Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INTELLIGENT SYSTEM FOR PRECISION TREATMENT MANAGEMENT FOR PATIENTS WITH BREAST CANCER METASTASIS

SARAH KHROUCH [1], LOUBNA CHERRAT[2], MOSTAFA EZZIYYANI[3]

[1]Mathematic and Applications Laboratory, Computer Science Department Faculty of Sciences and Techniques, Tangier, Abdelmalek Essaâdi University, Morocco.

[2]National School of Commerce and Management, Abdelmalek Essaâdi University, Tangier, Morocco.

[3]Mathematic and Applications Laboratory, Computer Science Department Faculty of Sciences and Techniques, Tangier, Abdelmalek Essaâdi University, Morocco.

E-mail: [1]sarah.khrouch@etu.uae.ac.ma, [2]l.cherrat@uae.ac.ma, [3]mezziyyani@uae.ac.ma

ABSTRACT

This study focuses on developing a novel decision-making system for oncologists to select the most suitable treatment strategy for breast cancer patients. The proposed system, called the Hybrid Model Process of Treatments (HMPT), is designed to assist oncologists by incorporating patient history, experiences, and responses. We analyzed clinical, digital pathology, and genomic data from patients treated with neoadjuvant chemotherapy to construct a strategic approach for metastatic breast cancer treatments. The HMPT model integrates two components: a Predictive Model (PM) using Neural Network (NN) and Logistic Regression (RL) to accurately forecast treatment outcomes, and a Data Augmentation Model (DAM) that generates new data. This newly generated data is evaluated against the Predictive Model (PM) to ensure alignment with established patterns. Results demonstrate that the model can be applied effectively across various breast cancer types, showing potential to expand clinical trial evaluations and test novel hypotheses for metastatic breast cancer patients. The HMPT model offers a revolutionary approach to reducing recurrence rates and enhancing the treatment experience, while also lowering the associated healthcare costs for patients.

Keywords: *Breast Cancer, Treatment management intelligent system, LR, NN, GA.*

1. INTRODUCTION

According to the World Health Organization WHO, breast cancer is the most common cancer among women worldwide, in 2020, the most prevalent type of cancer in terms of new cases was breast cancer, with 2.26 million reported cases [1]. As demonstrated by the latest statistics, the national annual index rate is estimated at more than 50,000 new cases in women. However, predicting the progression of breast cancer is a challenging task due to the complex and dynamic nature of the disease [2].

The 5-year survival rate for locally advanced breast cancer is around 72% for women under the age of 50 and around 64% for women over the age of 50, based on the National Cancer Institute. Locally advanced breast cancer has a higher risk of recurrence compared to earlier stages of breast cancer[3]. However, these statistics have reached a motivating phase, highlighting the need to develop new approaches to overcome the breast cancer. The treatment of locally advanced breast cancer depends on various factors, such as the stage of cancer, the size and location of the tumor, and the overall health of the patient.

Neoadjuvant therapy (NAT) is effective in reducing the risk of breast cancer among patients, emerges as a pivotal strategy, particularly in the context of metastatic breast cancer offers the advantage of downstaging the disease, potentially reducing the extent of surgery in an era of individualization of therapy. NAT regime is likely to reduce the tumor size and can make patients for surgical resection or can make some patients for breast-conserving surgery rather than mastectomy [4]. Most breast cancer treatments, while necessary, often come with high costs and severe side effects. Moreover, factors like age, stage, and individual treatment histories lead to significant variability in the disease's progression. The inherent diversity presents great challenges and difficulty in the accurate analysis of breast cancer to design

individual optimal therapeutic strategies or treatment plan with significant psychological and physical implications.

Clinical trials, which have been recommended by the World Health Organization, remain the gold standard in assessing chemotherapy regimens. This is crucial to test new treatment protocols in an attempt to establish those are even more effective than the current standard. The exploration can be anything from the testing of new pharmaceuticals to devising novel surgical procedures or administering therapies using different techniques. One of the other reasons encompassing these trials is to find a cure that is not only better but also has less harmful effects upon patients' body. For this reason, a clinical trial is classified as research that thoroughly evaluates efficacy and safety in the realm of treatment. In fact, over the years of scientific evolution, clinical trials have been used voluminously in observing pharmacological as well as other types of therapeutic effects according to Ulrikke Lyng Beauchamp (2020).[5]. For many years, clinical trials have been increasingly used to test the effects of drugs and new therapeutic techniques (Ulrikke Lyng Beauchamp, 2020). This represents a transformative improvement in how breast cancer is treated, especially for some types of metastatic breast cancers. These advances allow patient-directed treatment algorithms to cater to therapies based on the specific medical needs and backgrounds of individual patients. Metastatic breast cancer is a good example of where this personalization can work well, as the tumor features and how resistant to therapy the patient has been in their disease history are important considerations for optimal management. Thus, such new strategies of drug management not only improve treatment but also help in prolonging the lives of women fighting against this disease. An individual approach to medicine built on current research and drug development will provide more effective control of, and improved outcomes for breast cancer patients including those with advanced stage disease.

Given the different treatments tap cancer weaknesses, switch is important as cancers continue to grow and respond in distinct ways. Yet, the key to efficient drug control in these cases is a good choice of sequential pairs. Instead, the aim is to look for a range of therapies that not only work well in controlling the cancer but also fit with an individual's overall health and their intended goals for treatment.

Although arguably the most essential phase of this decision-making process, there are no globally agreed-upon guidelines or algorithms on how best to come up with combinations. This absence of guidance is a major obstacle for oncologists who want to deliver the best care they can.

Clinicians often use pre-established chemotherapy regimens when treating a particular type of cancer in real-world practice. They develop such regimens on a robust foundation that considers results from multiple clinical trials, as well based established standards of management (i.e., practice guidelines), current biological knowledge, and collective determinants learned over years. [6],[7],[8].

Thus, a highly complex and resource-consuming process experimental design for clinical trials is required. A main challenge in this endeavor is identifying good drug regimens and doses. This is because the importance of efficacy versus safety can pose a difficult balancing act given that there are so many possible permutations of treatment in this phase.

Also, the majority of clinical trials that inform treatment guidelines are short-term in nature. The main aspect they are examining is whether drugs or distillations of them work quickly enough and if drugs comb input on response. This is helpful, but it usually does not present the full picture of these treatments and how patients benefit in the long-term. Over time, this short-term perspective can prevent a clear sense of how treatment with different agents may impact not only the quality of life but also long-term survival prospects for individual patients.

Therefore, it is essential to optimize the medication treatment plans for metastasis Breast Cancer MBC to elongate patients' lives, to enhance the quality of life, to reduce hospitalization and follow-up treatment costs of those complications.

Significant strides have been made in the field of artificial intelligence (AI) and machine learning (ML), particularly in their application to clinical trials and the development of treatment strategies for breast cancer. These technological advances in breast cancer discovery could transform healthcare and breast cancer research. AI and ML algorithms are increasingly being employed to analyze vast datasets, including genomic information, medical images, and patient records. With such advanced analytical tools, researchers find complex patterns and correlations within the data that are not immediately visible through traditional means. This data-driven paradigm offers a potential for incredible progress in personalized medicine, with improved diagnostics and prognostics to develop an individual treatment strategy. Integrating AI and ML in breast cancer research advances not just the rate but also the

individualization and efficacy of therapeutic interventions overall, turning this into a major disruptor of patient outcome-centric therapy.[9] [10].

Our proposed approach is a synergistic integration of two distinct complementary components. The first part is a Predictive Model (PM), designed to forecast outcomes with unique accuracy. This model serves as the spine of our system, offering reliable predictions that guide the decision-making process. The second part is a sophisticated Data Augmented Model DAM, a dynamic responsible for the generation of new data. Through this iterative process, the genetic algorithm enhances our system's capabilities by refining its predictions and ensuring that the model remains robust and adaptable to evolving scenarios. By exploiting the power of real-time data analysis and personalized treatment modeling, this methodology allows for a more nuanced understanding of individual patient responses and disease progression. This model is aimed at maximizing the efficacy of the treatment while minimizing adverse effects, thus ensuring a personalized and adaptive treatment plan. By employing this new Predictive Model (PM), healthcare professionals can make more informed and precise decisions, enhancing the overall effectiveness of treatments for breast cancer patients.

Our Model provide a dynamic overview of breast cancer's evolution, offering insights that can adapt to changing conditions. Together, these advanced techniques promise to enhance the precision and effectiveness of breast cancer treatments, surpassing the limitations of traditional, one-size-fits-all clinical trials.

The goal is to create a model that not only predicts outcomes but also guides decision-making towards optimal results.

The following is the organization of this paper: Section 2 will introduce the methods and findings of prior research related to the clinical trials of breast cancer. Section 3 will outline the methodology that we propose for our research. Section 4 will provide a detailed presentation and explanation of the experiment results. Finally, Section 5 will summarize the paper and provide concluding remarks and future work.

2. RELATED WORK

Manuscripts must be in English (all figures and text) and prepared on Letter size paper (8.5 X 11 inches) in two column-format with 1.3 margins from top and .6 from bottom, and 1.25cm from left and right, leaving a gutter width of 0.2 between columns.

Several research studies have been conducted to evaluate treatment strategies for various diseases by analyzing patient treatment histories and employing sophisticated machine learning techniques. These studies focus on key outcomes such as survival rates, recurrence of the condition, and the response to treatment to assess the efficacy of different treatment modalities. By examining the historical data of patients' treatment journeys, researchers can identify patterns and correlations that inform the effectiveness of specific treatments. Survival rates provide insight into the long-term impact of treatment strategies, while recurrence rates help evaluate the potential for the disease to return after treatment. Treatment response, measured through various clinical and biomarker assessments, offers immediate feedback on the effectiveness of a treatment regimen. Together, these significant indicators are crucial for evaluating the success of treatment strategies, guiding medical professionals in optimizing patient care and advancing the development of more effective treatments [11], [12].

Through the application of machine learning algorithms, they have successfully identified biomarkers, prognostic indicators, and treatment responses specific to individual patients. These studies not only exemplify the power of clinical trials and precision medicine but also showcase the potential of data-driven approaches to revolutionize healthcare decision-making. M. Rakhshaninejad [13], Lopez-Perez L[14].

AI technologies excel in processing this unstructured data, efficiently extracting crucial and pertinent information. They are adept at mining key details, interpreting underlying meanings, and organizing the data into a comprehensible and structured format. This refined data presentation significantly aids data analysts in their review and interpretation processes, enhancing the overall efficiency and accuracy of clinical trial analysis.

There are other advanced clinical trials, such as Adaptive designs, that have gained prominence in the realm of clinical trials and drug development. Unlike traditional trial designs, adaptive designs leverage accumulating data to modify ongoing trials in real-time, while still preserving the trial's integrity and validity [15].

Several impactful research studies have explored various approaches in the application of

artificial intelligence (AI) to optimize decision-making and treatment selection in healthcare.

We collected some studies applied AI in medical treatment is provided in Table1.

Table 1: An overview of the applications of AI and ML in medical treatment.

Studies	Content
Mehrnaz Abdollahian [16]	This study presents a model designed to yield optimal intervention strategies for BRCA1/2 mutation carriers with ages between 30 and 65 and any prior intervention history.
Shan Liu [17]	This work, developed a framework to guide optimal treatment decisions for a deteriorating chronic disease HCV, using a discrete-time, finite-horizon Markov decision process, they showed that the optimal treatment decision is more likely to be to accept currently available treatment despite expectations for future treatment improvement for patients who have a high-risk history, who are older, or who have more comorbidities.
Nagesh Shukla[18]	This paper aims to construct a comprehensive and a robust data analysis model designed to enhance their understanding of breast cancer survivability, particularly in cases where data may be incomplete. The objectives include gaining deeper insights into the factors that influence patient survival, as well as identifying and categorizing groups of patients with shared characteristics. Through this approach, the model seeks to improve predictions and outcomes for individuals battling breast cancer, facilitating more tailored and effective treatment strategies.
Sang-Ho Oh [19]	This study developed a medical treatment recommendation system for diabetes using Korean EHRs along with the Markov decision process (MDP) their results showed that MDP recommendations can maintain better health conditions by delaying the occurrence of diabetic complications. The patients who followed MDP recommendations were able to delay the onset of complications longer than those who did not follow MDP recommendations and they proved that their MDP recommendation system could help doctors prescribe appropriate diabetes medications.
Frank P. Y. Lin[20]	This work has developed a machine-learning model specifically aimed at predicting the outcomes of MDT deliberations regarding adjuvant treatments for breast

	cancer. This model seeks to standardize medical decision-making processes by encapsulating the collective wisdom of MDT meetings, thus bridging the gap between centers and ensuring consistent treatment approaches.
Rachel Choi[21]	This study proposes a model that has the ability to predict pathologic complete response (PCR) before starting treatment would enable healthcare providers and patients to concentrate on the most promising therapeutic options, thereby reducing exposure to potentially unnecessary chemotherapy-related toxicities. Using a deep neural network algorithm, this algorithm would be trained on breast MRI images acquired before the initiation of treatment, offering a novel approach to personalized therapy plans based on predicted treatment responses.
Woojae Kim[22]	This study develops an innovative prognostic model utilizing a support vector machine (SVM) to predict the likelihood of breast cancer recurrence within five years following surgery, specifically within the Korean population. Furthermore, the study sought to evaluate the predictive accuracy of this new model by comparing its performance against existing models. This comparison is intended to highlight the potential improvements in prediction accuracy and reliability, offering a more tailored approach to post-surgical management and monitoring in breast cancer patients.
Yusong Wang[23]	This study used the machine learning algorithm and logistic regression to select the features to improve the neoadjuvant therapy efficacy in patients with locally advanced cancer to reduce tumor and prolong survival for HER2 positive and triple negative breast cancer. And this work investigates how variations in peripheral immune markers correlate with therapeutic responses throughout neoadjuvant therapy (NAT), shedding light on their predictive value for treatment efficacy.
S. Liu[24]	This study explores predictive factors for neoadjuvant therapy (NAT) outcomes in breast cancer, focusing on achieving pathological complete response (pCR). It introduces discordant pathological complete response (DpCR), an

	intermediate response category. Using data from 789 patients, predictive models based on clinicopathologic features and inflammatory indexes (HALP, P53, FAR, Molecular Subtype) were developed. The models demonstrated high accuracy in predicting traditional non-pCR (T-NpCR), total-pCR (TpCR), and DpCR, with strong clinical utility shown through decision curve analysis (DCA).
Jian. Chen[25]	This study focuses on predicting responses to neoadjuvant therapy (NAT) in breast cancer (BC) using RNA-seq data. The authors developed two machine learning models, the Ipredictor (immune gene-based) and ICpredictor (immune gene and receptor status-based), achieving high predictive accuracy for pathological complete response (pCR). The models were validated on both RNA-seq and microarray platforms, with results showing a strong correlation between immune profiles and pCR. These models aim to improve the accuracy of NAT response prediction, contributing to more personalized treatment for BC patients.
Yusong. Wang[26]	This study investigates the impact of the tumor immune microenvironment on the response to neoadjuvant therapy (NAT) in HER2-positive breast cancer. In 295 patients, tumors were classified as "immune-rich" or "immune-poor." The "immune-rich" phenotype was strongly associated with achieving pathological complete response (pCR). Ten genes were identified as correlated with both pCR and immune status. A generalized non-linear predictive model was developed, showing strong predictive power across internal, external, and clinical validation sets, with high accuracy, indicating potential for improving treatment outcomes in HER2+ BC patients

3. PROPOSED MODEL

Our work proposes a new approach that helps the clinicals and the oncologists to choose the appropriate therapeutic strategies. Our proposed model is, known as the Hybrid Model Process of Treatment (HMPT), consists of two essential components. The first part is the prediction model, distinguished for its high performance and remarkable accuracy, derived from observed data. This predictive segment excels in forecasting outcomes and trends, providing a robust foundation for informed decision-making in the treatment process.

The second component of HMPT involves the application of genetic algorithm techniques for data augmentation.

The proposed hybrid model architecture is shown in figure. 1 below: this figure illustrates the interaction between the Predictive Model (PM) and the Data Augmented Model (DAM) within the system.

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

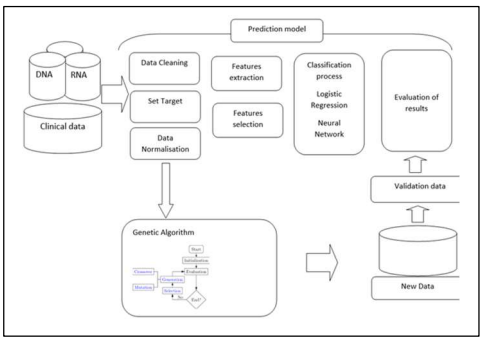


Figure 1: Architecture of Hybrid Model Process of Treatment

3.1 Data Source

This was a retrospective analysis of a series of patients who followed neoadjuvant treatment for metastatic breast cancer. Our study utilized data extracted from the article [27], the data from this source was chosen due to its relevance, comprehensiveness, and alignment with the study's objective of developing an intelligent dynamic graph system for precision drug management in patients with breast cancer metastasis. The data presented in Figure. 2.

Figure 2: Sample for breast cancer dataset

4. METHODOLOGY

Our methodology is divided into two key components: a Predictive Model (PM) and a Data Augmented Model (DAM).

Before training, the data, we start with data acquisition, followed by pre-processing, which includes the following steps: data cleaning, normalization, target setting and features extraction.

4.1 Predictive Model justification

We employed two algorithms: Neural Networks (NN) and Logistic Regression (RL), chosen for their ability to handle non-linear relationships and binary classification tasks, respectively.

After applying advanced machine learning algorithms, Logistic Regression (LR) and Neural Network (NR), on breast cancer data, we used several performance metrics to evaluate and compare the models to identify the best algorithm for breast cancer treatment response prediction. Specifically, we employed the Confusion Matrix, Accuracy, Precision, Sensitivity, F1 Score, and AUC as key metrics. The Confusion Matrix is a crucial tool in measuring the performance of classification problems, where the output can belong to two or more classes. It is a table with two dimensions, "Actual" and "Predicted," each containing values for "True Positives (TP)," "True Negatives (TN)," "False Positives (FP)," and "False Negatives (FN)." Among these metrics, Accuracy is the most common, defined as the number of correct predictions made as a ratio of all predictions made. Precision: The ratio of correctly predicted positive observations to the total predicted positives. Recall (Sensitivity): The ratio of correctly predicted positive observations to all observations in the actual class. F1 Score: The harmonic mean of precision and recall. By analyzing these metrics, we were able to determine the more effective model for predicting treatment response in breast cancer patients.

Table 2 below summarized the performance of the PM, neural networks and logistic regression. As we can see from the results in the table, the accuracy of both classifiers is close and consistently high, indicating strong performance in the PM.

Table 2: ML Performance.

	Accuracy	F1 Score	Error
Logistic Regression (LR)	1.000	1.000	0.0
Neural Network (NR)	0.980	0.990	0.02

Table 3 and Table 4 below present the confusion matrices and the calculated performance measures of the logistic regression and neural networks.

Table 3: Confusion matrix of logistic regression.

Logistic Regression (LR)	True Positive	False Negative	Accuracy	Precision
	49	0	1.000	1.000
	False Positive	True Negative	Recall	F1 Score
	0	1	1.000	1.000

From the results in the table, we can see that the precision is 1.00, the sensitivity is 0.98, and the F1 measure is 0.99, of LR, indicating that the model's performance is exceptionally high. These metrics demonstrate that the machine learning models are effectively capturing the true positives while maintaining a low rate of false positives and negatives. That is why the accuracy of both machine learning models is better. Figure. 3 presents the ROC curve of the Logistic Regression (LR) machine learning algorithm, which is an important metric for evaluating classifier performance. As we can see, the ROC curve for LR indicates that it is a near-perfect classifier, further confirming the model's high accuracy and reliability in making predictions.

Table 4: Confusion matrix of NR.

Neural Network (NR)	True Positive	False Negative	Accuracy	Precision
	48	1	0.980	1.000
	False Positive	True Negative	Recall	F1 Score
	0	1	0.980	0.990

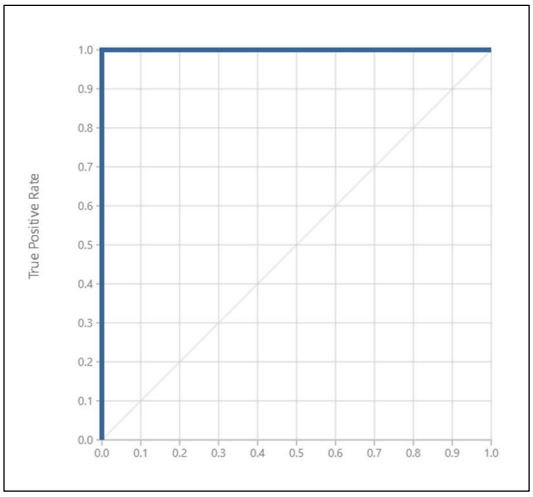


Figure 3: Roc curve for LR

This suggests that our PM is well-calibrated and can make reliable predictions. The minimal difference in accuracy between the classifiers further reinforces the model's robustness, as it performs consistently well regardless of the algorithm used. Given these results, we can confidently assert that our Predictive Model (PM) is good enough to perform effectively in other scenarios, delivering accurate and dependable outcomes.
Classifying our data after the prediction we noticed that the accuracy was perfect, then we suggested implementing the genetic data to ensure and validate the therapeutic strategy.

After executing the initial phase of prediction, our analysis revealed that the accuracy of the results was higher. This result was the crucial turning point, prompting us to explore additional avenues to enhance the model's performance. To address this challenge, we propose the integration of genetic data into our framework.
We aim to significantly prove the prediction accuracy and tailor the treatment or therapeutic strategies for breast cancer patients more effectively. This enhancement will be achieved through the genetic algorithm processes of selection, mutation, and crossover, which are designed to generate random and new individuals. These new individuals are expected to validate the model's accuracy, offering a more robust and effective approach to breast cancer treatment and care. By making these methodological advancements, our model is poised to provide more precise, personalized therapeutic strategies, fundamentally transforming the landscape of breast cancer treatment.

4.2 Appling genetic algorithm

A genetic algorithm is a search heuristic and evolutionary algorithm that is inspired by Charles Darwin's theory of natural evolution. This method reflects the process of natural selection, where the fittest individuals are selected for reproduction to produce offspring for the next generation. It is used to solve optimization and search problems.
A genetic algorithm was used to generate a synthetic data
To overcome the challenge of limited datasets, a genetic algorithm was implemented to generate synthetic data.
The application of a genetic algorithm in this manner allows for the analysis and optimization of treatment strategies based on the complex interplay of clinical and genetic data. By iterating through generations of patient data, the algorithm can help identify patterns

and correlations that might inform more effective and personalized treatment approaches for cancer patients.

To adapt the concept of a genetic algorithm for use in the oncology context, particularly for analyzing clinical, genetic, and treatment data, we need to redefine the essential elements of genetic algorithms - Generation, Population, Individual, Chromosomes, and Genes in terms relevant to oncology:

Generation: In this context, a 'Generation' refers to a specific set of patient data collected over a defined time. Each generation represents a unique cohort of patients, along with their associated clinical, genetic, and treatment data.

Population: The 'Population' comprises various patient cases within a particular generation. These cases are characterized by their clinical profiles, genetic and treatment responses. For instance, the initial population might consist of all patients treated in a specific year or those diagnosed with a particular type of cancer.

Individual: Each 'Individual' within the population represents a single patient case. An individual's data includes their specific clinical characteristics, genetic information, and treatment regimen. In this model, each patient's data set is analogous to an individual in a traditional genetic algorithm.

Chromosome: The 'Chromosome' in this setting is a type of data within each patient's case, such as demographics data, clinical data, DNA, RNA, treatment, and digital pathology, ect.

Gene: 'Genes' correspond to specific elements within each 'chromosome'. In the context of patient data, (like age, or stage), in chromosome treatment (number of chemeotheray).

Figure 4, represents the genetic algorithm implementation model adapted for patient data analysis in the context of metastatic breast cancer treatment.

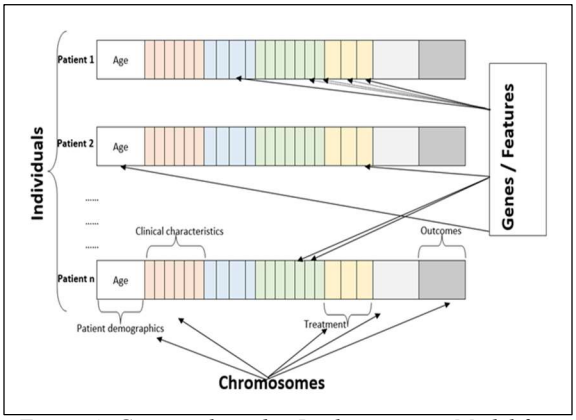


Figure 4: Genetic algorithm Implementation Model for Patient Data

Data Augmentation Model

Data augmentation using genetic algorithms is crucial in enhancing machine learning models, particularly in fields with limited datasets. Genetic algorithms simulate the process of natural selection to generate new, synthetic data points by combining attributes of existing data. This method mimics biological evolution, where "parent" data points are selected and combined to create "offspring" data points with shared and diverse characteristics. The main benefits of this approach include increased dataset diversity, which helps reduce overfitting, and improved model robustness. By augmenting the data, genetic algorithms enable models to learn more comprehensive and generalized patterns, leading to better performance on unseen data. This technique is instrumental in medical research, image processing, and other domains where acquiring large amounts of labeled data is challenging and expensive. Overall, data augmentation via genetic algorithms provides a powerful tool for enriching datasets, ultimately enhancing the efficacy and reliability of machine learning models.

This strategic integration enables the construction of more diversified therapeutic strategies and the formulation of personalized therapy plans, through repetitive application of the mutation, crossover and selection operators. Benefiting evolutionary principles, the genetic algorithm enhances adaptability and optimization in treatment approaches, contributing to the overall efficacy and customization of patient care within the model.

Following the implementation of the genetic algorithm, a crucial step in our methodology involves validating the new data within the Predictive Model (PM) (first part). This validation process ensures that the insights generated through the Data Augmented Model (DAM) seamlessly integrate into the PM. By validating the performance with the new data, we continuously refine and enhance the accuracy and adaptability of the HMPT framework, ensuring its reliability in real-world scenarios.

The key steps in a genetic algorithm include:

- Initial Population: The algorithm begins with a set of individuals, known as the population. Each individual is a solution to the problem you want to solve and is represented by a set of parameters (known as genes).
- Selection: The algorithm selects the fittest individuals from the current population. These

individuals are then used to produce the next generation.

• Crossover (Reproduction): Pairs of individuals (parents) are crossed over to produce new offspring (children), which inherit some of the genes from each parent. This process is meant to mimic biological reproduction and genetic crossover.

• Mutation: To maintain genetic diversity within the population and to avoid premature convergence, the algorithm introduces random changes to some individuals in the population. This step is akin to biological mutations.

• Fitness Function: Each individual in the population is evaluated using a fitness function. This function determines how fit or how good the solution is at solving the problem.

• Termination: This process is repeated over multiple generations. The algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

Data encoding in genetic algorithm

The dataset comprises 168 records, each representing an individual patient's clinical profile and treatment history. Key variables include patients' demographics, clinical, treatment Genetic, and outcomes as shown in the table 5 for data encoding.

Table 5: Table of data encoding

Chromosomes	Attributes	symbol	Encoding
Patient demographics	Age at diagnostic	A1 A2 A3 A4 A5 A6	[20,30] = 000 [30,40] = 001 [40,50] = 010 [50,60] = 011 [60,70] = 100 [70,80] = 101
Clinical characteristics	stage	T1, T2, T3, T4	T1=00, T2=01 T3=10, T4=11
	histology	H1, H2, H3	H1=00 H2=01 H3=10
	ER status	ER	ER=0, ER=1
	HER2 status	HER2	HER2 = 0 HER2 = 1
	Lymph node involvement	LNI1, LNI2	LNI1= 0, LNI2 = 1
Genetic Data	DNA	M1 M2	M1 = 0 M2 = 1
	RNA	S1 S2	S1 = 0 S2 = 1
	Digital Pathology	D1 D2	D1=1 D2=0
Treatment	Number of chemotherapy cycles	D1	D1 = 00 D2 = 01 D3= 10
	Drugs regimens	D2	
	Anti-her2 therapy	D3	
Outcomes	Response "yes or no"	Y1, Y2	Y1= 0 , Y2= 1

Chromosomes / Genes

process ensures that the data is accurately interpreted and categorized, laying a solid foundation for the subsequent phase. The second part of our model encompasses the implementation of a genetic algorithm called Data Augmented Model DAM. This sophisticated approach powers evolutionary principles to optimize solutions, enabling the model to adapt and evolve in response to dynamic datasets and complex problem-solving scenarios. The synergy between these two components "predictive analytics and data augmented" positions our model at the forefront of innovation in data analysis and algorithmic development.

Fitness function

Before proceeding to the selection phase, it is essential to define the fitness function, which serves as a metric for assessing the quality of each individual in our model. In this case, the fitness function is directly represented by the accuracy of the model for each individual (patient), expressed as:

$$f(I) = Accuracy(I)$$

In our model, this fitness function provides a quantitative evaluation of the quality of each individual pattern within the dataset. By using accuracy as the fitness measure, we prioritize individuals (patients) whose patterns result in higher accuracy, indicating better model performance. Higher fitness values correspond to more accurate predictions, and as the fitness increases, the individual patterns reflect a more effective representation of patients with superior outcomes. This approach helps to enhance the overall accuracy and efficiency of our patient-specific evaluations, ensuring that the best-performing patterns are selected and emphasized.

Individual fitness calculation:

The individual fitness calculation is a critical step in the optimization process, where each patient's fitness is evaluated based on the accuracy of the Predictive Model (PM). In our proposed model, fitness is calculated as the accuracy of the predictions made by the machine learning algorithms, specifically Logistic Regression (LR) and Neural Networks (NN), for each patient.

This metric ensures that individuals (patients) whose predicted treatment responses are highly accurate are considered fit. The fitness values are used to guide the selection process in the genetic algorithm, where patients with higher

fitness values are selected for crossover and mutation processes. This allows the model to evolve towards more optimized treatment strategies, enhancing the overall treatment effectiveness and patient outcomes. The individual fitness scores for the patients in our study ranged from 0.67 to 0.99, demonstrating variability in model performance across different patient profiles. Table 6 showed the fitness of the individual.

Table 6: Fitness Table

Individual	Accuracy	F(I)
Patient 1	0.88	0.88
Patient 2	0.9	0.9
Patient 3	0.76	0.76
Patient 4	0.99	0.99
Patient 5	0.67	0.67
Patient 6	0.87	0.87

Selection

During the application phase, our approach involves setting the fitness threshold as the average of the fitness scores across all individuals (patients) within the population. Consequently, individuals surpassing this threshold are considered the fittest and are retained for subsequent processes such as reproduction through crossover and mutation. These selected individuals play a pivotal role in shaping the genetic makeup of the next generation.

In our specific case, we adopt an initial strategy to ensure equal opportunities for all individuals within the population. The process begins with the application of mutation and crossover during the first iteration. This initial step injects variability into the population and lays the foundation for the subsequent selection process, which is initiated from the first-generation offspring. This sequential approach aims to promote diversity and adaptability within the population, enhancing the overall evolutionary dynamics of the genetic algorithm. See figure 5 illustrated the pseudo code of selection.

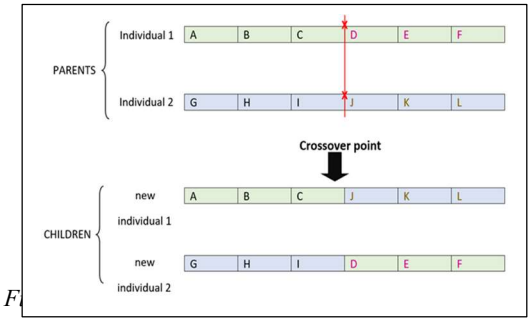
```
Function SelectBestIndividuals(InitialPopulation, NumberOfSelections) Returns

Population
    Define selectedPopulation as new Population
    Initialize selectedPopulation to contain NumberOfSelections individuals
    Sort InitialPopulation by fitnessScore in descending order
    For i from 0 to NumberOfSelections - 1 do
        selectedPopulation.patients[i] = InitialPopulation.patients[i]
    End For
    Return selectedPopulation

End Function
```

Figure 5: Selection Pseudo code

Crossover: Initially, we will employ the one-point method, selecting a crossing point at random. This chosen point will serve as the marker for exchanging traits between the parents, facilitating the creation of offspring. For each gene in each chromosome, the function determines whether the gene index is before or after the crossover point. If the index is before the crossover point, child1 inherits the corresponding gene from parent1 and child2 from parent2. If the index is after the crossover point, child1 inherits the corresponding gene from parent2 and child2 from parent1. The figure 6 shows the one-point crossover, shows how the chromosomes of two individuals are combined to create new offspring. In this process, we take chromosomes D, E and F of patient 1 from individual 1 and the chromosomes J, K and L of patient 2 from individual 2. The crossover occurs at one designated point along the chromosome sequences, effectively swapping segments between the two parents. This results in the formation of two new individuals: new individual 1 and new individual 2.



F

Then we created the two new individuals each having a unique combination of modules from their parents.

Similarly, the 2-point crossover method is employed, enabling us to assess and compare the outcomes derived from each of these techniques. See figure 7 showed the pseudo code of crossover.

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

```
Function Crossover (ParentPopulation) Returns Population
    Initialize newPopulation as empty
    Determine numberOfPairs as size of ParentPopulation / 2

    For i from 1 to numberOfPairs do
        Initialize child as new Patient
        Select parent1 and parent2 from ParentPopulation randomly ensuring no
repetition
        child.name = Concatenate("Child of ", parent1.name, " and ", parent2.name)
        child.demographics.ageAtDiagnostic =
RandomChoice(parent1.demographics.ageAtDiagnostic,
parent2.demographics.ageAtDiagnostic)
        child.clinical.stage = RandomChoice(parent1.clinical.stage,
parent2.clinical.stage)
        child.clinical.histology = RandomChoice(parent1.clinical.histology,
parent2.clinical.histology)
        child.clinical.erStatus = RandomChoice(parent1.clinical.erStatus,
parent2.clinical.erStatus)
        child.clinical.her2Status = RandomChoice(parent1.clinical.her2Status,
parent2.clinical.her2Status)
        child.clinical.lymphNode = RandomChoice(parent1.clinical.lymphNode,
parent2.clinical.lymphNode)
        child.clinical.tumorSize = (parent1.clinical.tumorSize +
parent2.clinical.tumorSize) / 2
        child.genetic.dna = MergeAttributes(parent1.genetic.dna, parent2.genetic.dna)
        child.genetic.rna = MergeAttributes(parent1.genetic.rna, parent2.genetic.rna)
        child.genetic.digitalPathology =
RandomChoice(parent1.genetic.digitalPathology, parent2.genetic.digitalPathology)
        child.treatment.chemotherapyCycles =
Average(parent1.treatment.chemotherapyCycles, parent2.treatment.chemotherapyCycles)
        child.treatment.antiHer2 = RandomChoice(parent1.treatment.antiHer2,
parent2.treatment.antiHer2)
        child.treatment.surgery = RandomChoice(parent1.treatment.surgery,
parent2.treatment.surgery)
        child.outcomes.pcr = RandomChoice(parent1.outcomes.pcr, parent2.outcomes.pcr)
        Add child to newPopulation
    End For
    Return newPopulation
End Function
Function RandomChoice(value1, value2)
    If random() < 0.5 then
        Return value1
    Else
        Return value2
    End If
End Function
Function MergeAttributes(attribute1, attribute2)
    Return SomeFormOfCombination(attribute1, attribute2)
End Function
Function Average(value1, value2)
    Return (value1 + value2) / 2
End Function
        Add child to newPopulation
    End For
        Return newPopulation
```

Figure 7: Crossover pseudo code

Crossover two point:

As shown at the figure 8 below the two-point crossover, demonstrates how the chromosomes of two individuals are combined to create new offspring. In this process, we take chromosomes C and D of patient 1 from individual 1 and the chromosomes I and J of patient 2 from individual 2. The crossover occurs at two designated points along the chromosome sequences, effectively swapping segments between the two parents. This results in the formation of two new individuals: new individual 1 and new individual 2.

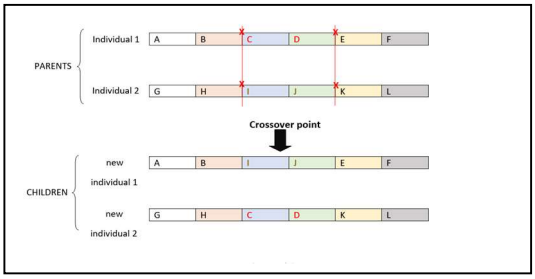


Figure 8: Two-point crossover in genetic algorithm

Mutation

After the crossover, the next step is mutation, where random alteration are introduced to the genetic makeup of the new individuals, as illustrated at Figure 9. By incorporating both crossover and mutation techniques, we enhance genetic diversity within the population, which is crucial improving the effectiveness of our genetic algorithm in optimizing treatment strategies.

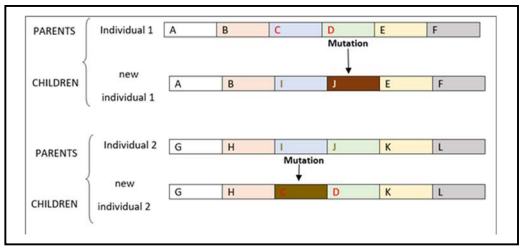


Figure 9: Schema of mutation

The former involves altering genetic information randomly within the same individual, while the latter pertains to genetic modifications occurring between different individuals in the population. As we can see below figure 10 Mutation pseudo code.

```
function Mutate (Individual)
    Define mutationRate = 0.05
    If Random () < mutationRate Then
        Individual.demographics.ageAtDiagnostic += RandomNumber(-2, 2)
    If Random() < mutationRate Then
        Individual.clinical.tumorSize += RandomNumber(-0.5, 0.5)
        If Random() < mutationRate Then
        Randomize Attribute Individual.clinical.stage from predefined list of stages
    If Random() < mutationRate Then
        Modify Individual.genetic.dna by randomly altering string characters
 If Random() < mutationRate Then
        Toggle Individual.treatment.chemotherapyCycles between predefined set of cycle
counts
    Return Individual
End Function
```

Figure 10: Mutation Pseudo Code

Mutation rate

In the context of genetic algorithms (GAs), the mutation rate is a crucial parameter that influences the algorithm's ability to explore the solution space and avoid premature convergence to local optima. After careful consideration and review of empirical studies, we have determined that a mutation rate of 0.01 (1%) is optimal for our problem. This decision is based on several key factors:

Balance Between Exploration and Exploitation:

A mutation rate of 0.01 strikes a balance between exploration (searching new areas of the solution space) and exploitation (refining existing solutions). This balance is essential to ensure that the algorithm can discover innovative solutions while improving the quality of the existing ones.

Maintaining Genetic Diversity:

Genetic diversity is vital for the robustness and adaptability of the population. A mutation rate of 0.01 helps maintain this diversity by introducing small variations in the chromosomes. This prevents the population from becoming too homogenous, which can

lead to stagnation and premature convergence to suboptimal solutions Cervantes J [28].

Empirical Success in Similar Problems:

Numerous studies and practical implementations of genetic algorithms across various domains have found that a mutation rate around 0.01 works well. These studies provide a solid empirical foundation that supports the effectiveness of this mutation rate in maintaining a good balance between solution quality and diversity. [29].

Preventing Overfitting and Enhancing Generalization:

In problems where overfitting is a concern, a mutation rate of 0.01 ensures that the algorithm does not become overly specialized to the current population's characteristics. By continuously introducing small mutations, the algorithm can better generalize to new data and avoid overfitting to specific patterns.

Algorithm Stability and Convergence:

High mutation rates can introduce too much randomness, leading to instability and slower convergence rates. Conversely, very low mutation rates may not introduce enough variation, causing the algorithm to get stuck in local optima. A mutation rate of 0.01 provides a stable convergence path while still allowing the algorithm to escape local optima through gradual exploration.

Parameter Tuning and Problem-Specific Considerations:

While a mutation rate of 0.01 is a good starting point, it is also flexible enough to be fine-tuned based on specific problem requirements. This adaptability makes it a practical choice for a wide range of problems, allowing further optimization as more is learned about the problem space.

Choosing a mutation rate of 0.01 is based on achieving a delicate balance between exploration and exploitation, maintaining genetic diversity, leveraging empirical evidence, and ensuring algorithm stability. This rate is generally effective for a broad spectrum of problems, providing a robust and adaptable foundation for optimizing the performance of genetic algorithms.

Validation of augmented data:

The newly acquired data points were validated by an expert to ensure their coherence and consistency. This expert validation helped to confirm the accuracy of the data, ensuring it aligned with established knowledge and clinical expectations, which further enhanced the reliability of the dataset. As we see the figure 11 below the sample of generated data by DAM.

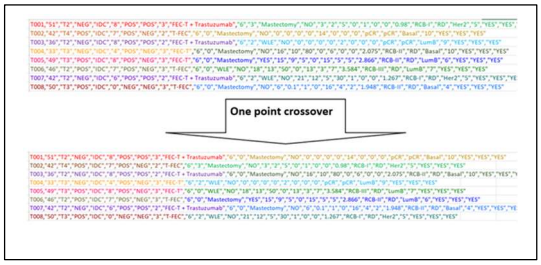


Figure 11: Sample of generated data

5. RESULTS & DISCUSSION

After implementing the genetic algorithm, which involved processes of genetic crossover and mutation, we successfully generated a new generation of data. This new dataset was meticulously compared with the old data to avoid any duplication, ensuring that each individual was unique and contributed to the diversity of the population. Subsequently, we evaluated the fitness of each individual within the new generation, comparing these values against the average fitness of the individuals from the previous dataset. This comparative analysis allowed us to identify and select individuals with higher fitness levels. By prioritizing these individuals, we aimed to enhance the overall quality and performance of the population in the next generations, driving the evolutionary process forward with individuals who exhibit superior traits and characteristics. Subsequently, we proceeded to validate these newly created individuals within our performance Predictive Model PM. This validation process was critical, as it ensured that the selected individuals not only demonstrated higher fitness levels but also exhibited predictive performance metrics that aligned with our model's expectations. Table 7 shows the performance of NN and LR of the Data Augmented Model DAM.

Table 7: ML PERFORMANCE

	Accuracy	F1 Score
Logistic Regression (LR)	0.990	0.990
Neural Network (NR)	0.990	0.990

In this study, we observed that the initial dataset had a 22 percent incidence of PCR=0, indicating that 20 percent of the patients achieved a pathological complete response, where 68% had not achieve it. This baseline statistic is crucial for understanding the current effectiveness of existing treatment strategies. However, when we applied our data augmentation model DAM to generate new data, we observed a marked increase in the percentage of cases with PCR=0, rising dramatically to 60 percent.

This substantial improvement highlights the potential of genetic algorithms in optimizing treatment strategies. Genetic algorithms, which are inspired by the process of natural selection, iteratively evolve solutions to complex problems through processes such as selection, crossover, and mutation. By simulating these evolutionary processes, the algorithm explores a vast space of potential treatment strategies and converges on patterns that are more effective than those derived from traditional methods.

The increase to 60 percent PCR=0 in the generated data suggests that the genetic algorithm can uncover novel treatment strategies that significantly improve patient outcomes. The validation of these strategies by an expert oncologist further reinforces their clinical relevance and potential applicability in real-world scenarios. This expert validation is a critical step, ensuring that the proposed strategies are not only statistically sound but also practically viable and aligned with current medical knowledge and practices. The figure 12 shows the graph of fitness value from generation 9.

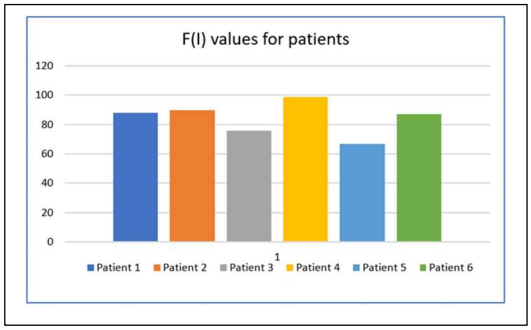


Figure 12: Trend of F(I) Values Among Patients

7. CONCLUSION

This research introduced the Hybrid Model Process of Treatment (HMPT), an innovative and unique approach to breast cancer treatment management, integrating predictive analytics with data augmented. The study demonstrated that HMPT can accurately predict treatment outcomes and optimize therapy strategies for metastatic breast cancer patients. By generating additional datasets through genetic algorithms, this approach holds significant value for medical research, particularly in oncology, where patient data is often scarce and highly individualized.

This technique minimizes the dependence on large-scale clinical trials, enabling advancements in treatment through smaller, more targeted datasets. By leveraging advanced computational methods, such as machine learning models like Neural Networks (NN) and Regression Logistic (RL), this study has shown improved precision in enhancing predictive accuracy.

Moreover, the integration of advanced algorithms offers a revolutionary solution to data-driven challenges in medical research, with the potential to significantly increase the rates of complete pathological responses. This approach represents a substantial leap forward in personalized medicine, offering tailored treatment strategies that maximize patient outcomes and mark a significant advancement in breast cancer care.

The next step in our study is to assist oncologists in incorporating our system into their treatment workflows, making it a practical tool for enhancing patient care and optimizing treatment decisions.

6. FUTERE WORK

In our coming endeavors, we plan to advance and refine our Predictive Model PM to ensure greater reliability and effectiveness. The primary objective is to harness advanced technologies and methodologies to elevate the model's capabilities. This entails incorporating state-of-the-art algorithms, leveraging the latest advancements in artificial intelligence and machine learning, and embracing cutting-edge data processing techniques. By doing so, we aim to enhance the model's predictive accuracy and robustness, enabling it to provide more precise and valuable insights. Additionally, our commitment extends to continuous improvement, staying abreast of emerging

technologies and methodologies to ensure that our Predictive Model remains at the forefront of innovation in its quest to contribute significantly to the oncology field.

In terms of clinical applicability, the next practical steps will involve supporting oncologists in integrating this system into their treatment workflows. By providing actionable insights directly within clinical practice, this system aims to optimize treatment decisions and improve patient outcomes. The focus will be on ensuring seamless adoption, providing training, and customizing the model to meet the specific needs of individual healthcare settings.

REFERENCES:

[1] Breast cancer n.d. https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed January 8, 2024).

[2] Khrouch S, Ezziyyani M, Ezziyyani M. Decision System for the Selection of the Best Therapeutic Protocol for Breast Cancer Based on Advanced Data-Mining: A Survey. In: Ezziyyani M, editor. Advanced Intelligent Systems for Sustainable Development (AI2SD'2018), vol. 914, Cham: Springer International Publishing; 2019, p. 120–8. https://doi.org/10.1007/978-3-030-11884-6_10.

[3] Cancer genome research and precision medicine - NCI 2022. https://www.cancer.gov/ccg/research/cancer-genomics-overview (accessed January 8, 2024).

[4] Wang Y, Wang M, Yu K, Xu S, Qiu P, Lyu Z, et al. A machine learning model to predict efficacy of neoadjuvant therapy in breast cancer based on dynamic changes in systemic immunity. Cancer Biol Med 2023;20:218–28. https://doi.org/10.20892/j.

[5] The Use of Wearables in Clinical Trials During Cancer Treatment: Systematic Review - PMC n.d. https://pmc.ncbi.nlm.nih.gov/articles/PMC7688381/ (accessed October 19, 2024). https://doi.org/10.2196/22006

[6] Scherrer A, Schwidde I, Dinges A, Rüdiger P, Kümmel S, Küfer K-H. Breast cancer therapy planning – a novel support concept for a sequential decision making problem. Health Care Manag Sci 2015;18:389–405. https://doi.org/10.1007/s10729-014-9302-2.

[7] Hughes TP, Branford S, White DL, Reynolds J, Koelmeyer R, Seymour JF, et al. Impact of early dose intensity on cytogenetic and molecular responses in chronic- phase CML patients receiving 600 mg/day of imatinib as initial therapy. Blood 2008;112:3965–73. https://doi.org/10.1182/blood-2008-06-161737.

[8] Efficacy and Safety of Trastuzumab as a Single Agent in First-Lin...: Ingenta Connect n.d. https://www.ingentaconnect.com/content/wk/jco/2023/00000041/00000009/art00003 (accessed January 9, 2024).

[9] Luchini C, Pea A, Scarpa A. Artificial intelligence in oncology: current applications and future perspectives. Br J Cancer 2022;126:4–9. https://doi.org/10.1038/s41416-021-01633-1.

[10] O'Shaughnessy J. Extending Survival with Chemotherapy in Metastatic Breast Cancer. The Oncologist 2005;10:20–9. https://doi.org/10.1634/theoncologist.10-90003-20.

[11] Huang S, Hu P, Lakowski TM. Predicting breast cancer drug response using a multiple-layer cell line drug response network model. BMC Cancer 2021;21:648. https://doi.org/10.1186/s12885-021-08359-6.

[12] Olow AK, Veer L van 't, Wolf DM. Toward developing a metastatic breast cancer treatment strategy that incorporates history of response to previous treatments. BMC Cancer 2021;21:212. https://doi.org/10.1186/s12885-021-07912-7.

[13] Refining breast cancer biomarker discovery and drug targeting through an advanced data-driven approach | BMC Bioinformatics n.d. https://link.springer.com/article/10.1186/s12859-024-05657-1 (accessed October 18, 2024).

[14] Lopez-Perez L, Georga E, Conti C, Vicente V, García R, Pecchia L, et al. Statistical and machine learning methods for cancer research and clinical practice: A systematic review. Biomedical Signal Processing and Control 2024;92:106067. https://doi.org/10.1016/j.bspc.2024.106067.

[15] Zang Y, Lee JJ. Adaptive clinical trial designs in oncology. Chinese Clinical Oncology 2014;3:49–49. https://doi.org/10.3978/j.issn.2304-3865.2014.06.04.

[16] Abdollahian M, Das TK. A MDP Model for Breast and Ovarian Cancer Intervention Strategies for BRCA1/2 Mutation Carriers. IEEE Journal of Biomedical and Health Informatics 2015;19:720–7. https://doi.org/10.1109/JBHI.2014.2319246.

[17] Liu S, Brandeau ML, Goldhaber-Fiebert JD. Optimizing patient treatment decisions in an era

of rapid technological advances: the case of hepatitis C treatment. Health Care Manag Sci 2017;20:16–32. https://doi.org/10.1007/s10729-015-9330-6.

[18] Shukla N, Hagenbuchner M, Win KT, Yang J. Breast cancer data analysis for survivability studies and prediction. Computer Methods and Programs in Biomedicine 2018;155:199–208. https://doi.org/10.1016/j.cmpb.2017.12.011.

[19] Oh S-H, Lee SJ, Noh J, Mo J. Optimal treatment recommendations for diabetes patients using the Markov decision process along with the South Korean electronic health records. Sci Rep 2021;11:6920. https://doi.org/10.1038/s41598-021-86419-4.

[20] Lin FPY, Pokorny A, Teng C, Dear R, Epstein RJ. Computational prediction of multidisciplinary team decision-making for adjuvant breast cancer drug therapies: a machine learning approach. BMC Cancer 2016;16:929. https://doi.org/10.1186/s12885-016-2972-z.

[21] Choi R, Joel M, Hui M, Aneja S. Deep learning algorithm to predict pathologic complete response to neoadjuvant chemotherapy for breast cancer prior to treatment. JCO 2022;40:600–600. https://doi.org/10.1200/JCO.2022.40.16_suppl.600.

[22] Kim W, Kim KS, Lee JE, Noh D-Y, Kim S-W, Jung YS, et al. Development of Novel Breast Cancer Recurrence Prediction Model Using Support Vector Machine. J Breast Cancer 2012;15:230–8. https://doi.org/10.4048/jbc.2012.15.2.230.

[23] [Wang Y, Wang M, Yu K, Xu S, Qiu P, Lyu Z, et al. A machine learning model to predict efficacy of neoadjuvant therapy in breast cancer based on dynamic changes in systemic immunity. Cancer Biol Med 2023;20:218–28. https://doi.org/10.20892/j.issn.2095-3941.2022.0513.

[24] Liu S, Jiang C, Wu D, Zhang S, Qiao K, Yang X, et al. Development of predictive models for pathological response status in breast cancer after neoadjuvant therapy based on peripheral blood inflammatory indexes. BMC Women's Health 2024;24:560. https://doi.org/10.1186/s12905-024-03400-9.

[25] Frontiers | Machine learning models based on immunological genes to predict the response to neoadjuvant therapy in breast cancer patients n.d. https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2022.948601/full (accessed October 21, 2024).

[26] A generalized non-linear model predicting efficacy of neoadjuvant therapy in HER2+ breast cancer: iScience n.d. https://www.cell.com/iscience/fulltext/S2589-0042(23)00407-8 (accessed October 21, 2024).

[27] Sammut S-J, Crispin-Ortuzar M, Chin S-F, Provenzano E, Bardwell HA, Ma W, et al. Multi-omic machine learning predictor of breast cancer therapy response. Nature 2022;601:623–9. https://doi.org/10.1038/s41586-021-04278-5.

[28] Cervantes J, Stephens CR. Limitations of Existing Mutation Rate Heuristics and How a Rank GA Overcomes Them. IEEE Transactions on Evolutionary Computation 2009;13:369–97. https://doi.org/10.1109/TEVC.2008.927707.

[29] Experimental Exploration of Evolutionary Algorithms and their Applications in Complex Problems: Genetic Algorithm and Particle Swarm Optimization Algorithm. JBEMI 2023;10. https://doi.org/10.14738/bjhmr.102.14427.