Learning Models to Diagnose Breast Cancer

José Manuel Martínez-Ramírez ¹, Cristobal Carmona ^{1,2,3}, María Jesús Ramírez-Expósito ⁴
and José Manuel Martínez-Martos ^{4,*}

¹ Department of Computer Science, University of Jaén, E-23071 Jaén, Spain; jmmr0049@red.ujaen.es (J.M.M.-R.); ccarmona@ujaen.es (C.C.)

² Andalusian Research Institute in Data Science and Computational Intelligence, DASCI, University of Jaén, E-23071 Jaén, Spain

³ Leicester School of Pharmacy, DeMontfort University, Leicester LE1 7RH, UK

⁴ Experimental and Clinical Physiopathology Research Group CVI-1039, Department of Health Sciences, University of Jaén, E-23071 Jaén, Spain; mramirez@ujaen.es

* Correspondence: jmmartos@ujaen.es; Tel.: +34-953-212-600

Abstract: This study explored the application of explainable machine learning models to enhance breast cancer diagnosis using serum biomarkers, contrary to many studies that focus on medical images and demographic data. The primary objective was to develop models that are not only accurate but also provide insights into the factors driving predictions, addressing the need for trustworthy AI in healthcare. Several classification models were evaluated, including OneR, JRIP, the FURIA, J48, the ADTree, and the Random Forest, all of which are known for their explainability. The dataset included a variety of biomarkers, such as electrolytes, metal ions, marker proteins, enzymes, lipid profiles, peptide hormones, steroid hormones, and hormone receptors. The Random Forest model achieved the highest accuracy at 99.401%, followed closely by JRIP, the FURIA, and the ADTree at 98.802%. OneR and J48 achieved 98.204% accuracy. Notably, the models identified oxytocin as a key predictive biomarker, with most models featuring it in their rules. Other significant parameters included GnRH, β -endorphin, vasopressin, IRAP, and APB, as well as factors like iron, cholinesterase, the total protein, progesterone, 5-nucleotidase, and the BMI, which are considered clinically relevant to breast cancer pathogenesis. This study discusses the roles of the identified parameters in cancer development, thus underscoring the potential of explainable machine learning models for enhancing early breast cancer diagnosis by focusing on explainability and the use of serum biomarkers. The combination of both can lead to improved early detection and personalized treatments, emphasizing the potential of these methods in clinical settings. The identified markers also provide additional research and therapeutic targets for breast cancer pathogenesis and a deep understanding of their interactions, advancing personalized approaches to breast cancer management.

Keywords: breast cancer; serum biomarkers; explainable AI; oxytocin; early diagnosis; peptide hormones; IRAP; progesterone



Academic Editors: Alejandro Martin Sanchez, Angela Bucaro and Flavia de Lauretis

Received: 31 December 2024

Revised: 17 January 2025

Accepted: 29 January 2025

Published: 31 January 2025

Citation: Martínez-Ramírez, J.M.; Carmona, C.; Ramírez-Expósito, M.J.; Martínez-Martos, J.M. Extracting Knowledge from Machine Learning Models to Diagnose Breast Cancer. *Life* **2025**, *15*, 211. <https://doi.org/10.3390/life15020211>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Breast cancer is a disease in which breast cells grow abnormally and uncontrollably [1–3]. These cells can form a tumor, which may become palpable as a lump [2]. Invasive ductal carcinoma is the most common type of breast cancer. It originates in the cells lining the ducts of the breast and spreads outside the duct into the surrounding breast

tissue [1]. Approximately 80% of breast cancers are ductal carcinomas. Ductal carcinoma in situ (DCIS) is a non-invasive cancer that originated in the cells lining the ducts of the breast and has not spread outside the duct into the surrounding breast tissue [4]. Invasive lobular carcinoma originates in the lobules of the breast, where milk is produced, and spreads into the surrounding breast tissue [2,3].

The risk of breast cancer increases with age. Approximately 66.6% of breast cancer patients are older than 55 years [1,5–8]. Having a close relative with breast cancer, especially a first-degree relative, is also a risk factor [6,9]. Certain genetic mutations, such as mutations in the BRCA1 and BRCA2 genes, significantly increase the risk of breast cancer [8]. Prolonged exposure to estrogen, such as the early onset of menstruation, late menopause, or the use of hormone therapy, also increases the risk [6,8]. With regard to lifestyle, alcohol consumption, obesity, and a lack of physical activity are associated with an increased risk of the disease [2,6].

Prevention currently includes breast self-examination to detect any changes, such as lumps, skin changes, or nipple discharge. Screening and early detection methods include mammography [9–11], an X-ray of the breast used to detect breast cancer in the early stages, even before a lump can be felt, breast ultrasounds [8,11–13], which use sound waves to determine whether a lump in the breast is solid or a fluid-filled cyst, and breast MRI [8,13,14], an imaging test that uses a magnetic field and radio waves to create more detailed images of breast tissue and which is used to detect cancer in women at high risk or to better evaluate a finding from mammography or a physical examination. For definitive diagnosis, a biopsy is performed [5,10]. Ultimately, the treatment of breast cancer depends on several factors, such as the stage of the cancer, the type of cancer, and the age and general health of the patient. The treatment options include surgery, which may involve the removal of the tumor, part of the breast, or the whole breast (mastectomy) [5,8,10]; radiotherapy, which uses high-energy X-rays to destroy cancer cells [10,11]; chemotherapy, which uses drugs to destroy cancer cells [8,10,13,15]; hormone therapy, which is used to block the effects of hormones on cancer cells; and targeted therapy, which uses drugs that target specific characteristics of cancer cells [8,10,15].

Breast cancer is the most common cancer among women in the United States. It is estimated that there were about 2.9 million new cases in 2022 [7,16,17]. Globally, breast cancer accounts for 11.7% of cancer diagnoses and contributes significantly to mortality, especially among women over the age of 50 [1,8,18]. In 2020, there were 2.2 million new cases and 684,996 deaths from breast cancer [18]. The breast cancer incidence in 2020 was highest in Asia (45.4%) and Europe (23.5%), followed by North America (12.5%) and Latin America and the Caribbean (9.29%) [18]. Africa and Oceania reported lower incidence rates of 8.25% and 1.14%, respectively [7,18]. Asia also led in terms of breast cancer mortality with 50.5% of deaths, while Europe's mortality rate stood at 20.7%, probably reflecting better healthcare outcomes [18]. Africa's considerable mortality rate of 12.5% highlights critical challenges in medical care. In Spain, breast cancer causes 35,000 new cases per year and is the leading cause of cancer-related mortality in women, with 6,651 deaths per year [6,18].

An early diagnosis of breast cancer is directly associated with a higher chance of survival [5,10–12]. This is because early detection allows for treatment to be initiated in a timely manner, which increases the chances of success and reduces the risk of the disease spreading to other parts of the body (metastasis) [1,5,11,12]. The treatment options are also broader and less aggressive [6,11]. Late diagnosis may imply the need for more invasive treatments, such as a mastectomy instead of breast-conserving surgery [5,9,10]. In addition, chemotherapy and radiotherapy, which can have considerable side effects, may be necessary at later stages [5,10,15]. In general, breast cancer treatment is most effective

when initiated early in the disease because cancer cells are more sensitive when the tumor is small and has not spread [5,10–12]. The early diagnosis of breast cancer therefore allows patients to preserve a better quality of life as treatments are less invasive and have fewer side effects, allowing patients to maintain a more active and healthy life [11]. Finally, the early detection of breast cancer can reduce long-term healthcare costs because early-stage treatments are generally less expensive than treatments for advanced breast cancer [6,11].

In any case, it is important to emphasize the importance of early diagnosis in breast cancer management because it is crucial for improving survival rates and preventing breast cancer recurrence [5,6,10–12]. In fact, technology-based diagnostic tools are emerging. For example, a liquid biopsy is a non-invasive method for breast cancer screening that utilizes cancer-related biomolecules in the blood. Surface-enhanced Raman spectroscopy (SERS) combined with machine learning algorithms has shown great potential for rapid and accurate breast cancer detection. This technology enhances the Raman spectroscopy signal of serum using a specialized substrate, allowing for the discovery of new biomarkers. SERS combined with machine learning could achieve 100% accuracy, sensitivity, and specificity in distinguishing breast cancer patients from controls [19,20]. Similarly, Raman spectroscopy has been used with deep learning algorithms for screening triple-negative or HER-2-positive breast cancers [21].

Most data mining and machine learning studies mainly focus on the use of medical images and demographic data in studies of breast cancer diagnosis and prognosis [1,5,8,9,12–14,16,22–24]. However, serum biochemical parameters have been much less studied. Nevertheless, it is important to note that the integration of serum biomarkers can improve the accuracy of machine learning models for breast cancer risk prediction, early diagnosis, treatment response prediction, or recurrence detection [8,13,16,17]. Furthermore, the use of serum biomarkers offers a less invasive and more efficient approach for the early detection and monitoring of disease, also allowing for a binary classification of whether a patient has breast cancer or not. Unlike imaging techniques, which may have limitations in terms of their sensitivity and specificity, serum biomarkers can reflect early molecular changes associated with cancer development. In the context of the medical workflow, the extraction of serum biomarkers is performed through a blood test, a routine and minimally invasive procedure. This sample can be obtained at various stages of the healthcare process, from initial screening to monitoring the response to treatment. While conventional screening focuses primarily on mammography and other imaging techniques such as ultrasounds and MRI, it is important to note that these techniques have certain limitations. For example, mammography may not be as effective in women with dense breast tissue, and both ultrasounds and MRI can be expensive and are not available everywhere. In addition, these imaging techniques often detect cancer when a palpable tumor has already formed. Therefore, several artificial intelligence (AI) tools could be applied to a dataset of serum biomarkers with information on both patients and healthy individuals. By defining a classification problem [25], models could be trained in order to extract information from the previously mentioned dataset to diagnose whether a patient had breast cancer based exclusively on these biomarkers, which could prove useful to diagnose breast cancer when used along with other detection methods.

Though these tools often achieve highly accurate predictions, given the context of the problem, it is of the utmost importance that the results are not only precise but also explainable. The implemented systems should be able to explicitly define the reasons why their predictions occur, which in turn will allow the specialists who make use of them to discover new patterns in the data that may aid further research. For this purpose, black-box models, such as deep learning models, were ignored, and only models that offer rules or

trees that grant information on specific markers that are relevant for the diagnosis of breast cancer were used in an effort to achieve trustworthy AI (TAI) [26].

2. Materials and Methods

2.1. Subjects

A total of 198 women were recruited in the Unit of Breast Pathology at the University Hospital of Jaén, and 78 volunteer women without breast cancer were also included as a control group. This study was approved by the Ethical Committee of the University Hospital of Jaén, and all subjects signed a declaration of free, informed consent. All women with breast cancer were diagnosed with ductal infiltrating carcinoma. A total of 83 of these women (39 premenopausal and 44 postmenopausal) did not receive neoadjuvant chemotherapy, whereas 115 of them (63 premenopausal and 52 postmenopausal) received neoadjuvant chemotherapy before surgery. The clinicopathological characteristics of the studied patients have been previously described [27–31]. The patient characterization included the age at diagnosis, tumor size, tumor histology, pathologic T classification, Scarff–Bloom–Richardson grade, hormonal and HER-2/neu status, and molecular subtype. The control group consisted of 78 women, aged 28 to 69 years old (premenopausal women with regular menstrual periods, $n = 38$; postmenopausal women with spontaneous menopause for at least one year, $n = 40$), with no previous history of any type of cancer, chemotherapy, hormonal or antioxidant therapy, or chronic diseases. Women were excluded if they were current smokers, regular alcohol consumers, antioxidant supplement users, pregnant or lactating, presented hepatic, cardiac, or renal dysfunction, had undergone hormonal therapy, used drugs, or had hypertension, diabetes, and other eventual chronic conditions. Samples from patients treated with NACT were obtained after the completion of chemotherapy treatment and in parallel to samples from patients not treated with NACT and control volunteers in order to be processed under the same conditions. Blood samples were obtained after an overnight fast by venous arm puncture in tubes without anticoagulants. Blood specimens were allowed to clot and centrifuged at $3000 \times g$ for 10 min at 4°C to obtain the serum. Serum samples were collected, rapidly frozen in liquid nitrogen, and kept at -80°C until usage for assays. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

2.2. Datasets and Pre-Processing

The study subjects were all women of varying ages, as described in Section 2.1. This study aimed to make use of explainable models to accurately predict if a patient had breast cancer.

The dataset included data previously described in [27–35] and additional unpublished results. The dataset included electrolytes (sodium, potassium, chloride, calcium, phosphorus, and magnesium), metal ions (iron), albumin, the total protein, and marker proteins (Cerb2, also known as HER2/neu, P53, Ki67, and Bcl-2). It also included enzymes such as adenosine deaminase, CKMB (creatine kinase muscle–brain isoform), amylase, gamma-glutamyl-transferase (GGT), ALAT (alanine aminotransferase) and ASAT (aspartate aminotransferase), acid and alkaline phosphatases, CPK (creatine phosphokinase), and LDH (lactate dehydrogenase). Additionally, it included glucose; non-protein nitrogenous compounds like creatinine, urea, and uric acid; and a lipid profile with the total cholesterol, HDL (high-density lipoprotein), LDL (low-density lipoprotein), and triglycerides. Furthermore, it contained renin–angiotensin system-regulating aminopeptidases (APN, aminopeptidase N; APB, aminopeptidase B; ASAP, aspartyl aminopeptidase; APA,

aminopeptidase A; and IRAP, insulin-regulated aminopeptidase). The dataset also contained apolipoproteins A1 and B; hormones including GnRH (gonadotropin-releasing hormone), TSH (thyroid-stimulating hormone), FSH (follicle-stimulating hormone), LH (luteinizing hormone), GH (growth hormone), thyroxine, cortisol, angiotensin, vasopressin, oxytocin, β -endorphin, estradiol, progesterone, and prolactin, along with oxytocinase and pyrrolidone carboxypeptidase hormone-regulatory enzymes; and hormone receptors (ERs, estrogen receptors; PRs, progesterone receptors).

The dataset had several issues, and as such, several pre-processing techniques were applied. The first issue that the dataset had was that some of the attributes in it were specific for patients that had breast cancer and were not applicable to people that belonged to the control class. Consequently, all those attributes were ignored, since they would negatively impact the results of the model due to a bias towards one class if a patient presented any value whatsoever for those attributes. However, no examples were removed from the dataset, as it contained no errors and all data were relevant to the study. As such, it was not necessary to filter the dataset beyond removing those values.

Moreover, many of the attributes had several missing values. In order to fix this issue, the average of the existing values was calculated for each individual class (control or breast cancer), and the result was used to get rid of missing values.

Additionally, since the dataset was quite unbalanced, including roughly twice as many examples of breast cancer patients as healthy people, SMOTE [36] was used as an oversampling technique on the minority class, increasing the amount of examples it included by 50%. There were still more examples of the majority class, but it made the difference less notable. Though more examples could have been added, it would have caused the accuracy of the results to decrease due to an increase in noise in the dataset.

Lastly, the strategy used for the training of the models was a leave-one-out strategy, in which each model was trained with all examples in the dataset but one, which was used to test its accuracy. This process was repeated a number of times equal to the amount of examples in the dataset, so that every individual example could be used as a test once, as seen in Figure 1.

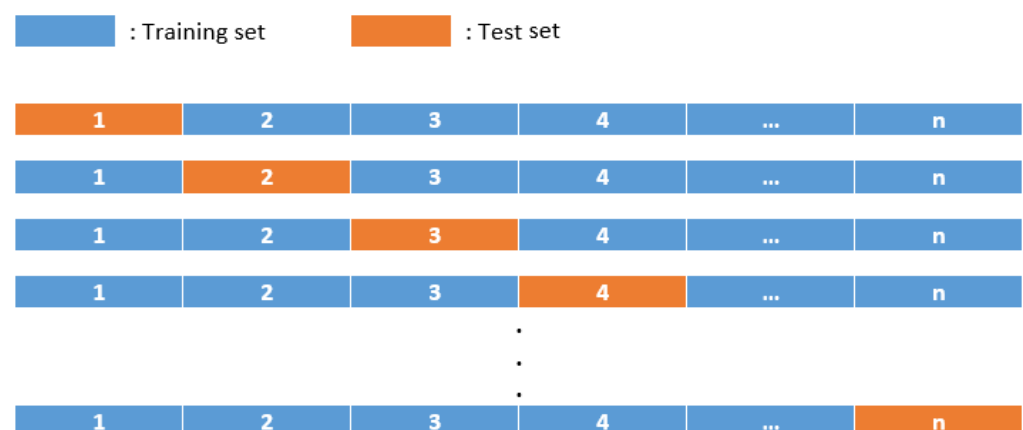


Figure 1. Behavior of the leave-one-out strategy.

2.3. Classification Models

In order to solve the classification problem, several models were used in order to compare the results and discover which was most appropriate for the problem at hand. This research focused on models that are both easy to understand and offer highly accurate results.

In order to determine which explainable models to use, the main criterion that was considered for the present study was the output of the model, so that all models provided rules or trees that justified their results and that could be directly linked to the biomarkers that influenced them, thus offering evidence of what biomarkers can provide relevant information that can be easily understood by doctors, potentially aiding them when diagnosing patients. Within these two categories, models that presented highly accurate results (as described in Section 3.1) were selected.

This section includes an explanation of all the models used, as well as the relevant hyperparameters for each model. These hyperparameters were selected due to them providing the most accurate results for each model and were obtained by fine-tuning the models on the training and test datasets.

2.3.1. OneR

One Rule (OneR) [37] is a very simple model that attempts to predict the label of a given dataset by generating one singular rule. It is the natural evolution of the ZeroR model, which simply predicts the majority label. While ZeroR offers no valuable information for predictions, it can be used as a baseline for all the other models, which should achieve higher results.

OneR generates an output, that is, a rule, in two steps. At first, it analyzes all the attributes of the dataset and creates a rule for each of them. Afterwards, it tests every single generated rule to determine which one achieves the highest accuracy. That rule is the one that the model will use for all future predictions.

In spite of its simplicity, OneR can achieve remarkable results on some datasets, particularly those that contain classes that heavily depend on a given attribute or that are relatively simple.

For this specific research, OneR was tested out with a minimum bucket size of 6 for discretizing numerical values, and 4 decimal places were considered.

2.3.2. JRIP

JRIP is the java implementation of the RIPPER [38] algorithm, which generates a set of rules that are then used to predict the labels of later examples. The algorithm creates the ruleset in two stages:

- Building stage: This stage generates rules until the error rate increases beyond a given threshold or no examples remain. A rule is generated by using one attribute, and antecedents are added to it by following a greedy approach [39] until the rule is perfect. Afterwards, rules are incrementally pruned to prevent overfitting.
- Optimization stage: after generating the original ruleset, several variants of each rule are generated and pruned by following the same process as the building stage. One variant is generated by adding more antecedents, and the other is generated from an empty rule. The minimal description length for each original rule and its variants is computed, and the rule that has the lowest description length is selected and added into the final ruleset.

The result is a series of rules that are executed in order and that can accurately describe the behavior of a dataset, which allows for the prediction of items but most importantly allows experts to learn and understand how and why the algorithm labels each example.

In this research, the error threshold was 50%, and two variant rules were generated during the optimization stage.

2.3.3. FURIA

The Fuzzy Unordered Rule Induction Algorithm [40], or FURIA for short, is an algorithm for generating a ruleset that allows for the prediction of labels. It was designed as an extension of JRIP.

The main difference between both algorithms is that the FURIA makes use of fuzzy logic. This implies that rules do not have specific values that need to be met for the antecedent to become relevant; instead, a rule can be triggered with values that are in a certain range with a given confidence factor. It also includes an efficient rule-stretching method to deal with uncovered examples and unordered rulesets.

All of these changes allow the FURIA to improve the results of JRIP in regard to certain problems, particularly for those examples with a large number of attributes or that would require a large ruleset in order to achieve accurate predictions.

As the FURIA is based on JRIP, the hyperparameters used for the FURIA were akin to those used in JRIP. Furthermore, the fuzzy AND operator was defined as the Product T-Norm.

2.3.4. J48

J48 [41] is a decision tree that originated from C4.5 [42]. A J48 tree consists of several nodes which can lead to two or more nodes based on the value of certain attributes of the data. The lowest level nodes, also called leaf nodes, return a given label as the output. Whenever a new example is to be labeled, one must simply follow the nodes of the tree by assessing the values of the attributes of the example.

The tree is built through several iterations of the same algorithm, which takes place in three main stages:

- Stage 1: If all or almost all the elements of a node are of the same class or label, such a node is considered a leaf node, and the algorithm stops for that node.
- Stage 2: Otherwise, for each element in the node, all attributes are considered and tested. The attribute that presents the largest gain in data is selected, and the node is split into several child nodes. Each child node is defined by the threshold values of the selected attribute, so that the items of the parent node are distributed accordingly between all child nodes.
- Stage 3: Once the entire tree has been built by repeating stages 1 and 2, it can be pruned [43] by removing some nodes, thus preventing overfitting and sometimes increasing the accuracy.

J48 has been thoroughly used in many applications and has proven to be able to deal with particular characteristics, lost or missing attribute estimations of the data, and varying attribute costs [44].

The J48 tree used in this research had a confidence factor of 0.25 and a minimum of 2 objects per leaf node. It was also designed not to use the actual value of nodes when it came to splitting them, instead using the value that achieved the highest accuracy. Lastly, during the pruning phase, subtree raising was considered so as to not fully remove them from the resulting tree.

2.3.5. ADTree

The Alternating Decision Tree (ADTree) [45] is a tree-based algorithm. However, unlike decision trees such as J48, the ADTree has two different types of nodes: decision nodes and prediction nodes. Decision nodes work much like nodes in a decision tree would, specifying a singular predicate condition. Prediction nodes, however, have a singular value associated with them.

The ADTree represents each possible label as a numerical value. This way, when an example is examined, it follows the path of the tree, using the decision nodes to decide which prediction nodes are taken into consideration. Once this process is over, the numbers associated with each decision node are added up. The result of this operation is compared to the value of each label. The label with the closest value will be the one predicted. Moreover, this result can double as a confidence factor: if it is very close to the value of the label, it is more likely that the prediction is accurate.

The ADTree has proven to be as robust as boosted decision trees. Although it tends to be more complex than other algorithms that can achieve similar results, several optimizations have been developed to overcome this issue [46].

All of the previously discussed improvements were taken into consideration when designing the ADTree model for this research. Furthermore, the tree was built in 10 boosting iterations, each of which was designed to attempt to expand all possible paths.

2.3.6. Random Forest

The Random Forest model [47] is an ensemble model, and as such, it combines several simpler tree-based models, such as J48, to compute a final result.

What differentiates the Random Forest from other ensemble algorithms is the fact that, for every tree, instead of the full set of attributes, only a fraction of them is considered. Afterwards, the results of each individual tree are combined by using a weighted sum or a majority vote in order to predict the label for a given example.

Since each tree considers different attributes and the information is combined later on, the algorithm becomes more robust, particularly when compared with the individual trees that make up the model. Furthermore, this allows for the partial mitigation of overfitting, as well as any other issues that may arise as a consequence of a high number of attributes.

The Random Forest was generated by using J48 trees, and all previously applicable hyperparameters were still used here. Furthermore, the model was tested by using several different numbers of trees. The best results were achieved by using 20 trees, which were the ones that were considered.

3. Results

3.1. Metrics

As the accuracy may not be sufficient to determine the quality of a model that deals with the diagnosis of cancer, several metrics were used to compare the models to one another. To understand these metrics, a series of basic definitions upon which they are based is presented:

- True Positive (TP): The amount of patients with breast cancer that were labeled as such.
- True Negative (TN): The amount of patients without breast cancer that were labeled as such.
- False Positive (FP): The amount of patients without breast cancer that were labeled as having breast cancer.
- False Negative (FN): The amount of patients with breast cancer that were labeled as healthy.

With these definitions, we can define the following metrics, which were used to measure the performance of our models:

- Accuracy: $(TP + TN) / (TP + FP + TN + FN)$
- Precision: $TP / (TP + FP)$
- Recall: $TP / (TP + FN)$
- F1 Score: $2 \times (Precision \times Recall) / (Precision + Recall)$

3.2. General Results

This section includes the results for all models when attempting to detect if a patient had cancer, which can be found in Table 1. Further subsections show the specific results of each model.

Table 1. Results for each model when diagnosing a patient.

	Accuracy (%)	Precision	Recall	F1
OneR	98.204 ± 0.134	0.991	0.982	0.987
JRIP	98.802 ± 0.110	0.991	0.991	0.991
FURIA	98.802 ± 0.111	0.991	0.991	0.991
J48	98.204 ± 0.134	0.991	0.982	0.987
ADTree	98.802 ± 0.110	0.991	0.991	0.991
Random Forest	99.401 ± 0.103	0.991	1.000	0.996

3.2.1. OneR

OneR achieved an accuracy of 98.204% by using one singular rule, which implies that the attributes in it could be extremely relevant when attempting to diagnose breast cancer. The rule it generated can be found in Table 2.

Table 2. Rule generated by OneR.

				Precision	Recall
IF	Oxytocin < 35.75	THEN	diagnosis = BreastCancer	0.991	0.982

3.2.2. JRIP

JRIP generated a ruleset that achieved an accuracy of 98.802% when diagnosing breast cancer. This ruleset can be found in Table 3.

It should be noted that JRIP only generated one rule, which was the same as the one generated by OneR except for some tweaking of specific values, further proving that oxytocin is a key value to consider when diagnosing breast cancer.

Table 3. Rule generated by JRIP.

				Precision	Recall
IF	Oxytocin < 36.7	THEN	diagnosis = BreastCancer	0.991	0.991

3.2.3. FURIA

Unlike the previous algorithms, the FURIA applies fuzzy logic and a confidence factor (CF) to determine how likely a predicted diagnosis is based on the specific values of the data it uses. For each attribute, the entire range of values was divided evenly across five labels: very low, low, medium, high, and very high. Each label was defined by four values; if the example value was between the first and second values or between the third and fourth values, the chosen label was not certain, and the confidence factor was considered. If no other labels were more likely, that label was used. If it was between the second and third values, the label was certain, acting much like those of the previous algorithms. These rules, which achieved an accuracy of 98.802%, are described in Table 4.

The relevant values for the labels are presented hereunder:

- Oxytocin: Very low oxytocin levels were defined by $[-\text{Inf}, -\text{Inf}, 34.8, 36.7]$. Very high oxytocin levels were defined by $[31.6, 36.7, \text{inf}, \text{inf}]$.

- IRAP: Very low levels were defined by $[-inf, -inf, 360, 363.859944]$.

Table 4. Rules generated by FURIA.

					Precision	Recall
IF	Oxytocin IS very low	THEN	diagnosis = BreastCancer (CF = 0.99)		0.991	0.991
IF	Oxytocin IS very high AND IRAP IS very low	THEN	diagnosis = Control (CF = 0.98)		0.982	0.982

3.2.4. J48

Unlike the previous algorithms, J48 generates a tree which must be followed in order to determine the diagnosis of a given patient. However, such a tree can be easily transformed into an algorithmic set of rules, which can be found in Table 5.

It should be noted that, once more, J48 generated one single rule while achieving an accuracy of 98.204%. However, this time a different attribute was used.

Table 5. Rules generated by J48.

				Precision	Recall
IF	GnRH < 17.5	THEN	diagnosis = BreastCancer	0.991	0.982

3.2.5. ADTree

The ADTree, much like the previous algorithm, also generates a tree. However, in this case there are two kinds of nodes: decision nodes, to determine which branch of the tree is followed, and prediction nodes, which modify the value of the diagnosis score. Based on the final score, a final diagnosis is provided. The ADTree achieved an accuracy of 98.802%. The results can be represented as a ruleset, which is shown in Table 6. It should be noted that a base value for the score was provided, so that *defaultScore* = -0.545 .

Table 6. Rules generated by ADTree.

IF	oxytocin < 35.75	THEN	score = score − 2.258
IF	oxytocin >= 35.75	THEN	score = score + 2.012
IF	oxytocin < 85.9	THEN	score = score − 1.151
IF	oxytocin >= 85.9	THEN	score = score + 1.185
IF	APB < 133.396	THEN	score = score + 0.684
IF	APB >= 133.396	THEN	score = score − 0.711
IF	β -Endorphin < 133.396	THEN	score = score − 0.742
IF	β -Endorphin >= 133.396	THEN	score = score + 0.858
IF	GnRH < 20.6	THEN	score = score − 0.303
IF	GnRH >= 20.6	THEN	score = score + 0.378
IF	5'-Nucleotidase < 1.45	THEN	score + 0.26
IF	5'-Nucleotidase >= 1.45	THEN	score = score − 0.224
IF	APB < 114.584	THEN	score = score − 0.179
IF	APB >= 114.584	THEN	score = score − 0.177
IF	score < 0	THEN	diagnosis = BreastCancer
ELSE	diagnosis = Control		

3.2.6. Random Forest

The Random Forest created a total of 25 J48 trees in order to analyze the data and diagnose breast cancer in patients. By using these 25 trees, it achieved an accuracy of 99.401%. By using only the five most relevant trees, the accuracy was over 98%. Those five trees can be found in Tables 7, 8, 9, 10, and 11, respectively.

It should be noted that the model correctly diagnosed all cancer patients, as shown by the perfect score for its recall.

Table 7. First tree generated by Random Forest.

				Precision	Recall
IF	Vasopressin < 2.68 AND β -Endorphin < 23.3 AND cholinesterase < 7186	THEN	diagnosis = BreastCancer	1.000	0.223
IF	Vasopressin < 2.68 AND β -Endorphin < 23.3 AND cholinesterase $\geq$ 7186	THEN	diagnosis = control	1.000	0.055
IF	Vasopressin < 2.68 AND β -Endorphin $\geq$ 23.3	THEN	diagnosis = control	0.975	0.710
IF	Vasopressin $\geq$ 2.68 AND GnRH < 22.48	THEN	diagnosis = BreastCancer	0.989	0.830
IF	Vasopressin $\geq$ 2.68 AND GnRH $\geq$ 22.48	THEN	diagnosis = Control	1.000	0.127

Table 8. Second tree generated by Random Forest.

				Precision	Recall
IF	GnRH < 20.6 AND Cholinesterase < 8342.5	THEN	diagnosis = BreastCancer	0.991	1.000
IF	GnRH < 20.6 AND Cholinesterase $\geq$ 8342.5	THEN	diagnosis = Control	1.000	0.019
IF	GnRH $\geq$ 20.6	THEN	diagnosis = Control	0.981	0.964

Table 9. Third tree generated by Random Forest.

				Precision	Recall
IF	β -Endorphin < 32.55 AND Protein < 6.79	THEN	diagnosis = BreastCancer	0.984	0.580
IF	β -Endorphin < 32.55 AND Protein $\geq$ 6.79 AND Progesterone < 2.15 AND BMI < 28.71 AND oxytocin < 104.24	THEN	diagnosis = BreastCancer	1.000	0.054
IF	β -Endorphin < 32.55 AND Protein $\geq$ 6.79 AND Progesterone < 2.15 AND BMI < 28.71 AND oxytocin $\geq$ 104.24	THEN	diagnosis = Control	1.000	0.109
IF	β -Endorphin < 32.55 AND Protein $\geq$ 6.79 AND Progesterone < 2.15 AND BMI $\geq$ 28.71	THEN	diagnosis = BreastCancer	1.000	0.348
IF	β -Endorphin < 32.55 AND Protein $\geq$ 6.79 AND Progesterone $\geq$ 2.15	THEN	diagnosis = Control	1.000	0.145
IF	β -Endorphin $\geq$ 32.55	THEN	diagnosis = Control	0.980	0.890

Table 10. Fourth tree generated by Random Forest.

				Precision	Recall
IF	Cholinesterase < 5174	THEN	diagnosis = BreastCancer	0.987	0.688
IF	Cholinesterase $\geq$ 5174 AND oxytocin < 35.75	THEN	diagnosis = BreastCancer	1.000	0.277
IF	Cholinesterase $\geq$ 5174 AND oxytocin $\geq$ 35.75	THEN	diagnosis = control	0.982	0.982

Table 11. Fifth tree generated by Random Forest.

				Precision	Recall
IF	β -Endorphin < 28.2 AND Iron < 78.87	THEN	diagnosis = BreastCancer	0.987	0.679
IF	β -Endorphin < 28.2 AND Iron $\geq$ 78.87 AND IRAP < 224.96	THEN	diagnosis = BreastCancer	1.000	0.036
IF	β -Endorphin < 28.2 AND Iron $\geq$ 78.87 AND 224.96 < IRAP < 264.45	THEN	diagnosis = Control	1.000	0.109
IF	β -Endorphin < 28.2 AND Iron $\geq$ 78.87 AND IRAP $\geq$ 264.45	THEN	diagnosis = BreastCancer	1.000	0.303
IF	β -Endorphin $\geq$ 28.2	THEN	diagnosis = Control	0.980	0.909

4. Discussion

Though deep learning and AI techniques have been used for the detection of breast cancer in several recent articles in the literature [23,24,48], most of them are fully focused on the prediction of the cancer in itself and less so on the information that can be extracted during the process.

As such, many of the models in the literature oftentimes make use of structures such as autoencoders (AEs) for anomaly detection in images or convolutional neural networks (CNNs) for the analysis of screening mammograms. In [49,50], AEs were used to great effect, achieving an accuracy of roughly 98.5%. In [51], CNNs were used to predict breast cancer using ultrasound images, achieving an accuracy of up to 94%. Similarly, in [52], a CNN model achieved 95.7% accuracy on mammography images. Lastly, in [53], a Quantum Hybrid CNN achieved results of 97.661%.

Though these models can achieve highly accurate results on their own, they lack explainability, since they are black-box models that offer no reasons as to why they predict a given label. Though in many applications of these models this is not an issue, when it comes to healthcare, it is of the utmost importance that experts make informed decisions based on data that back them up. As such, using explainable and trustworthy methods, even if their results are slightly less accurate than those provided by more advanced models, may be preferable. Several studies have made use of machine learning models, such as the Random Forest, to predict breast cancer: in [54], a perfect result was achieved using the Random Forest to diagnose breast cancer according to cell attributes extracted from images, while in [55], a result of 93% was achieved when predicting whether a patient was at risk of developing breast cancer, but the results were based on data gathered from images as well.

Though many other articles have previously used machine learning techniques to detect and prevent breast cancer in the literature [56–60], their results are based on the analysis of images or the exclusive use of medical parameters such as symptoms, weight, and height, among others. Furthermore, whilst several of the proposed models make use of machine learning to predict breast cancer or its recurrence risk, none of them make use of biomarkers and, while the most relevant factors for predictions are sometimes mentioned, the relation and interactions between these factors and breast cancer or its recurrence are hardly explained. This manuscript not only uses biomarkers as a key indicator to predict breast cancer, but also properly defines the given interactions between different biomarkers and the specific threshold values that could be relevant to researchers attempting to diagnose breast cancer.

As can be seen, most of the literature revolves around the use of mammographies and symptom data and extracting information from them, while clinical biomarkers are seldom used. Regardless, the accuracy provided by all the models in this contribution, particularly the Random Forest, which achieved an accuracy of 99.401%, rivals that of some of the state-of-the-art literature while additionally providing valuable information on biomarkers that can guide researchers when it comes to predicting breast cancer.

The used models also provide another advantage: they are easy to understand once generated, and as long as the data that were used to train them are not biased, their results will also be trustworthy. Once a model has generated a ruleset, all that is needed is to check all the generated rules one by one to determine which one is to be used whilst considering biomarker data from a new patient who is to be diagnosed. This prevents biases and limitations of the models when interpreting the results of the models if they have been properly trained.

Nevertheless, these models are not without their own limitations: they require large amounts of quality information in order to properly define the interactions between biomarkers and their potential effect on breast cancer. If data for a specific patient lack

values for biomarkers, the models may offer the wrong predictions. Furthermore, model training also has to be carefully revised to avoid biases and overfitting, which may cause models to perfectly predict the training set but offer predictions that are lacking when used outside of it. However, as long as these limitations are taken into consideration and fixed by using pre-processing techniques such as balancing the datasets and leave-one-out validation, the models are able to provide valuable information and accurate results.

In this regard, several important parameters related to breast cancer pathogenesis appear to be involved in several models described here. These parameters can be seen in Table 12 and include peptide hormones, their regulatory proteolytic enzymes, steroid hormones, enzyme activities, and the BMI, among others, being the basis for the explainability of the different models. Of these, oxytocin appeared in all the proposed models except J48.

The following sections provide a thorough explanation of these parameters and their relevance when predicting breast cancer according to the literature.

Table 12. Parameters deemed relevant for the prediction of breast cancer for each tested model.

Parameter	OneR	JRIP	FURIA	J48	ADTree	Random Forest
Hormones	Oxytocin	Oxytocin	Oxytocin	GnRH	Oxytocin, GnRH, β -endorphin	Oxytocin, GnRH, β -endorphin, vasopressin, progesterone
Hormone-regulating proteolytic enzymes			IRAP		APB	IRAP
Other enzyme activities					5'-Nucleotidase	Cholinesterase
Metal ions						Iron
Other						Total protein, BMI

IRAP: insulin-regulated aminopeptidase; GnRH: gonadotrophin-releasing hormone; APB: aminopeptidase B; BMI: body mass index.

4.1. Peptide Hormones

Oxytocin, a neuropeptide hormone known for its roles in reproduction and social behavior, has emerged as a potentially influential agent in breast cancer development and progression. Its role in this disease, however, is characterized by remarkable complexity, manifesting both proliferative and antiproliferative effects depending on an intricate network of factors including the cell type, hormone concentration, duration of treatment, and interactions with the tumor microenvironment [61–67]. In fact, oxytocin can inhibit breast cancer cell proliferation through the activation of its receptor (OTR), a member of the G protein-coupled receptor family, triggering an intracellular signaling cascade that can lead to apoptosis, cell cycle arrest, and the suppression of pro-tumor pathways such as PI3K/Akt and ERK [68–73]. In vivo studies using animal models of breast cancer have corroborated these findings, demonstrating a reduction in the tumor growth and volume following oxytocin administration [63,64,68,73]. On the other hand, the ability of oxytocin to modulate estrogen receptor (ER) expression and function adds another layer of complexity to its interaction with breast cancer [74,75]. ERs, which play a crucial role in breast cell growth and development, are also important therapeutic targets in breast cancer treatment. Thus, oxytocin, through the modulation of ER expression and activity, could influence the response of tumor cells to hormonal therapies [74,76]. The impact of oxytocin on the expression of microRNAs (miRNAs) has also been the subject of research. miRNAs, small non-coding RNA molecules that regulate gene expression, have emerged as important regulators in cancer. Studies have suggested that oxytocin may influence the expression of specific miRNAs, such as miR-195, which are involved in suppressing tumor growth [73,77]. Despite evidence pointing towards a potential anti-tumor effect of oxytocin, several studies have suggested that this peptide hormone, under certain conditions, may

promote breast cancer cell proliferation. The stimulation of angiogenesis, the formation of new blood vessels that nourish the tumor, is one of the mechanisms proposed to explain this effect [78]. In the same way, the overexpression of OTRs, observed in some studies, has been associated with an increase in the metastatic potential of triple-negative breast cancer cells [64,79]. The discrepancy in the results obtained in different studies highlights the importance of context in assessing the effects of OT in breast cancer and could also explain the influence of oxytocin levels in the rules generated in the different models.

GnRH also appeared as an explainable hormonal factor in the J48, ADtree, and Random Forest models. This peptide hormone, produced in the hypothalamus, stimulates the release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) by the pituitary gland, which in turn induces steroid production in the gonads, playing a key role in the regulation of the reproductive system but also enhancing the growth and progression of estrogen-dependent tumors, as occurs with breast cancer [80,81]. In fact, GnRH analogs are used to suppress ovarian function based on the premise that reducing the circulating estrogen levels can inhibit hormone-dependent tumor growth [82]. It has been recently described that, in premenopausal women with hormone receptor-positive (HR+) breast cancer, ovarian function suppression using GnRH analogs such as goserelin and leuprolide has become an important therapeutic strategy [82]. A meta-analysis of 25 randomized clinical trials, involving nearly 15,000 premenopausal women with HR+ breast cancer, also demonstrated that ovarian ablation or ovarian function suppression using GnRH significantly reduced the 15-year risk of recurrence and mortality [82]. Finally, in addition to the role of GnRH in suppressing ovarian function, there has been research on the use of GnRH analogs as vectors for the targeted delivery of chemotherapeutic drugs to tumor cells. Preclinical studies have shown that the conjugation of GnRH analogs to cytotoxic agents, such as doxorubicin, can increase the treatment efficacy and reduce systemic toxicity [83,84].

The ADtree and Random Forest models also included β -endorphin as an explainable parameter. Neurobehavioural stress can influence the growth and progression of various types of cancer, including breast cancer [35,85]. This is due to the ability of psychological factors to alter the functioning of the immune and endocrine systems, which can affect the tumor growth and spread. Indeed, stress relief in cancer patients promotes faster recovery and a sense of well-being in those receiving radiotherapy or chemotherapy [35,85]. Within this context, opioid peptides, and in particular β -endorphin, have emerged as possible mediators of immune alterations that may contribute to tumor development [35]. β -endorphin plays a crucial role in the bidirectional connection between the immune and neuroendocrine systems, which could explain the effects of stress on the immune capacity against cancer [35,86]. Several studies have shown that healthy women have elevated levels of β -endorphin, with these being even higher in postmenopausal women. However, in women with breast cancer, significantly lower levels are observed, with no differences between premenopausal and postmenopausal women [35]. These findings suggest that decreased levels of β -endorphin may be associated with the development of breast cancer. Neoadjuvant chemotherapy, a common treatment for breast cancer, only improves β -endorphin levels in postmenopausal women, without reaching the levels of healthy women. Furthermore, both premenopausal and postmenopausal women treated with neoadjuvant chemotherapy maintain elevated cortisol levels, indicating a persistent stressful situation [35]. Our findings underscore the significance of our explainable artificial intelligence models in elucidating the critical role of β -endorphin modulation in breast cancer patients to the medical and research community. They could not only facilitate the potential therapeutic implications of β -endorphin regulation, thereby informing clinicians and researchers about novel avenues for intervention in breast cancer management, but could also enhance the understanding of the neuroendocrine basis of breast cancer,

potentially guiding future research directions and clinical strategies. In fact, Sánchez et al. (2023) have pointed out that the regulation of stress levels by modulation with β -endorphin could be an alternative pharmacological therapy against tumor growth and development.

Vasopressin was other explainable parameter found in the Random Forest model. Also known as antidiuretic hormone, it has been studied in breast cancer since the presence of vasopressin receptors (V1a, V1b, and V2) in breast cancer cell lines, including MCF-7, were demonstrated [87,88]. Although it has traditionally been attributed a proliferative effect, recent research has suggested that its influence in breast cancer may be more complex and dependent on factors such as the dose, receptor subtype, and cellular context. Thus, a recent study by Alkafaas et al. [89] explored the effects of vasopressin on luminal A breast cancer cells, using a concentration of 100 nM. Their results suggested an anti-proliferative and anti-metastatic effect of vasopressin, with increased apoptosis, the increased expression of apoptosis markers such as Bax and Caspase-3, and decreased cell invasion. In addition, the increased expression of the autophagy marker LC3II and decreased Akt protein activation were found, suggesting the involvement of these pathways in the response to AVP. Garona et al. [90,91] investigated the desmopressin analog [V4Q5]dDAVP, a selective V2 receptor agonist, in preclinical models of breast cancer. They found that this analog inhibited tumor growth, angiogenesis, and metastasis in in vitro and in vivo models. Their findings in MDA-MB-231 xenografts showed reduced tumor growth and angiogenesis with its administration, while in immunocompetent mice with F3II mammary tumors, the complete inhibition of metastatic progression was observed.

4.2. Peptide Hormone-Regulating Proteolytic Enzymes

Apart from the previously described peptide hormones, our FURIA, ADTree, and Random Forest models also included the peptide hormone-regulating proteolytic enzymes insulin-regulated aminopeptidase (IRAP) and aminopeptidase B (APB). IRAP, also known as oxytocinase, is a zinc metallopeptidase with important regulatory functions. IRAP is the only known enzyme that cleaves oxytocin and vasopressin in vivo [92,93], two important peptide hormones also included in our proposed models. In addition to its role in hormone regulation, IRAP is also the high-affinity binding site for angiotensin IV (AngIV) type 4 (AT4) receptor ligands [94,95] and is linked to insulin-dependent glucose transporters via the translocation of glucose transporter type 4 (GLUT4) [96]. Previous studies have shown an association between the IRAP activity and the number and size of mammary tumors in an animal model of breast cancer [97,98]. A highly significant increase in IRAP activity has also been found in breast cancer tissue from female patients [99]. Another study by Ramírez-Expósito et al. [31] analyzed the IRAP activity in premenopausal and postmenopausal women with breast cancer, whether or not they had been treated with neoadjuvant chemotherapy. The results showed significant changes in the specific activity of circulating IRAP, with variations dependent on the hormonal status and chemotherapy administration. Thus, IRAP is involved in mechanisms related not only to the regulation of oxytocin and/or vasopressin, but also to the local mammary renin–angiotensin system (RAS) through AngIV and its role in glucose transport through the IRAP/GLUT4 system. The local RAS has been implicated in several of the hallmarks of cancer, including the processes of tumor cell survival, proliferation, and dissemination. The best known RAS signaling is mediated by angiotensin II (AngII) through its type 1 receptor (AT1), which increases cell proliferation and modulates vascular cell growth during angiogenesis [100]. The study by Ramírez-Expósito et al. [31,101] suggested an increase in the circulating levels of AngIV, a substrate of IRAP. AT4 receptor ligands dose-dependently inhibit the catalytic activity of IRAP [102]. Therefore, the changes in the IRAP-specific activity observed in women with breast cancer may be related to local RAS dysregulation. Additionally,

mammalian cells use glucose as a major source of energy production. The presence of GLUT4 has been demonstrated in insulin-independent tissues, such as the mammary gland, and also in breast cancer cells, where it plays a role in cell cycle progression [103,104]. IRAP and GLUT4 colocalize in specialized vesicles that translocate to the cell membrane in response to stimuli such as insulin. GLUT4 inhibition in breast cancer cells has been shown to inhibit cell proliferation and decrease cell viability [105]. The changes in IRAP activity observed in [31] could affect the availability of IRAP and, consequently, the translocation of GLUT4 to the cell membrane, which would influence glucose metabolism in tumor cells.

Aminopeptidase B (APB) has also been studied in relation to breast cancer because of its role in the RAS [106]. Thus, APB activity has shown significant patterns of alteration [107]. Martinez et al. [99] found that the membrane-bound APB activity was significantly elevated in neoplastic tissue compared to unaffected tissue, suggesting that APB may play an active role in the tumor microenvironment. Furthermore, in a study by Carrera et al. [108], a decrease in soluble APB was observed in mammary tissue from rats with N-methyl-nitrosourea (NMU)-induced tumors, suggesting that the subcellular location of this enzyme could be an important factor in its influence on a tumor. APB is an enzyme involved in the conversion of angiotensin III (AngIII) to AngIV [108]. Despite the important role of Ang II in cell proliferation, angiogenesis, and tumor growth [106], the metabolism of Ang II, via RAS regulatory enzymes such as APB, is thought to affect tumor progression. Ruíz-Sanjuan et al. [106] suggested that elevated levels of APB may indicate an increased conversion of AngIII to Ang IV, a peptide that binds to AT4 receptors and is also associated with IRAP activity, as previously described. In a two-year follow-up study, Ramirez-Exposito et al. [28] found that APB levels remained significantly unchanged in women with breast cancer. However, a small increase in the APB levels was observed in premenopausal women treated with chemotherapy but was not sustained afterwards. These data suggest that the initial changes in APB activity, associated with tumor development or the effect of chemotherapy, are largely stable over time. Therefore, APB may influence breast cancer development through its role in angiotensin regulation, which may alter cell proliferation and angiogenesis in the tumor microenvironment.

4.3. Progesterone

The steroid hormone progesterone was another explainable variable proposed by the Random Forest model. The relationship between progesterone and breast cancer is complex and has been the subject of numerous studies for decades, generating controversy as to whether progesterone is a stimulating or inhibitory factor in the development of this disease [109,110]. Initially, the presence of progesterone receptors (PRs) was recognized as a marker of a response to endocrine therapies and as an indicator of the presence of functional estrogen receptors [109,110]. However, subsequent research has revealed that the action of progesterone can vary depending on factors such as the dose, the type of progestin used, the presence or absence of estrogens, and the clinical context [110,111]. Progesterone exerts its effects through its progesterone receptors (PRs), which are transcription factors that regulate various physiological processes in mammary cells [110]. PRs are found in two isoforms, A and B, which differ in their structure and function [112,113]. These isoforms can operate through different signaling pathways in target cells [112,113]. The activity of PRs is also modulated by the binding of cofactors and by post-translational modifications, such as phosphorylation [113,114]. Progesterone plays a crucial role in the regulation of the cell cycle and proliferation in mammary tissue [111]. It has been observed that progesterone can induce cell proliferation through the activation of the RANKL and WNT4 pathways [110,111,115–119]. In addition, it has been proposed that progesterone can influence the expansion of cancer stem cells, which could contribute to

tumor progression [110,115,118,119]. However, there is also evidence that progesterone can have anti-proliferative effects in certain circumstances [120]. The interaction between progesterone and estrogen is fundamental in the context of breast cancer. PRs are target genes regulated by estrogen receptors (ERs) [110,116]. PRs can modulate ER activity, which can have both agonist and antagonist effects [121]. Some studies have suggested that the presence of PRs in ER+ breast tumors is associated with a better response to endocrine therapy [122]. However, the loss of PRs in these tumors has been linked to resistance to treatments such as tamoxifen [123–125]. The relationship between endogenous progesterone and the breast cancer risk is not entirely clear [111]. Some studies suggest that elevated levels of progesterone during the luteal phase of the menstrual cycle may increase the risk of breast cancer, especially in women carrying BRCA mutations [111]. It has also been proposed that cumulative exposure to progesterone throughout a woman's reproductive life could influence the risk [111]. Therefore, progesterone exerts a complex role in the development of breast cancer, with actions that can be both promoting and inhibitory, depending on the cellular and molecular context [110,111].

4.4. Other Enzyme Activities

The ADTree model included the 5'-nucleotidase activity as an explainable variable, whereas the Random Forest included the cholinesterase activity as an explainable variable. Ecto-5'-nucleotidase (CD73) is emerging as an enzyme of great interest due to its involvement in the regulation of the tumor microenvironment and its influence on disease progression [126–129]. This enzyme, located on the cell surface, catalyzes the hydrolysis of 5'-AMP to adenosine and inorganic phosphate [128,130]. Adenosine, the product of this reaction, acts as a potent immunosuppressor in the tumor microenvironment, promoting immune evasion and tumor growth [127,129,131,132]. Its presence in breast cancer is correlated with a worse prognosis, a higher risk of metastasis, and resistance to chemotherapy [127,133]. In particular, in TNBC, a high expression of CD73 has been observed, contributing to treatment resistance and tumor progression [127,134]. Also, CD73 participates in the promotion of tumor angiogenesis, tumor growth, and metastasis [135,136]. On the other hand, some studies have suggested the involvement of cholinesterases in mammary tumorigenesis. In fact, Ruiz-Espejo et al. [137] found that acetylcholinesterase activity doubled in breast tumor tissue compared to in normal tissue. Additionally, the glycosylation of acetylcholinesterase was affected in breast cancer [137]. Bernardi et al. [138] observed a high frequency of amplifications in the gene encoding this enzyme in sporadic breast tumors, which could be responsible for its increased expression. Furthermore, the promoter of the acetylcholinesterase gene contains binding sites for estrogen receptors [139], and the abundance of estrogen receptors in mammary tissue also suggests that these could be involved in regulating acetylcholinesterase expression. Other factors, such as the stimulation of muscarinic receptors, can induce the synthesis of transcription factors that activate the transcription of the acetylcholinesterase gene [140]. Additionally, stimulation by cytokines like interleukin-1 β (IL-1 β), which is actively synthesized in breast cancer [141], could increase the expression of acetylcholinesterase mRNA and, consequently, its enzymatic activity [142]. Therefore, tumor-specific genetic and biological alterations could influence acetylcholinesterase activity, or its alteration may be a consequence of neoplastic transformations that contribute to the maintenance of the tumor process [138].

4.5. Iron

The Random Forest model also proposed serum-circulating iron as an explainable variable for breast cancer diagnosis. Iron is an essential micronutrient for the human body, playing a crucial role in various physiological processes such as oxygen transport, immune

function, and energy production [143]. However, excess iron can generate oxidative stress and promote the activation of oncogenes, making it a potentially implicated factor in cancer development [144,145]. Previous epidemiological studies have yielded conflicting results on the association between the serum iron status and breast cancer risk [146]. Some studies have reported a positive association, while others have found no significant relationship [147–152]. This inconsistency in results may be attributed to limitations inherent to observational studies, such as the presence of confounding factors and the possibility of reverse causality [153]. To address these limitations, Hou et al. [146] conducted a two-sample Mendelian randomization study to explore the causal relationship between the serum iron status and breast cancer risk. This method used genetic variants as instruments to estimate the causal effect of exposure on an outcome, minimizing bias from confounding factors [154,155]. The results of this study suggested that serum transferrin levels might be positively associated with the risk of estrogen receptor-negative (ER-negative) breast cancer, although no significant effect was observed on the risk of breast cancer in general or its ER-positive subtypes [146]. In another study, Hua et al. [156] retrospectively analyzed the prognostic value of the serum iron level in women with early-stage triple-negative breast cancer (TNBC). They divided patients into two groups based on a cut-off point for the serum iron level and found that those with lower iron levels had significantly longer disease-free survival and overall survival compared to patients with higher levels. These findings suggest that the serum iron level could be a potential prognostic biomarker in patients with early-stage TNBC. It is important to note that most studies assess the iron status in the body by measuring the circulating levels of iron-binding proteins, such as transferrin and ferritin [157,158]. However, the direct measurement of serum iron might provide a more accurate assessment of the iron status in the body [150–152,159].

4.6. Body Mass Index

The BMI was another of the explainable variables proposed by the Random Forest model. One of the primary mechanisms by which the BMI influences breast cancer outcomes is through its impact on the mammographic imaging quality. Increased adipose tissue in the breast can significantly affect the sensitivity and specificity of mammography, potentially leading to delayed or missed diagnoses. Calvo et al. [160] have described significant differences between the BMI and mammographic density, with a higher BMI in patients with type A density (“fatty breast”) and a lower BMI in patients with type D density (“dense breast”). These differences were more relevant in premenopausal women than in postmenopausal women, although no association was found between the mammographic density and the molecular subtype of cancer [161]. On the other hand, Cao et al. [162] found a positive association between the BMI and breast cancer risk. Women with obesity (BMI ≥ 28 kg/m²) had a 2.09-fold higher risk of developing breast cancer compared to women with a normal weight (BMI of 18.5–23.9 kg/m²). A family history of cancer also increased the risk of breast cancer, and an additive interaction was observed between the BMI and family history. In other study, Jiang et al. [163] analyzed the association between the BMI and breast cancer risk in a large cohort of Chinese women, confirming the association with an increased risk of breast cancer, especially in postmenopausal women. Wada et al. [164] also found that in postmenopausal women, a high BMI was significantly associated with an increased risk of breast cancer, but no significant association was observed in premenopausal women. Cubelos-Fernand et al. [165] used the so-called “Clínica Universidad de Navarra Body Fatness Estimator (CUN-BAE)” to assess the relationship between body fat and postmenopausal breast cancer. An increased risk of breast cancer was also observed in women with a higher BMI and CUN-BAE score, with covariates such as the age at recruitment, age at menarche, number of children, and breastfeeding duration

also appearing in the multivariate model. Holm et al. [166] investigated the relationship between C-reactive protein (CRP) quartiles and the disease-free survival and overall survival in patients with stage I-III breast cancer. A high CRP level was found to be associated with worse survival in all BMI groups, but the association was more pronounced in overweight and obese women. Jiang et al. [163] focused their study on the relationship between the BMI and the incidence of chronic arm lymphedema (BCRL) after breast cancer surgery. Patients with a BMI ≥ 30 kg/m² had a significantly higher rate of BCRL at 1, 2, and 3 years compared with patients with a BMI < 25 kg/m². Ogiya et al. [167] also investigated the incidence of lymphedema after breast cancer surgery and its relationship with the BMI. A BMI ≥ 25 kg/m² was found to be significantly associated with an increased risk of lymphedema at 1 year after surgery. Quartuccio et al. [168] found that the BMI appears to correlate with an increased recurrence rate, especially in the breast, and increased glucose uptake in postmenopausal patients with recurrent breast cancer. Therefore, a complex relationship also exists between the BMI and breast cancer, with a high BMI, especially in postmenopausal women, increasing the risk of breast cancer development, an important variable that must also be taken into account [169].

4.7. Total Protein

Very few studies have investigated the relationship between the total serum protein levels and breast cancer. One study revealed that breast cancer patients had significantly higher total serum protein levels compared to the control group [170]. This finding is in agreement with previous studies that have explored the assessment of the total serum protein in patients with brain tumors [171] and with significant increases in the serum protein levels in patients with lung cancer [172]. The increase in the total serum protein concentration may be due to the fact that the total serum protein is composed of albumin and other proteins, collectively called globulins. It is known that the serum albumin concentration can change under oxidative stress, such as that associated with cancer [173]. Furthermore, as plasma circulates through tissues, it picks up proteins released due to physiological events such as tissue remodeling, trauma, and cell death, leading to an increase in the total serum protein.

5. Conclusions

This research provides a solid foundation for future clinical research and practical applications, directly challenging conventional approaches to breast cancer diagnosis. The use of XAI models and the identification of relevant serum biomarkers offer a new perspective for early detection, understanding pathogenesis, and developing more effective and personalized therapies. The results not only enrich the scientific knowledge but also open a wide array of possibilities for transforming breast cancer management. They include new approaches to early detection, proposing an alternative to traditional imaging techniques, such as mammography, by focusing on the analysis of serum biomarkers and providing new targets for research and the development of more precise and less invasive diagnostic tools.

Secondly, they provide a deeper understanding of pathogenesis due to the ability of XAI models to generate rules and decision trees that allow researchers to better understand the underlying biological mechanisms in breast cancer development. For example, the identification of oxytocin as a primary predictive factor across several models suggests that this hormone plays a more complex and multifaceted role in pathogenesis than previously thought, which could impact the development of targeted therapies. The detection of complex interactions among hormonal, enzymatic, and metabolic factors opens new avenues for understanding the molecular alterations that drive the disease.

Thirdly, this study lays the groundwork for the design of more individualized treatments by providing information on the specific biomarkers relevant to a patient's breast cancer risk, an approach that could be especially valuable in managing breast cancers with high heterogeneity and resistance to conventional therapies. Ultimately, this research proposes a transformation in breast cancer management by promoting a more integrated and personalized approach, based on a deeper understanding of the disease and the use of more precise diagnostic tools. The adoption of XAI models could lead to earlier detection, which translates to better survival rates and a reduction in the use of more invasive and costly treatments. Moreover, the models' capacity to generate explainable information promotes greater confidence in the diagnostic process.

Lastly, the practical challenges of implementing these models in real-world clinical settings should be addressed. All of the used models are highly efficient and simple to implement and can be run even on low-end computers. For these models to be used in a real-world clinical setting, all that would be required would be to store the trained models on a computer, gather all relevant biomarker data from a given patient who has to be diagnosed, and follow the rules generated by a given model. Since several models have been proposed, the data could be run through any of them. This additionally prevents biases and limitations: if all models predict that the patient is likely to develop breast cancer, the chance of a misdiagnosis severely decreases.

Author Contributions: Conceptualization, J.M.M.-R., C.C., M.J.R.-E. and J.M.M.-M.; methodology, J.M.M.-R. and C.C.; software, J.M.M.-R. and C.C.; validation, M.J.R.-E. and J.M.M.-M.; formal analysis, J.M.M.-R.; investigation, J.M.M.-R. and J.M.M.-M.; resources, C.C.; data curation, M.J.R.-E. and J.M.M.-M.; writing—original draft preparation, J.M.M.-R. and J.M.M.-M.; writing—review and editing, J.M.M.-R., C.C., M.J.R.-E. and J.M.M.-M.; supervision, J.M.M.-R., C.C., M.J.R.-E. and J.M.M.-M.; project administration, C.C.; funding acquisition, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financed by the Ministry of Science, Innovation and Universities, Spain, with the code PID2023-149511OB-I00.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the local Clinical Research Ethical Committee at the University Hospital of Jaén, as stated in the corresponding references.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data are unavailable due to privacy or ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ADTree	Alternating Decision Tree
AI	Artificial intelligence
ALAT	Alanine aminotransferase
AngII	Angiotensin II
AngIII	Angiotensin III
AngIV	Angiotensin IV
APA	Aminopeptidase A
APB	Aminopeptidase B
APN	Aminopeptidase N

ASAP	Aspartyl aminopeptidase
ASAT	Aspartate aminotransferase
AT1	Angiotensin II type 1 receptor
BCRL	Breast cancer-related lymphedema
CF	Confidence factor
CKMB	Creatine kinase muscle–brain isoform
CNN	Convolutional neural network
CPK	Creatine phosphokinase
CRP	C-reactive protein
CUN-BAE	Clínica Universidad de Navarra Body Fatness Estimator
DCIS	Ductal carcinoma in situ
ER	Estrogen receptor
FN	False Negative
FP	False Positive
FSH	Follicle-stimulating hormone
FURIA	Fuzzy Unordered Rule Induction Algorithm
GGT	Gamma-glutamyl-transferase
GH	Growth hormone
GLUT4	Glucose transporter type 4
GnRH	Gonadotropin-releasing hormone
HDL	High-density lipoprotein
HR+	Hormone receptor-positive
IRAP	Insulin-regulated aminopeptidase
JRIP	Java implementation of the RIPPER algorithm
LDH	Lactate dehydrogenase
LDL	Low-density lipoprotein
LH	Luteinizing hormone
miRNAs	MicroRNAs
MRI	Magnetic resonance imaging
NMU	N-methyl-nitrosourea
OneR	One Rule
OTR	Oxytocin receptor
PR	Progesterone receptor
RANKL	Receptor activator of nuclear factor kappa-B ligand
RAS	Renin–angiotensin system
RNN	Recurrent neural network
TAI	Trustworthy artificial intelligence
TN	True Negative
TNBC	Triple-negative breast cancer
TP	True Positive
TSH	Thyroid-stimulating hormone
WNT4	Wingless-related integration site 4

References

1. Alsheikhy, A.A.; Said, Y.; Shawly, T.; Alzahrani, A.K.; Lahza, H. Biomedical Diagnosis of Breast Cancer Using Deep Learning and Multiple Classifiers. *Diagnostics* **2022**, *12*, 2863. <https://doi.org/10.3390/diagnostics12112863>.
2. Shankar, G.S.; Onyema, E.M.; Kavın, B.P.; Gude, V.; Prasad, B.S. Breast Cancer Diagnosis Using Virtualization and Extreme Learning Algorithm Based on Deep Feed Forward Networks. *Biomed. Eng. Comput. Biol.* **2024**, *15*, 11795972241278907. <https://doi.org/10.1177/11795972241278907>.
3. Teoh, J.R.; Hasikin, K.; Lai, K.W.; Wu, X.; Li, C. Enhancing early breast cancer diagnosis through automated microcalcification detection using an optimized ensemble deep learning framework. *PeerJ Comput. Sci.* **2024**, *10*, e2082. <https://doi.org/10.7717/peerj-cs.2082>.
4. Kanbayashi, C.; Iwata, H. Update on the management of ductal carcinoma in situ of the breast: current approach and future perspectives. *Jpn. J. Clin. Oncol.* **2024**, *55*, 4–11. <https://doi.org/10.1093/jjco/hyae122>.

5. Dewangan, K.K.; Dewangan, D.K.; Sahu, S.P.; Janghel, R. Breast cancer diagnosis in an early stage using novel deep learning with hybrid optimization technique. *Multimed. Tools Appl.* **2022**, *81*, 13935–13960. <https://doi.org/10.1007/s11042-022-12385-2>.
6. Sanchez Galindo, B.; Lopez-Torres Hidalgo, J.; Lopez Gonzalez, A.; Sanchez Martinez, C.M.; Castano Diaz, M.; Rabanales Sotos, J. [Time intervals for diagnosis and treatment in breast cancer patients]. *Aten. Primaria* **2024**, *57*, 103153. <https://doi.org/10.1016/j.aprim.2024.103153>.
7. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. <https://doi.org/10.3322/caac.21660>.
8. Uchikov, P.; Khalid, U.; Dedaj-Salad, G.H.; Ghale, D.; Rajadurai, H.; Kraeva, M.; Kraev, K.; Hristov, B.; Doykov, M.; Mitova, V.; et al. Artificial Intelligence in Breast Cancer Diagnosis and Treatment: Advances in Imaging, Pathology, and Personalized Care. *Life* **2024**, *14*, 1451. <https://doi.org/10.3390/life14111451>.
9. Kavitha, T.; Mathai, P.P.; Karthikeyan, C.; Ashok, M.; Kohar, R.; Avanija, J.; Neelakandan, S. Deep Learning Based Capsule Neural Network Model for Breast Cancer Diagnosis Using Mammogram Images. *Interdiscip. Sci.* **2022**, *14*, 113–129. <https://doi.org/10.1007/s12539-021-00467-y>.
10. Aljuaid, H.; Alturki, N.; Alsubaie, N.; Cavallaro, L.; Liotta, A. Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. *Comput. Methods Programs Biomed.* **2022**, *223*, 106951. <https://doi.org/10.1016/j.cmpb.2022.106951>.
11. Devi, S.; Kaul Ghanekar, R.; Pande, J.A.; Dumbre, D.; Chavan, R.; Gupta, H. Prediction and Diagnosis of Breast Cancer Using Machine and Modern Deep Learning Models. *Asian Pac. J. Cancer Prev.* **2024**, *25*, 1077–1085. <https://doi.org/10.31557/APJCP.2024.25.3.1077>.
12. Alam, T.; Shia, W.C.; Hsu, F.R.; Hassan, T. Improving Breast Cancer Detection and Diagnosis through Semantic Segmentation Using the Unet3+ Deep Learning Framework. *Biomedicines* **2023**, *11*, 1536. <https://doi.org/10.3390/biomedicines11061536>.
13. Hu, Q.; Whitney, H.M.; Giger, M.L. A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI. *Sci. Rep.* **2020**, *10*, 10536. <https://doi.org/10.1038/s41598-020-67441-4>.
14. Kaur, J.; Kaur, P. A systematic literature analysis of multi-organ cancer diagnosis using deep learning techniques. *Comput. Biol. Med.* **2024**, *179*, 108910. <https://doi.org/10.1016/j.compbiomed.2024.108910>.
15. Schettini, F.; Braso-Maristany, F.; Pascual, T.; Lorman-Carbo, N.; Nucera, S.; Bergamino, M.; Galvan, P.; Conte, B.; Segui, E.; Garcia Fructuoso, I.; et al. Identifying predictors of treatment response and molecular changes induced by neoadjuvant chemotherapy and endocrine therapy in hormone receptor-positive/HER2-negative breast cancer: The NEOENDO translational study. *ESMO Open* **2024**, *9*, 103989. <https://doi.org/10.1016/j.esmoop.2024.103989>.
16. Zhang, M.; Li, S.; Xue, M.; Zhu, Q. Two-stage classification strategy for breast cancer diagnosis using ultrasound-guided diffuse optical tomography and deep learning. *J. Biomed. Opt.* **2023**, *28*, 086002. <https://doi.org/10.1117/1.JBO.28.8.086002>.
17. Zhang, Y.; Liu, Y.L.; Nie, K.; Zhou, J.; Chen, Z.; Chen, J.H.; Wang, X.; Kim, B.; Parajuli, R.; Mehta, R.S.; et al. Deep Learning-based Automatic Diagnosis of Breast Cancer on MRI Using Mask R-CNN for Detection Followed by ResNet50 for Classification. *Acad. Radiol.* **2023**, *30* (Suppl. 2), S161–S171. <https://doi.org/10.1016/j.acra.2022.12.038>.
18. Ejiyi, C.J.; Qin, Z.; Agbesi, V.K.; Yi, D.; Atwereboannah, A.A.; Chikwendu, I.A.; Bamisile, O.F.; Bakanina Kissanga, G.M.; Bamisile, O.O. Advancing cancer diagnosis and prognostication through deep learning mastery in breast, colon, and lung histopathology with ResoMergeNet. *Comput. Biol. Med.* **2024**, *185*, 109494. <https://doi.org/10.1016/j.compbiomed.2024.109494>.
19. Cheng, Z.; Li, H.; Chen, C.; Lv, X.; Zuo, E.; Xie, X.; Li, Z.; Liu, P.; Li, H.; Chen, C. Application of serum SERS technology based on thermally annealed silver nanoparticle composite substrate in breast cancer. *Photodiagnosis Photodyn. Ther.* **2023**, *41*, 103284. <https://doi.org/10.1016/j.pdpdt.2023.103284>.
20. Ma, X.; Cheng, H.; Hou, J.; Jia, Z.; Wu, G.; Lü, X.; Li, H.; Zheng, X.; Chen, C. Detection of breast cancer based on novel porous silicon Bragg reflector surface-enhanced Raman spectroscopy-active structure. *Chin. Opt. Lett.* **2020**, *18*, 051701.
21. Zeng, Q.; Chen, C.; Chen, C.; Song, H.; Li, M.; Yan, J.; Lv, X. Serum Raman spectroscopy combined with convolutional neural network for rapid diagnosis of HER2-positive and triple-negative breast cancer. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2023**, *286*, 122000. <https://doi.org/10.1016/j.saa.2022.122000>.
22. Samala, R.K.; Heang-Ping, C.; Hadjiiski, L.; Helvie, M.A.; Richter, C.D.; Cha, K.H. Breast Cancer Diagnosis in Digital Breast Tomosynthesis: Effects of Training Sample Size on Multi-Stage Transfer Learning Using Deep Neural Nets. *IEEE Trans. Med. Imaging* **2019**, *38*, 686–696. <https://doi.org/10.1109/TMI.2018.2870343>.
23. Yang, J.; Hussein Kadir, D. Data mining techniques in breast cancer diagnosis at the cellular-molecular level. *J. Cancer Res. Clin. Oncol.* **2023**, *149*, 12605–12620. <https://doi.org/10.1007/s00432-023-05090-6>.
24. Zhang, S.; Yang, B.; Yang, H.; Zhao, J.; Zhang, Y.; Gao, Y.; Monteiro, O.; Zhang, K.; Liu, B.; Wang, S. Potential rapid intraoperative cancer diagnosis using dynamic full-field optical coherence tomography and deep learning: A prospective cohort study in breast cancer patients. *Sci. Bull.* **2024**, *69*, 1748–1756. <https://doi.org/10.1016/j.scib.2024.03.061>.

25. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.
26. Cuadros, M. The EU Trustworthy Artificial Intelligence: How the EU Wants to Regulate the Artificial Intelligence Practices. In *Actas del Congreso Internacional de Ingeniería de Sistemas*; University of Lima: Lima Peru, 2021; pp. 173–181. <https://doi.org/10.26439/ciis2021.5585>.
27. Ramirez-Exposito, M.J.; Carrera-Gonzalez Mdel, P.; Mayas, M.D.; Duenas, B.; Martinez-Ferrol, J.; Martinez-Martos, J.M. Neoadjuvant chemotherapy modifies serum angiotensinase activities in women with breast cancer. *Maturitas* **2012**, *72*, 79–83. <https://doi.org/10.1016/j.maturitas.2012.02.007>.
28. Ramirez-Exposito, M.J.; Duenas-Rodriguez, B.; Martinez-Martos, J.M. Circulating renin-angiotensin system-regulating specific aminopeptidase activities in pre- and post- menopausal women with breast cancer treated or not with neoadjuvant chemotherapy. A two years follow up study. *Breast* **2019**, *43*, 28–30. <https://doi.org/10.1016/j.breast.2018.10.010>.
29. Ramirez-Exposito, M.J.; Sanchez-Lopez, E.; Cueto-Urena, C.; Duenas, B.; Carrera-Gonzalez, P.; Navarro-Cecilia, J.; Mayas, M.D.; Arias de Saavedra, J.M.; Sanchez-Agesta, R.; Martinez-Martos, J.M. Circulating oxidative stress parameters in pre- and post-menopausal healthy women and in women suffering from breast cancer treated or not with neoadjuvant chemotherapy. *Exp. Gerontol.* **2014**, *58*, 34–42. <https://doi.org/10.1016/j.exger.2014.07.006>.
30. Ramirez-Exposito, M.J.; Urbano-Polo, N.; Duenas, B.; Navarro-Cecilia, J.; Ramirez-Tortosa, C.; Martin-Salvago, M.D.; Martinez-Martos, J.M. Redox status in the sentinel lymph node of women with breast cancer. *Upsala J. Med. Sci.* **2017**, *122*, 207–216. <https://doi.org/10.1080/03009734.2017.1403522>.
31. Ramirez-Exposito, M.J.; Martinez-Martos, J.M.; Duenas-Rodriguez, B.; Navarro-Cecilia, J.; Carrera-Gonzalez, M.P. Neoadjuvant chemotherapy modifies serum pyrrolidone carboxypeptidase specific activity in women with breast cancer and influences circulating levels of GnRH and gonadotropins. *Breast Cancer Res. Treat.* **2020**, *182*, 751–760. <https://doi.org/10.1007/s10549-020-05723-1>.
32. Carrera-Gonzalez Mdel, P.; Ramirez-Exposito, M.J.; Duenas, B.; Martinez-Ferrol, J.; Mayas, M.D.; Martinez-Martos, J.M. Putative relationship between hormonal status and serum pyrrolidone carboxypeptidase activity in pre- and post- menopausal women with breast cancer. *Breast* **2012**, *21*, 751–754. <https://doi.org/10.1016/j.breast.2012.02.001>.
33. Martinez-Martos, J.M.; del Pilar Carrera-Gonzalez, M.; Duenas, B.; Mayas, M.D.; Garcia, M.J.; Ramirez-Exposito, M.J. Renin angiotensin system-regulating aminopeptidase activities in serum of pre- and postmenopausal women with breast cancer. *Breast* **2011**, *20*, 444–447. <https://doi.org/10.1016/j.breast.2011.04.008>.
34. Navarro-Cecilia, J.; Duenas-Rodriguez, B.; Luque-Lopez, C.; Ramirez-Exposito, M.J.; Martinez-Ferrol, J.; Ruiz-Mateas, A.; Urena, C.; Carrera-Gonzalez, M.P.; Mayas, M.D.; Martinez-Martos, J.M. Intraoperative sentinel node biopsy by one-step nucleic acid amplification (OSNA) avoids axillary lymphadenectomy in women with breast cancer treated with neoadjuvant chemotherapy. *Eur. J. Surg. Oncol.* **2013**, *39*, 873–879. <https://doi.org/10.1016/j.ejso.2013.05.002>.
35. Ramirez-Exposito, M.J.; Duenas-Rodriguez, B.; Carrera-Gonzalez, M.P.; Navarro-Cecilia, J.; Martinez-Martos, J.M. Circulating levels of beta-endorphin and cortisol in breast cancer. *Compr. Psychoneuroendocrinol.* **2021**, *5*, 100028. <https://doi.org/10.1016/j.cpnec.2021.100028>.
36. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. <https://doi.org/10.1613/jair.953>.
37. Holte, R.C. Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. *Mach. Learn.* **1993**, *11*, 63–90.
38. Cohen, W.W. Fast Effective Rule Induction. In *Machine Learning Proceedings 1995, Proceedings of the Twelfth International Conference on Machine Learning, Tahoe City, California, 9–12 July 1995*; Morgan Kaufmann: Burlington, MA, USA, 1995; pp. 115–123.
39. Wang, Y. Review on greedy algorithm. *Theor. Nat. Sci.* **2023**, *14*, 233–239. <https://doi.org/10.54254/2753-8818/14/20241041>.
40. Hühn, J.; Hüllermeier, E. FURIA: An algorithm for unordered fuzzy rule induction. *Data Min. Knowl. Discov.* **2009**, *19*, 293–319. <https://doi.org/10.1007/s10618-009-0131-8>.
41. Ihya, R.; Namir, A.; Filali, S.; Aitdaoud, M.; Guerss, f.Z. J48 algorithms of machine learning for predicting user's the acceptance of an E-orientation systems. In *Proceedings of the SCA '19: The 4th International Conference on Smart City Applications, Casablanca, Morocco, 2–4 October 2019*; pp. 1–8. <https://doi.org/10.1145/3368756.3368995>.
42. Quinlan, J.R. *C4.5: Programs for Machine Learning*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1993.
43. Venkatesan, E.; Thambusamy, V. Performance Analysis of Decision Tree Algorithms for Breast Cancer Classification. *Indian J. Sci. Technol.* **2015**, *8*, IPL0625. <https://doi.org/10.17485/ijst/2015/v8i1/84646>.
44. Saravanan, N.; Gayathri, V. Performance and Classification Evaluation of J48 Algorithm and Kendall's Based J48 Algorithm (KNJ48). *Int. J. Comput. Trends Technol.* **2018**, *59*, 73–80. <https://doi.org/10.14445/22312803/IJCTT-V59P112>.
45. Freund, Y.; Mason, L. The Alternating Decision Tree Learning Algorithm. In *Proceedings of the Sixteenth International Conference on Machine Learning, Bled, Slovenia, 27–30 June 1999*; ICML '99; pp. 124–133.
46. Pfahringer, B.; Holmes, G.; Kirkby, R. Optimizing the Induction of Alternating Decision Trees. In *Proceedings of the 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining, Hong Kong, China, 16–18 April 2001*; PAKDD '01; pp. 477–487.

47. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. <https://doi.org/10.1023/A:1010933404324>.
48. Nasser, M.; Yusof, U.K. Deep learning based methods for breast cancer diagnosis: A systematic review and future direction. *Diagnostics* **2023**, *13*, 161.
49. Manimurugan, S.; Karthikeyan, P.; Aborokbah, M.; Narmatha, C.; Ganesan, S. Breast cancer diagnosis model using stacked autoencoder with particle swarm optimization. *Ain Shams Eng. J.* **2024**, *15*, 102734. <https://doi.org/10.1016/j.asej.2024.102734>.
50. Singh, S.J.; Rajaraman, R.; Verlekar, T.T. Breast Cancer Prediction Using Auto-Encoders. In Proceedings of the Data Management, Analytics and Innovation, Pune, India, 14–16 January 2022; Goswami, S.; Barara, I.S.; Goje, A.; Mohan, C.; Bruckstein, A.M., Eds.; Springer: Singapore, 2023; pp. 121–132.
51. Alotaibi, M.; Aljouie, A.; Alluhaidan, N.; Qureshi, W.; Almatar, H.; Alduhayan, R.; Alsomaie, B.; Almazroa, A. Breast cancer classification based on convolutional neural network and image fusion approaches using ultrasound images. *Heliyon* **2023**, *9*, e22406. <https://doi.org/10.1016/j.heliyon.2023.e22406>.
52. Das, H.S.; Das, A.; Neog, A.; Mallik, S.; Bora, K.; Zhao, Z. Breast cancer detection: Shallow convolutional neural network against deep convolutional neural networks based approach. *Front. Genet.* **2022**, *13*, 1097207.
53. Xiang, Q.; Li, D.; Hu, Z.; Yuan, Y.; Sun, Y.; Zhu, Y.; Fu, Y.; Jiang, Y.; Hua, X. Quantum classical hybrid convolutional neural networks for breast cancer diagnosis. *Sci. Rep.* **2024**, *14*, 24699.
54. Minnoor, M.; Baths, V. Diagnosis of Breast Cancer Using Random Forests. *Procedia Comput. Sci.* **2023**, *218*, 429–437. <https://doi.org/10.1016/j.procs.2023.01.025>.
55. Whig, P.; Kouser, S.; Bhatia, A.B.; Nadikattu, R.R. Chapter 3—Prediction of breast cancer diagnosis using random forest classifier. In *Computational Intelligence and Modelling Techniques for Disease Detection in Mammogram Images*; Hemanth, D.J., Ed.; Academic Press: Cambridge, MA, USA, 2024; pp. 55–73. <https://doi.org/10.1016/B978-0-443-13999-4.00011-0>.
56. Clift, A.K.; Dodwell, D.; Lord, S.; Petrou, S.; Brady, M.; Collins, G.S.; Hippisley-Cox, J. Development and internal-external validation of statistical and machine learning models for breast cancer prognostication: Cohort study. *BMJ* **2023**, *381*, e073800. <https://doi.org/10.1136/bmj-2022-073800>.
57. Khalid, A.; Mehmood, A.; Alabrah, A.; Alkhamees, B.F.; Amin, F.; AlSalman, H.; Choi, G.S. Breast Cancer Detection and Prevention Using Machine Learning. *Diagnostics* **2023**, *13*, 3113. <https://doi.org/10.3390/diagnostics13193113>.
58. Sammut, S.J.; Crispin-Ortuzar, M.; Chin, S.F.; Provenzano, E.; Bardwell, H.A.; Ma, W.; Cope, W.; Dariush, A.; Dawson, S.J.; Abraham, J.E.; et al. Multi-omic machine learning predictor of breast cancer therapy response. *Nature* **2022**, *601*, 623–629. <https://doi.org/10.1038/s41586-021-04278-5>.
59. Stark, G.F.; Hart, G.R.; Nartowt, B.J.; Deng, J. Predicting breast cancer risk using personal health data and machine learning models. *PLoS ONE* **2019**, *14*, e0226765. <https://doi.org/10.1371/journal.pone.0226765>.
60. Zuo, D.; Yang, L.; Jin, Y.; Qi, H.; Liu, Y.; Ren, L. Machine learning-based models for the prediction of breast cancer recurrence risk. *BMC Med. Inf. Decis. Mak.* **2023**, *23*, 276. <https://doi.org/10.1186/s12911-023-02377-z>.
61. Ito, Y.; Kobayashi, T.; Kimura, T.; Matsuura, N.; Wakasugi, E.; Takeda, T.; Shimano, T.; Kubota, Y.; Nobunaga, T.; Makino, Y.; et al. Investigation of the oxytocin receptor expression in human breast cancer tissue using newly established monoclonal antibodies. *Endocrinology* **1996**, *137*, 773–779. <https://doi.org/10.1210/endo.137.2.8593829>.
62. Lerman, B.; Harricharran, T.; Ogunwobi, O.O. Oxytocin and cancer: An emerging link. *World J. Clin. Oncol.* **2018**, *9*, 74–82. <https://doi.org/10.5306/wjco.v9.i5.74>.
63. Liu, H.; Gruber, C.W.; Alewood, P.F.; Moller, A.; Muttenthaler, M. The oxytocin receptor signalling system and breast cancer: A critical review. *Oncogene* **2020**, *39*, 5917–5932. <https://doi.org/10.1038/s41388-020-01415-8>.
64. Liu, S.; Huang, R.; Li, A.; Yu, S.; Yao, S.; Xu, J.; Tang, L.; Li, W.; Gan, C.; Cheng, H. The role of the oxytocin system in the resilience of patients with breast cancer. *Front. Oncol.* **2023**, *13*, 1187477. <https://doi.org/10.3389/fonc.2023.1187477>.
65. Miranda-Olvera, A.D.; Ferro-Flores, G.; Pedraza-Lopez, M.; de Murphy, C.A.; De Leon-Rodriguez, L.M. Synthesis of oxytocin HYNIC derivatives as potential diagnostic agents for breast cancer. *Bioconjug. Chem.* **2007**, *18*, 1560–1567. <https://doi.org/10.1021/bc070047a>.
66. Murrell, T.G. The potential for oxytocin (OT) to prevent breast cancer: A hypothesis. *Breast Cancer Res. Treat.* **1995**, *35*, 225–229. <https://doi.org/10.1007/BF00668213>.
67. Suijkerbuijk, K.P.; van der Wall, E.; Meijrink, H.; Pan, X.; Borel Rinkes, I.H.; Ausems, M.G.; van Diest, P.J. Successful oxytocin-assisted nipple aspiration in women at increased risk for breast cancer. *Fam. Cancer* **2010**, *9*, 321–325. <https://doi.org/10.1007/s10689-010-9344-7>.
68. Alizadeh, A.M.; Heydari, Z.; Rahimi, M.; Bazgir, B.; Shirvani, H.; Alipour, S.; Heidarian, Y.; Khalighfard, S.; Isanejad, A. Oxytocin mediates the beneficial effects of the exercise training on breast cancer. *Exp. Physiol.* **2018**, *103*, 222–235. <https://doi.org/10.1113/EP086463>.
69. Ariana, M.; Pornour, M.; Mehr, S.S.; Vaseghi, H.; Ganji, S.M.; Alivand, M.R.; Salari, M.; Akbari, M.E. Preventive effects of oxytocin and oxytocin receptor in breast cancer pathogenesis. *Per. Med.* **2019**, *16*, 25–34. <https://doi.org/10.2217/pme-2018-0009>.

70. Cassoni, P.; Sapino, A.; Negro, F.; Bussolati, G. Oxytocin inhibits proliferation of human breast cancer cell lines. *Virchows Arch.* **1994**, *425*, 467–472. <https://doi.org/10.1007/BF00197549>.
71. Cassoni, P.; Sapino, A.; Fortunati, N.; Munaron, L.; Chini, B.; Bussolati, G. Oxytocin inhibits the proliferation of MDA-MB231 human breast-cancer cells via cyclic adenosine monophosphate and protein kinase A. *Int. J. Cancer* **1997**, *72*, 340–344. [https://doi.org/10.1002/\(sici\)1097-0215\(19970717\)72:2<340::aid-ijc23>3.0.co;2-i](https://doi.org/10.1002/(sici)1097-0215(19970717)72:2<340::aid-ijc23>3.0.co;2-i).
72. Kalaba, P.; Sanchez de la Rosa, C.; Moller, A.; Alewood, P.F.; Muttenthaler, M. Targeting the Oxytocin Receptor for Breast Cancer Management: A Niche for Peptide Tracers. *J. Med. Chem.* **2024**, *67*, 1625–1640. <https://doi.org/10.1021/acs.jmedchem.3c01089>.
73. Khorri, V.; Alizadeh, A.M.; Khalighfard, S.; Heidarian, Y.; Khodayari, H. Oxytocin effects on the inhibition of the NF-kappaB/miR195 pathway in mice breast cancer. *Peptides* **2018**, *107*, 54–60. <https://doi.org/10.1016/j.peptides.2018.07.007>.
74. Amico, J.A.; Rauk, P.N.; Cai, H.M. Estradiol and progesterone regulate oxytocin receptor binding and expression in human breast cancer cell lines. *Endocrine* **2002**, *18*, 79–84. <https://doi.org/10.1385/ENDO:18:1:79>.
75. Kalinina, T.S.; Kononchuk, V.V.; Sidorov, S.V.; Obukhova, D.A.; Abdullin, G.R.; Gulyaeva, L.F. [Oxytocin receptor expression is associated with estrogen receptor status in breast tumors]. *Biomed. Khim.* **2021**, *67*, 360–365. <https://doi.org/10.18097/pbmc20216704360>.
76. Cassoni, P.; Catalano, M.G.; Sapino, A.; Marrocco, T.; Fazzari, A.; Bussolati, G.; Fortunati, N. Oxytocin modulates estrogen receptor alpha expression and function in MCF7 human breast cancer cells. *Int. J. Oncol.* **2002**, *21*, 375–378.
77. Behtaji, S.; Ghafouri-Fard, S.; Sayad, A.; Sattari, A.; Rederstorff, M.; Taheri, M. Identification of oxytocin-related lncRNAs and assessment of their expression in breast cancer. *Sci. Rep.* **2021**, *11*, 6471. <https://doi.org/10.1038/s41598-021-86097-2>.
78. Cassoni, P.; Marrocco, T.; Deaglio, S.; Sapino, A.; Bussolati, G. Biological relevance of oxytocin and oxytocin receptors in cancer cells and primary tumors. *Ann. Oncol.* **2001**, *12* (Suppl. 2), S37–S39. https://doi.org/10.1093/annonc/12.suppl_2.s37.
79. Liu, H.; Muttenthaler, M. High Oxytocin Receptor Expression Linked to Increased Cell Migration and Reduced Survival in Patients with Triple-Negative Breast Cancer. *Biomedicines* **2022**, *10*, 1595. <https://doi.org/10.3390/biomedicines10071595>.
80. Alfari, I.; Asselah, J.; Aziz, H.; Bouganin, N.; Mousavi, N. The Cardiovascular Risks Associated with Aromatase Inhibitors, Tamoxifen, and GnRH Agonists in Women with Breast Cancer. *Curr. Atheroscler. Rep.* **2023**, *25*, 145–154. <https://doi.org/10.1007/s11883-023-01085-3>.
81. Sukumar, J.S.; Quiroga, D.; Kassem, M.; Grimm, M.; Shinde, N.V.; Appiah, L.; Palettas, M.; Stephens, J.; Gatti-Mays, M.E.; Pariser, A.; et al. Patient preferences and adherence to adjuvant GnRH analogs among premenopausal women with hormone receptor positive breast cancer. *Breast Cancer Res. Treat.* **2021**, *190*, 183–188. <https://doi.org/10.1007/s10549-021-06368-4>.
82. LeVasseur, N.; Manna, M.; Jerzak, K.J. An Overview of Long-Acting GnRH Agonists in Premenopausal Breast Cancer Patients: Survivorship Challenges and Management. *Curr. Oncol.* **2024**, *31*, 4209–4224. <https://doi.org/10.3390/curroncol31080314>.
83. Bergami, M.; Manfrini, O.; Cenko, E.; Bugiardini, R. Combined Therapy with Anthracyclines and GnRH Analogues for Breast Cancer: Impact on Ischemic Heart Disease. *J. Clin. Med.* **2023**, *12*, 6791. <https://doi.org/10.3390/jcm12216791>.
84. Ma, L.; Yang, B.; Wu, J. Revisiting ovarian function suppression with GnRH agonists for premenopausal women with breast cancer: Who should use and the impact on survival outcomes. *Cancer Treat. Rev.* **2024**, *129*, 102770. <https://doi.org/10.1016/j.ctrv.2024.102770>.
85. Sarkar, D.K.; Zhang, C. Beta-endorphin neuron regulates stress response and innate immunity to prevent breast cancer growth and progression. *Vitam. Horm.* **2013**, *93*, 263–276. <https://doi.org/10.1016/B978-0-12-416673-8.00011-3>.
86. Sanchez, M.L.; Rodriguez, F.D.; Covenas, R. Involvement of the Opioid Peptide Family in Cancer Progression. *Biomedicines* **2023**, *11*, 1993. <https://doi.org/10.3390/biomedicines11071993>.
87. Iannucci, N.B.; Ripoll, G.V.; Garona, J.; Cascone, O.; Ciccia, G.N.; Gomez, D.E.; Alonso, D.F. Antiproliferative effect of 1-deamino-8-D-arginine vasopressin analogs on human breast cancer cells. *Future Med. Chem.* **2011**, *3*, 1987–1993. <https://doi.org/10.4155/fmc.11.152>.
88. North, W.G.; Fay, M.J.; Du, J. MCF-7 breast cancer cells express normal forms of all vasopressin receptors plus an abnormal V2R. *Peptides* **1999**, *20*, 837–842. [https://doi.org/10.1016/S0196-9781\(99\)00070-4](https://doi.org/10.1016/S0196-9781(99)00070-4).
89. Alkafaas, S.S.; Loutfy, S.A.; Diab, T.; Hessien, M. Vasopressin induces apoptosis but does not enhance the antiproliferative effect of dynamin 2 or PI3K/Akt inhibition in luminal A breast cancer cells. *Med. Oncol.* **2022**, *40*, 35. <https://doi.org/10.1007/s12032-022-01889-4>.
90. Garona, J.; Pifano, M.; Orlando, U.D.; Pastrian, M.B.; Iannucci, N.B.; Ortega, H.H.; Podesta, E.J.; Gomez, D.E.; Ripoll, G.V.; Alonso, D.F. The novel desmopressin analogue [V4Q5]dDAVP inhibits angiogenesis, tumour growth and metastases in vasopressin type 2 receptor-expressing breast cancer models. *Int. J. Oncol.* **2015**, *46*, 2335–2345. <https://doi.org/10.3892/ijo.2015.2952>.
91. Garona, J.; Alonso, D.F. Perioperative biology in primary breast cancer: Selective targeting of vasopressin type 2 receptor using desmopressin as a novel therapeutic approach. *Breast Cancer Res. Treat.* **2016**, *158*, 597–599. <https://doi.org/10.1007/s10549-016-3899-4>.
92. Keller, S.R. The insulin-regulated aminopeptidase: A companion and regulator of GLUT4. *Front. Biosci.* **2003**, *8*, s410–s420. <https://doi.org/10.2741/1078>.

93. Wallis, M.G.; Lankford, M.F.; Keller, S.R. Vasopressin is a physiological substrate for the insulin-regulated aminopeptidase IRAP. *Am. J. Physiol. Endocrinol. Metab.* **2007**, *293*, E1092–E1102. <https://doi.org/10.1152/ajpendo.00440.2007>.
94. Albiston, A.L.; McDowall, S.G.; Matsacos, D.; Sim, P.; Clune, E.; Mustafa, T.; Lee, J.; Mendelsohn, F.A.; Simpson, R.J.; Connolly, L.M.; et al. Evidence that the angiotensin IV (AT(4)) receptor is the enzyme insulin-regulated aminopeptidase. *J. Biol. Chem.* **2001**, *276*, 48623–48626. <https://doi.org/10.1074/jbc.C100512200>.
95. Albiston, A.L.; Mustafa, T.; McDowall, S.G.; Mendelsohn, F.A.; Lee, J.; Chai, S.Y. AT4 receptor is insulin-regulated membrane aminopeptidase: Potential mechanisms of memory enhancement. *Trends Endocrinol. Metab.* **2003**, *14*, 72–77. [https://doi.org/10.1016/s1043-2760\(02\)00037-1](https://doi.org/10.1016/s1043-2760(02)00037-1).
96. Keller, S.R. Role of the insulin-regulated aminopeptidase IRAP in insulin action and diabetes. *Biol. Pharm. Bull.* **2004**, *27*, 761–764. <https://doi.org/10.1248/bpb.27.761>.
97. Carrera, M.P.; Ramirez-Exposito, M.J.; Valenzuela, M.T.; Duenas, B.; Garcia, M.J.; Mayas, M.D.; Martinez-Martos, J.M. Renin-angiotensin system-regulating aminopeptidase activities are modified in the pineal gland of rats with breast cancer induced by N-methyl-nitrosourea. *Cancer Investig.* **2006**, *24*, 149–153. <https://doi.org/10.1080/07357900500524389>.
98. Carrera-Gonzalez, M.P.; Ramirez-Exposito, M.J.; de Saavedra, J.M.; Sanchez-Agosta, R.; Mayas, M.D.; Martinez-Martos, J.M. Hypothalamus-pituitary-thyroid axis disruption in rats with breast cancer is related to an altered endogenous oxytocin/insulin-regulated aminopeptidase (IRAP) system. *Tumour Biol.* **2011**, *32*, 543–549. <https://doi.org/10.1007/s13277-010-0149-y>.
99. Martinez, J.M.; Prieto, I.; Ramirez, M.J.; Cueva, C.; Alba, F.; Ramirez, M. Aminopeptidase activities in breast cancer tissue. *Clin. Chem.* **1999**, *45*, 1797–1802.
100. Arrieta, O.; Villarreal-Garza, C.; Vizcaino, G.; Pineda, B.; Hernandez-Pedro, N.; Guevara-Salazar, P.; Wegman-Ostrosky, T.; Villanueva-Rodriguez, G.; Gamboa-Dominguez, A. Association between AT1 and AT2 angiotensin II receptor expression with cell proliferation and angiogenesis in operable breast cancer. *Tumour Biol.* **2015**, *36*, 5627–5634. <https://doi.org/10.1007/s13277-015-3235-3>.
101. Ramirez-Exposito, M.J.; Duenas-Rodriguez, B.; Carrera-Gonzalez, M.P.; Navarro-Cecilia, J.; Martinez-Martos, J.M. Insulin-Regulated Aminopeptidase in Women with Breast Cancer: A Role beyond the Regulation of Oxytocin and Vasopressin. *Cancers* **2020**, *12*, 3252. <https://doi.org/10.3390/cancers12113252>.
102. Lew, R.A.; Mustafa, T.; Ye, S.; McDowall, S.G.; Chai, S.Y.; Albiston, A.L. Angiotensin AT4 ligands are potent, competitive inhibitors of insulin regulated aminopeptidase (IRAP). *J. Neurochem.* **2003**, *86*, 344–350. <https://doi.org/10.1046/j.1471-4159.2003.01852.x>.
103. Medina, R.A.; Meneses, A.M.; Vera, J.C.; Guzman, C.; Nualart, F.; Astuya, A.; Garcia, M.A.; Kato, S.; Carvajal, A.; Pinto, M.; et al. Estrogen and progesterone up-regulate glucose transporter expression in ZR-75-1 human breast cancer cells. *Endocrinology* **2003**, *144*, 4527–4535. <https://doi.org/10.1210/en.2003-0294>.
104. Harmon, A.W.; Patel, Y.M. Naringenin inhibits glucose uptake in MCF-7 breast cancer cells: A mechanism for impaired cellular proliferation. *Breast Cancer Res. Treat.* **2004**, *85*, 103–110. <https://doi.org/10.1023/B:BREA.0000025397.56192.e2>.
105. Garrido, P.; Osorio, F.G.; Moran, J.; Cabello, E.; Alonso, A.; Freije, J.M.; Gonzalez, C. Loss of GLUT4 induces metabolic reprogramming and impairs viability of breast cancer cells. *J. Cell Physiol.* **2015**, *230*, 191–198. <https://doi.org/10.1002/jcp.24698>.
106. Ruiz-Sanjuan, M.D.; Martinez-Martos, J.M.; Carrera-Gonzalez, M.P.; Mayas, M.D.; Garcia, M.J.; Arrazola, M.; Ramirez-Exposito, M.J. Normolipidic dietary fat modifies circulating Renin-Angiotensin system-regulating aminopeptidase activities in rat with breast cancer. *Integr. Cancer Ther.* **2015**, *14*, 149–155. <https://doi.org/10.1177/1534735414562956>.
107. Ma, Y.; Yang, X.; Pan, P.; Yang, J.; Wu, X.; Wang, D.; Gao, H. Bestatin attenuates breast cancer stemness by targeting puromycin-sensitive aminopeptidase. *Discov. Oncol.* **2024**, *15*, 197. <https://doi.org/10.1007/s12672-024-01063-4>.
108. Carrera, M.P.; Ramirez-Exposito, M.J.; Mayas, M.D.; Garcia, M.J.; Martinez-Martos, J.M. Mammary renin-angiotensin system-regulating aminopeptidase activities are modified in rats with breast cancer. *Tumour Biol.* **2010**, *31*, 583–588. <https://doi.org/10.1007/s13277-010-0072-2>.
109. Dwyer, A.R.; Truong, T.H.; Ostrander, J.H.; Lange, C.A. 90 YEARS OF PROGESTERONE: Steroid receptors as MAPK signaling sensors in breast cancer: Let the fates decide. *J. Mol. Endocrinol.* **2020**, *65*, T35–T48. <https://doi.org/10.1530/JME-19-0274>.
110. Horwitz, K.B.; Sartorius, C.A. 90 YEARS OF PROGESTERONE: Progesterone and progesterone receptors in breast cancer: Past, present, future. *J. Mol. Endocrinol.* **2020**, *65*, T49–T63. <https://doi.org/10.1530/JME-20-0104>.
111. Trabert, B.; Sherman, M.E.; Kannan, N.; Stanczyk, F.Z. Progesterone and Breast Cancer. *Endocr. Rev.* **2020**, *41*, 320–344. <https://doi.org/10.1210/endrev/bnz001>.
112. Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.; Gronemeyer, H.; Chambon, P. Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. *EMBO J.* **1990**, *9*, 1603–1614. <https://doi.org/10.1002/j.1460-2075.1990.tb08280.x>.
113. Li, Z.; Wei, H.; Li, S.; Wu, P.; Mao, X. The Role of Progesterone Receptors in Breast Cancer. *Drug Des. Dev. Ther.* **2022**, *16*, 305–314. <https://doi.org/10.2147/DDDT.S336643>.
114. Beato, M.; Wright, R.H.G.; Dily, F.L. 90 YEARS OF PROGESTERONE: Molecular mechanisms of progesterone receptor action on the breast cancer genome. *J. Mol. Endocrinol.* **2020**, *65*, T65–T79. <https://doi.org/10.1530/JME-19-0266>.

115. Bighetti-Trevisan, R.L.; Sousa, L.O.; Castilho, R.M.; Almeida, L.O. Cancer Stem Cells: Powerful Targets to Improve Current Anticancer Therapeutics. *Stem Cells Int.* **2019**, 2019, 9618065. <https://doi.org/10.1155/2019/9618065>.
116. Briskin, C.; Scabia, V. 90 YEARS OF PROGESTERONE: Progesterone receptor signaling in the normal breast and its implications for cancer. *J. Mol. Endocrinol.* **2020**, 65, T81–T94. <https://doi.org/10.1530/JME-20-0091>.
117. Tanos, T.; Sflomos, G.; Echeverria, P.C.; Ayyanan, A.; Gutierrez, M.; Delaloye, J.F.; Raffoul, W.; Fiche, M.; Dougall, W.; Schneider, P.; et al. Progesterone/RANKL is a major regulatory axis in the human breast. *Sci. Transl. Med.* **2013**, 5, 182ra55. <https://doi.org/10.1126/scitranslmed.3005654>.
118. Truong, T.H.; Lange, C.A.; Ostrander, J.H. Targeting steroid receptor co-activators to inhibit breast cancer stem cells. *Oncoscience* **2018**, 5, 281–282. <https://doi.org/10.18632/oncoscience.469>.
119. Joshi, P.A.; Jackson, H.W.; Beristain, A.G.; Di Grappa, M.A.; Mote, P.A.; Clarke, C.L.; Stingl, J.; Waterhouse, P.D.; Khokha, R. Progesterone induces adult mammary stem cell expansion. *Nature* **2010**, 465, 803–807. <https://doi.org/10.1038/nature09091>.
120. Diep, C.H.; Charles, N.J.; Gilks, C.B.; Kalloger, S.E.; Argenta, P.A.; Lange, C.A. Progesterone receptors induce FOXO1-dependent senescence in ovarian cancer cells. *Cell Cycle* **2013**, 12, 1433–1449. <https://doi.org/10.4161/cc.24550>.
121. Singhal, H.; Greene, M.E.; Tarulli, G.; Zarnke, A.L.; Bourgo, R.J.; Laine, M.; Chang, Y.F.; Ma, S.; Dembo, A.G.; Raj, G.V.; et al. Genomic agonism and phenotypic antagonism between estrogen and progesterone receptors in breast cancer. *Sci. Adv.* **2016**, 2, e1501924. <https://doi.org/10.1126/sciadv.1501924>.
122. Bardou, V.J.; Arpino, G.; Elledge, R.M.; Osborne, C.K.; Clark, G.M. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. *J. Clin. Oncol.* **2003**, 21, 1973–1979. <https://doi.org/10.1200/JCO.2003.09.099>.
123. Arpino, G.; Weiss, H.; Lee, A.V.; Schiff, R.; De Placido, S.; Osborne, C.K.; Elledge, R.M. Estrogen receptor-positive, progesterone receptor-negative breast cancer: Association with growth factor receptor expression and tamoxifen resistance. *J. Natl. Cancer Inst.* **2005**, 97, 1254–1261. <https://doi.org/10.1093/jnci/dji249>.
124. Cui, X.; Schiff, R.; Arpino, G.; Osborne, C.K.; Lee, A.V. Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy. *J. Clin. Oncol.* **2005**, 23, 7721–7735. <https://doi.org/10.1200/JCO.2005.09.004>.
125. Piva, M.; Domenici, G.; Iriondo, O.; Rabano, M.; Simoes, B.M.; Comaills, V.; Barredo, I.; Lopez-Ruiz, J.A.; Zabalza, I.; Kypta, R.; et al. Sox2 promotes tamoxifen resistance in breast cancer cells. *EMBO Mol. Med.* **2014**, 6, 66–79. <https://doi.org/10.1002/emmm.201303411>.
126. Bricard, G.; Cadassou, O.; Cassagnes, L.E.; Cros-Perrial, E.; Payen-Gay, L.; Puy, J.Y.; Lefebvre-Tournier, I.; Tozzi, M.G.; Dumontet, C.; Jordheim, L.P. The cytosolic 5'-nucleotidase cN-II lowers the adaptability to glucose deprivation in human breast cancer cells. *Oncotarget* **2017**, 8, 67380–67393. <https://doi.org/10.18632/oncotarget.18653>.
127. da Silva, F.C.; de Melo Neto, A.B.; Martins, C.A.; de Sousa Cardoso, T.C.; de Souza Gomes, M.; de Araujo, T.G.; Furstenau, C.R. Is the regulation by miRNAs of NTPDase1 and ecto-5'-nucleotidase genes involved with the different profiles of breast cancer subtypes? *Purinergic Signal.* **2022**, 18, 123–133. <https://doi.org/10.1007/s11302-021-09824-4>.
128. Jedrzejewska, A.; Kutryb-Zajac, B.; Krol, O.; Harasim, G.; Franczak, M.; Jablonska, P.; Slominska, E.; Smolinski, R.T. The decreased serum activity of cytosolic 5'-nucleotidase IA as a potential marker of breast cancer-associated muscle inflammation. *Nucleosides Nucleotides Nucleic Acids* **2022**, 41, 273–284. <https://doi.org/10.1080/15257770.2021.2007396>.
129. Yu, J.; Wang, X.; Lu, Q.; Wang, J.; Li, L.; Liao, X.; Zhu, W.; Lv, L.; Zhi, X.; Yu, J.; et al. Extracellular 5'-nucleotidase (CD73) promotes human breast cancer cells growth through AKT/GSK-3beta/beta-catenin/cyclinD1 signaling pathway. *Int. J. Cancer* **2018**, 142, 959–967. <https://doi.org/10.1002/ijc.31112>.
130. Rocha-Vieira, T.C.; Lacerda-Abreu, M.A.; Carvalho-Kelly, L.F.; Santos-Araujo, S.; Gondim, K.C.; Meyer-Fernandes, J.R. Comparative characterisation of an ecto-5'-nucleotidase (CD73) in non-tumoral MCF10-A breast cells and triple-negative MDA-MB-231 breast cancer cells. *Cell Biol. Int.* **2024**, 48, 1354–1363. <https://doi.org/10.1002/cbin.12202>.
131. Stagg, J.; Smyth, M.J. Extracellular adenosine triphosphate and adenosine in cancer. *Oncogene* **2010**, 29, 5346–5358. <https://doi.org/10.1038/onc.2010.292>.
132. Stagg, J.; Divisekera, U.; McLaughlin, N.; Sharkey, J.; Pommey, S.; Denoyer, D.; Dwyer, K.M.; Smyth, M.J. Anti-CD73 antibody therapy inhibits breast tumor growth and metastasis. *Proc. Natl. Acad. Sci. USA* **2010**, 107, 1547–1552. <https://doi.org/10.1073/pnas.0908801107>.
133. Wang, L.; Zhou, X.; Zhou, T.; Ma, D.; Chen, S.; Zhi, X.; Yin, L.; Shao, Z.; Ou, Z.; Zhou, P. Ecto-5'-nucleotidase promotes invasion, migration and adhesion of human breast cancer cells. *J. Cancer Res. Clin. Oncol.* **2008**, 134, 365–372. <https://doi.org/10.1007/s00432-007-0292-z>.
134. Qi, Y.; Li, M.; Li, S.; Zeng, D.; Xiao, Y.; Li, J.; Ye, Q.; Bremer, E.; Zhang, G.J. Notch1 promotes resistance to cisplatin by up-regulating Ecto-5'-nucleotidase (CD73) in triple-negative breast cancer cells. *Cell Death Discov.* **2023**, 9, 204. <https://doi.org/10.1038/s41420-023-01487-x>.
135. Wang, L.; Tang, S.; Wang, Y.; Xu, S.; Yu, J.; Zhi, X.; Ou, Z.; Yang, J.; Zhou, P.; Shao, Z. Ecto-5'-nucleotidase (CD73) promotes tumor angiogenesis. *Clin. Exp. Metastasis* **2013**, 30, 671–680. <https://doi.org/10.1007/s10585-013-9571-z>.

136. Yu, J.; Liao, X.; Li, L.; Lv, L.; Zhi, X.; Yu, J.; Zhou, P. A preliminary study of the role of extracellular-5'- nucleotidase in breast cancer stem cells and epithelial-mesenchymal transition. *Vitr. Cell Dev. Biol. Anim.* **2017**, *53*, 132–140. <https://doi.org/10.1007/s11626-016-0089-y>.
137. Ruiz-Espejo, F.; Cabezas-Herrera, J.; Illana, J.; Campoy, F.J.; Vidal, C.J. Cholinesterase activity and acetylcholinesterase glycosylation are altered in human breast cancer. *Breast Cancer Res. Treat.* **2002**, *72*, 11–22. <https://doi.org/10.1023/a:1014904701723>.
138. Bernardi, C.C.; Ribeiro Ede, S.; Cavalli, I.J.; Chautard-Freire-Maia, E.A.; Souza, R.L. Amplification and deletion of the ACHE and BCHE cholinesterase genes in sporadic breast cancer. *Cancer Genet. Cytogenet.* **2010**, *197*, 158–165. <https://doi.org/10.1016/j.cancergencyto.2009.10.011>.
139. Grisaru, D.; Sternfeld, M.; Eldor, A.; Glick, D.; Soreq, H. Structural roles of acetylcholinesterase variants in biology and pathology. *Eur. J. Biochem.* **1999**, *264*, 672–686. <https://doi.org/10.1046/j.1432-1327.1999.00693.x>.
140. Nitsch, R.M.; Rossner, S.; Albrecht, C.; Mayhaus, M.; Enderich, J.; Schliebs, R.; Wegner, M.; Arendt, T.; von der Kammer, H. Muscarinic acetylcholine receptors activate the acetylcholinesterase gene promoter. *J. Physiol. Paris* **1998**, *92*, 257–264. [https://doi.org/10.1016/s0928-4257\(98\)80029-6](https://doi.org/10.1016/s0928-4257(98)80029-6).
141. Speirs, V.; Kerin, M.J.; Newton, C.J.; Walton, D.S.; Green, A.R.; Desai, S.B.; Atkin, S.L. Evidence for transcriptional activation of ERalpha by IL-1beta in breast cancer cells. *Int. J. Oncol.* **1999**, *15*, 1251–1254. <https://doi.org/10.3892/ijo.15.6.1251>.
142. Li, Y.; Liu, L.; Kang, J.; Sheng, J.G.; Barger, S.W.; Mrak, R.E.; Griffin, W.S. Neuronal-glia interactions mediated by interleukin-1 enhance neuronal acetylcholinesterase activity and mRNA expression. *J. Neurosci.* **2000**, *20*, 149–155. <https://doi.org/10.1523/JNEUROSCI.20-01-00149.2000>.
143. Ponka, P.; Beaumont, C.; Richardson, D.R. Function and regulation of transferrin and ferritin. *Semin. Hematol.* **1998**, *35*, 35–54.
144. Toyokuni, S. Iron-induced carcinogenesis: The role of redox regulation. *Free Radic. Biol. Med.* **1996**, *20*, 553–566. [https://doi.org/10.1016/0891-5849\(95\)02111-6](https://doi.org/10.1016/0891-5849(95)02111-6).
145. Huang, X. Iron overload and its association with cancer risk in humans: evidence for iron as a carcinogenic metal. *Mutat. Res.* **2003**, *533*, 153–171. <https://doi.org/10.1016/j.mrfmmm.2003.08.023>.
146. Hou, C.; Hou, Q.; Xie, X.; Wang, H.; Chen, Y.; Lu, T.; Wu, Q.; Liang, Y.; Hu, Y.; Mao, Y. Serum iron status and the risk of breast cancer in the European population: A two-sample Mendelian randomisation study. *Genes Nutr.* **2021**, *16*, 9. <https://doi.org/10.1186/s12263-021-00691-7>.
147. Cui, Y.; Vogt, S.; Olson, N.; Glass, A.G.; Rohan, T.E. Levels of zinc, selenium, calcium, and iron in benign breast tissue and risk of subsequent breast cancer. *Cancer Epidemiol. Biomark. Prev.* **2007**, *16*, 1682–1685. <https://doi.org/10.1158/1055-9965.EPI-07-0187>.
148. Gaur, A.; Collins, H.; Wulaningsih, W.; Holmberg, L.; Garmo, H.; Hammar, N.; Walldius, G.; Jungner, I.; Van Hemelrijck, M. Iron metabolism and risk of cancer in the Swedish AMORIS study. *Cancer Causes Control* **2013**, *24*, 1393–1402. <https://doi.org/10.1007/s10552-013-0219-8>.
149. Quintana Pacheco, D.A.; Sookthai, D.; Graf, M.E.; Schubel, R.; Johnson, T.; Katzke, V.A.; Kaaks, R.; Kuhn, T. Iron status in relation to cancer risk and mortality: Findings from a population-based prospective study. *Int. J. Cancer* **2018**, *143*, 561–569. <https://doi.org/10.1002/ijc.31384>.
150. Sanagoo, A.; Kiani, F.; Saei Gharenaz, M.; Sayehmiri, F.; Koohi, F.; Jouybari, L.; Dousti, M. A systematic review and meta-analysis on the association of serum and tumor tissue iron and risk of breast cancer. *Casp. J. Intern. Med.* **2020**, *11*, 1–11. <https://doi.org/10.22088/cjim.11.1.1>.
151. Von Holle, A.; O'Brien, K.M.; Sandler, D.P.; Janicek, R.; Weinberg, C.R. Association Between Serum Iron Biomarkers and Breast Cancer. *Cancer Epidemiol. Biomark. Prev.* **2021**, *30*, 422–425. <https://doi.org/10.1158/1055-9965.EPI-20-0715>.
152. Von Holle, A.; Thompson, R.L.; O'Brien, K.M.; Sandler, D.P.; Weinberg, C.R. Associations Between Serum Iron Biomarkers and Breast Cancer Tumor Size. *Cancer Res. Commun.* **2024**, *4*, 182–185. <https://doi.org/10.1158/2767-9764.CRC-23-0205>.
153. Kabat, G.C.; Rohan, T.E. Does excess iron play a role in breast carcinogenesis? An unresolved hypothesis. *Cancer Causes Control* **2007**, *18*, 1047–1053. <https://doi.org/10.1007/s10552-007-9058-9>.
154. Ebrahim, S.; Davey Smith, G. Mendelian randomization: Can genetic epidemiology help redress the failures of observational epidemiology? *Hum. Genet.* **2008**, *123*, 15–33. <https://doi.org/10.1007/s00439-007-0448-6>.
155. Smith, G.D.; Ebrahim, S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **2003**, *32*, 1–22. <https://doi.org/10.1093/ije/dyg070>.
156. Hua, X.; Duan, F.; Huang, J.; Bi, X.; Xia, W.; Song, C.; Wang, L.; Jiang, C.; Yuan, Z. A Novel Prognostic Model Based on the Serum Iron Level for Patients With Early-Stage Triple-Negative Breast Cancer. *Front. Cell Dev. Biol.* **2021**, *9*, 777215. <https://doi.org/10.3389/fcell.2021.777215>.
157. Hambidge, M. Biomarkers of trace mineral intake and status. *J. Nutr.* **2003**, *133*, 948S–955S. <https://doi.org/10.1093/jn/133.3.948S>.
158. Fonseca-Nunes, A.; Jakszyn, P.; Agudo, A. Iron and cancer risk—a systematic review and meta-analysis of the epidemiological evidence. *Cancer Epidemiol. Biomark. Prev.* **2014**, *23*, 12–31. <https://doi.org/10.1158/1055-9965.EPI-13-0733>.

159. Torti, S.V.; Manz, D.H.; Paul, B.T.; Blanchette-Farra, N.; Torti, F.M. Iron and Cancer. *Annu. Rev. Nutr.* **2018**, *38*, 97–125. <https://doi.org/10.1146/annurev-nutr-082117-051732>.
160. Calvo, I.; Gonzalez-Rodriguez, M.; Neria, F.; Gallegos, I.; Garcia-Sanchez, L.; Sanchez-Gomez, R.; Perez, S.; Arenas, M.F.; Estevez, L.G. An analysis of the association between breast density and body mass index with breast cancer molecular subtypes in early breast cancer: Data from a Spanish population. *Clin. Transl. Oncol.* **2024**, *26*, 2541–2548. <https://doi.org/10.1007/s12094-024-03469-6>.
161. Johnson, K.S.; Conant, E.F.; Soo, M.S. Molecular Subtypes of Breast Cancer: A Review for Breast Radiologists. *J. Breast Imaging* **2021**, *3*, 12–24. <https://doi.org/10.1093/jbi/wbaa110>.
162. Cao, J.; Li, J.; Zhang, Z.; Qin, G.; Pang, Y.; Wu, M.; Gu, K.; Xu, H. Interaction between body mass index and family history of cancer on the risk of female breast cancer. *Sci. Rep.* **2024**, *14*, 4927. <https://doi.org/10.1038/s41598-024-54762-x>.
163. Jiang, Q.; Hu, H.; Liao, J.; Duan, P.; Li, Z.; Tan, J. Body Mass Index and Breast Cancer-Related Lymphedema: A Retrospective Cohort Study. *J. Surg. Oncol.* **2024**. <https://doi.org/10.1002/jso.27969>.
164. Wada, K.; Kuboyama, K.; Abe, S.K.; Rahman, M.S.; Islam, M.R.; Saito, E.; Nagata, C.; Sawada, N.; Tamakoshi, A.; Shu, X.O.; et al. Body mass index and breast cancer risk in premenopausal and postmenopausal East Asian women: A pooled analysis of 13 cohort studies. *Breast Cancer Res.* **2024**, *26*, 158. <https://doi.org/10.1186/s13058-024-01907-5>.
165. Cubelos-Fernandez, N.; Davila-Batista, V.; Fernandez-Villa, T.; Castano-Vinyals, G.; Perez-Gomez, B.; Amiano, P.; Ardanaz, E.; Delgado Sillero, I.; Llorca, J.; Tardon, G.F.; et al. Burden of postmenopausal breast cancer attributable to excess body weight: Comparative study of body mass index and CUN-BAE in MCC-Spain study. *J. Epidemiol. Community Health* **2024**, *79*, 64–71. <https://doi.org/10.1136/jech-2023-220706>.
166. Holm, J.B.; Baggesen, E.; Cronin-Fenton, D.; Frystyk, J.; Bruun, J.M.; Christiansen, P.; Borgquist, S. Circulating C-reactive protein levels as a prognostic biomarker in breast cancer across body mass index groups. *Sci. Rep.* **2024**, *14*, 14486. <https://doi.org/10.1038/s41598-024-64428-3>.
167. Ogiya, A.; Kimura, K.; Ueno, T.; Iwase, T.; Ohno, S. Time trend of breast cancer-related lymphedema according to body mass index. *Eur. J. Surg. Oncol.* **2024**, *50*, 108350. <https://doi.org/10.1016/j.ejso.2024.108350>.
168. Quartuccio, N.; Ialuna, S.; Pulizzi, S.; D’Oppido, D.; Antoni, M.; Moreci, A.M. The Correlation of Body Mass Index with Risk of Recurrence in Post-Menopausal Women with Breast Cancer Undergoing Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography. *J. Clin. Med.* **2024**, *13*, 1575. <https://doi.org/10.3390/jcm13061575>.
169. Van Cauwenberge, J.; Van Baelen, K.; Maetens, M.; Geukens, T.; Nguyen, H.L.; Nevelsteen, I.; Smeets, A.; Deblander, A.; Neven, P.; Koolen, S.; et al. Reporting on patient’s body mass index (BMI) in recent clinical trials for patients with breast cancer: A systematic review. *Breast Cancer Res.* **2024**, *26*, 81. <https://doi.org/10.1186/s13058-024-01832-7>.
170. Al-Muhtaseb, S.I. Serum and saliva protein levels in females with breast cancer. *Oncol. Lett.* **2014**, *8*, 2752–2756. <https://doi.org/10.3892/ol.2014.2535>.
171. Hamad, A.W.R.; Ibrahim, M.A.; Al-Mohtasib, S.I.; Al-Kobasi, K.N. Comparative Study on Saliva Proteins in Patients of Brain Tumors and Healthy Individuals. *Trends Med. Res.* **2009**, *4*, 16–23.
172. Gao, W.M.; Kuick, R.; Orzechowski, R.P.; Misek, D.E.; Qiu, J.; Greenberg, A.K.; Rom, W.N.; Brenner, D.E.; Omenn, G.S.; Haab, B.B.; et al. Distinctive serum protein profiles involving abundant proteins in lung cancer patients based upon antibody microarray analysis. *BMC Cancer* **2005**, *5*, 110. <https://doi.org/10.1186/1471-2407-5-110>.
173. Halliwell, B. Antioxidants: The basics—what they are and how to evaluate them. *Adv. Pharmacol.* **1997**, *38*, 3–20. [https://doi.org/10.1016/s1054-3589\(08\)60976-x](https://doi.org/10.1016/s1054-3589(08)60976-x).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.