

## Article

# Explainable Artificial Intelligence Reveals Novel Insight into Tumor Microenvironment Conditions Linked with Better Prognosis in Patients with Breast Cancer

Debaditya Chakraborty <sup>1,\*</sup>, Cristina Ivan <sup>2,3</sup> , Paola Amero <sup>2</sup> , Maliha Khan <sup>4</sup>, Cristian Rodriguez-Aguayo <sup>2,3</sup> , Hakan Başağaoğlu <sup>5</sup> and Gabriel Lopez-Berestein <sup>2,3</sup>

- <sup>1</sup> Department of Construction Science, The University of Texas at San Antonio, San Antonio, TX 78249, USA  
<sup>2</sup> Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA; Civan@mdanderson.org (C.I.); pamero@mdanderson.org (P.A.); CRodriguez2@mdanderson.org (C.R.-A.); glopez@mdanderson.org (G.L.-B.)  
<sup>3</sup> Center for RNA Interference and Non-Coding RNA, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA  
<sup>4</sup> Department of Lymphoma and Myeloma, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA; MKhan4@mdanderson.org  
<sup>5</sup> Evolution Online LLC, San Antonio, TX 78260, USA; hakanbasagaoglu@hotmail.com  
\* Correspondence: debaditya.chakraborty@utsa.edu



**Citation:** Chakraborty, D.; Ivan, C.; Amero, P.; Khan, M.; Rodriguez-Aguayo, C.; Başağaoğlu, H.; Lopez-Berestein, G. Explainable Artificial Intelligence Reveals Novel Insight into Tumor Microenvironment Conditions Linked with Better Prognosis in Patients with Breast Cancer. *Cancers* **2021**, *13*, 3450. <https://doi.org/10.3390/cancers13143450>

Academic Editor: Ognjen Arandjelović

Received: 4 May 2021

Accepted: 6 July 2021

Published: 9 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/>).

**Simple Summary:** Over the past decade, there has been a significant increase in the number of omics datasets that provide unprecedented opportunities to systematically characterize the underlying biological mechanisms involved in cancer evolution and to understand how the tumor microenvironment contributes to this evolution. Novel techniques in artificial intelligence (AI) can help determine areas of therapeutic need, enhance clinical trial interpretation, identify novel targets, and generate accurate predictions that are impossible with traditional statistical techniques. However, a major criticism of incorporating the highly accurate and nonlinear AI models into medical fields is the notion that AI is essentially a “black box”. We resolved this overarching problem with explainable artificial intelligence (XAI) to determine prognoses in patients with breast cancer and reveal valuable information about conditions in the tumor microenvironment that are associated with enhanced prognosis and patient survival. The benefits of using XAI in the development of new targeted therapies would be significant.

**Abstract:** We investigated the data-driven relationship between immune cell composition in the tumor microenvironment (TME) and the  $\geq 5$ -year survival rates of breast cancer patients using explainable artificial intelligence (XAI) models. We acquired TCGA breast invasive carcinoma data from the cBioPortal and retrieved immune cell composition estimates from bulk RNA sequencing data from TIMER2.0 based on EPIC, CIBERSORT, TIMER, and xCell computational methods. Novel insights derived from our XAI model showed that B cells, CD8<sup>+</sup> T cells, M0 macrophages, and NK T cells are the most critical TME features for enhanced prognosis of breast cancer patients. Our XAI model also revealed the inflection points of these critical TME features, above or below which  $\geq 5$ -year survival rates improve. Subsequently, we ascertained the conditional probabilities of  $\geq 5$ -year survival under specific conditions inferred from the inflection points. In particular, the XAI models revealed that the B cell fraction (relative to all cells in a sample) exceeding 0.025, M0 macrophage fraction (relative to the total immune cell content) below 0.05, and NK T cell and CD8<sup>+</sup> T cell fractions (based on cancer type-specific arbitrary units) above 0.075 and 0.25, respectively, in the TME could enhance the  $\geq 5$ -year survival in breast cancer patients. The findings could lead to accurate clinical predictions and enhanced immunotherapies, and to the design of innovative strategies to reprogram the breast TME.

**Keywords:** explainable artificial intelligence (XAI); machine learning; breast cancer; tumor microenvironment; survival analysis

## 1. Introduction

Breast cancer is the most common cancer and the leading cause of cancer death in women worldwide. The prognosis is dependent on the type of breast cancer and on the stage of disease at detection [1,2]. Breast cancer can be divided into several subtypes, based primarily on the expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). Triple-negative breast cancer (TNBC) is a heterogeneous category of breast cancer, characterized by negative ER, PR, and HER2. TNBC is highly metastatic and aggressive, with poor prognosis, poor patient survival, and limited therapeutic options [3].

Cancer development, progression, and treatment resistance are known to be influenced by genetic and epigenetic alterations as well as by crosstalk between tumor cells and the tumor microenvironment (TME) [4]. The TME involves a complex network of soluble factors, tumor cells, and stromal cells that plays a crucial role in the initiation, development, and progression of breast cancer. The TME in breast cancer consists of multilineage immune cells (e.g., T and B lymphocytes, myeloid cells, and dendritic cells), cancer-associated fibroblasts, and tumor endothelial cells [5]. These cell subtypes live in an ocean of hormones, growth factors, and cytokines in the breast TME [6]. Adding to this complexity is a myriad of pathways that dictate the fate of tumor, metastases, and patients' lives. The balance between tumor-infiltrating immune effector cells in the TME (such as CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells (or cytotoxic T lymphocytes, CTLs)) regulates the immune response to cytotoxic effects on tumor cells. In a mouse model of metastatic breast cancer, natural killer (NK) T cell activation was shown to enhance antitumor immunity by increasing cytotoxic responses and interferon- $\gamma$  production from NK and CD8<sup>+</sup> T cells [7]. In contrast, tumor-infiltrating myeloid cells, such as tumor-associated macrophages, promote the expansion and dissemination of cancer cells depending on their functional state [8]. Hypoxia in the TME stimulates macrophages to further produce proangiogenic factors such VEGF and suppress the T cell immune responses, hence enhancing the evasion of tumor cells and ultimately metastasis [8]. TME crosstalk potentially promotes cancer progression and TME plasticity. Adaptations to TME factors may be responsible for metastasis and immune evasion.

Multiple targeted therapies have been used for TNBC but to no avail. Aberrant signaling by VEGFR2 and cMET, as well as modifications of the immune cell population in response to an immune-suppressive phenotype, resulted in a failure of immunotherapy for TNBC [9]. In TNBC, multiple genomic instabilities and mutations have been associated with immune responses [10]. A comparative study between TNBC and non-TNBC (NTNBC) showed that TNBC is characterized by higher expression levels of functional gene sets associated with 15 types of immune cells [11,12]. Unfortunately, innovative strategies to reprogram the TME of breast cancer patients is challenging, in part because of conflicting findings in the literature. For example, tumor-infiltrating B lymphocytes (B cells) have been reported to be associated with positive, negative, or no significance in breast cancer prediction and prognosis [13]. Other studies have reported that targeting the regulatory B cell activity may be used to enhance immunotherapeutic outcomes [14].

Augmentation of CTL-induced antitumor immune reactions has been considered to be an attractive therapeutic modality for lethal solid tumors due to the tumor-killing ability of CD8<sup>+</sup> CTL [15]. In the adaptive immune system, T-helper cells (CD4<sup>+</sup> cells) play a critical role in releasing cytokines and primarily help CD8<sup>+</sup> CTL and antibody responses to mediate antitumor immunity. However, the interplay between polyfunctional CD4<sup>+</sup> T cells and other immune cell lineages within the context of tumor immunity is not well understood [16]. In addition to immune cells, tumor-associated factors in the TME have also been targeted in cancer therapies. The recognition that tumor-associated endothelial cells and cancer-associated fibroblasts are important mediators of immune suppression has led to the development of cell-specific targeting drugs in an attempt to enhance the immune response [5]. Tumor-associated macrophages are able to suppress the functions of CD8<sup>+</sup> T and NK cells and promote tumor cell growth in the TME [17].

Over the past decade, there has been a major increase in the number of large and complex omics datasets [18,19], especially through large consortium projects such as TCGA, which has sampled multiomics measurements from more than 30,000 patients and dozens of cancer types [20]. These rich omics data provide unprecedented opportunities to systematically characterize the underlying biological mechanisms involved in the evolution of cancer and to understand how the TME (hosting stromal cells, immune cells, and other types of cells) contributes to this evolution [19,21].

Novel techniques in AI can bring together diverse data types to expand novel insights gained from the multiomics datasets. It is well acknowledged that enrichment of high-quality data coupled with machine learning, a subset of AI, can help investigate the areas of changing patients' unhealthy behaviors [22], risk prediction or recurrence prediction of chronic diseases after a surgery [23] and curative treatment [24], progression and survivability of patients with chronic diseases [25], therapeutic need, enhanced clinical trial interpretation, and novel targets [26]. However, a major criticism of incorporating AI, particularly deep learning, into medical fields is the idea that AI is essentially a mechanistically uninterpretable opaque "black box" [27–29], and hence may not meet the required high level of accountability, transparency, and reliability in medical decisions [30]. The assumed lack of interpretability of AI models has been a debated topic within the field, with models cited that have achieved high accuracy due to factors that are not useful in prospective predictions [27,31,32]. The crux of the problem is that linear models, although interpretable, produce less accurate models when the datasets are complex and inherently nonlinear. In such cases, tree-based ensemble models, which are interpretable models [24], can be used in lieu of deep learning models that allow scientists and clinicians to understand the underlying reasoning behind the decisions and predictions. Recently, Gu et al. [33] used a tree-based ensemble model, called an extreme gradient boosting (XGBoost) model [34], to predict the risk of breast cancer relapse from clinical data (e.g., age, tumor size, treatment) and then use case-based reasoning—which solves new problems by constructing a historical case base and using the results of similar historical cases—to explain the reason for the prediction. In addition, the authors used a game theory-based Shapley additive explanation model called SHAP [35,36] for global explainability of the results to identify the order of importance of the clinical features considered.

In this article, we develop explainable artificial intelligence (XAI) models to establish and investigate the data-driven relationship between TME features, which comprise a vast variety of immune cells, and the  $\geq 5$ -year survival rates of breast cancer patients. The XAI models also determine relative influence of immune cells (e.g., T cells, B cells) and tumor-associated cells (e.g., macrophages) in the TME on the  $\geq 5$ -year survival rates of patients. In addition, using XAI models and conditional probabilities, we identified and analyzed the inflection points of the critical microenvironment features, above or below which the  $\geq 5$ -year survival rates could potentially improve. The resulting new perspective on favorable or deleterious microenvironmental conditions could lead to improved prognoses through well-informed clinical management and therapeutics, including the design of innovative strategies to reprogram the TME of breast cancer patients.

## 2. Materials and Methods

We downloaded patient clinical information for TCGA breast-invasive carcinoma cohort (BRCA) from two projects on the cBioPortal (<http://www.cbioportal.org/>) (accessed on 1 February 2021) [37,38]—the PanCanAtlas [20,21]—and the Firehose Legacy (<https://gdac.broadinstitute.org/>) (accessed on 1 February 2021) that provided clinical information for 1101 patients. We found 1015 patients with primary tumor samples common to both projects. For these 1015 primary tumors, we searched TIMER2.0 (<http://timer.cistrome.org/>) (accessed on 1 February 2021) for immune infiltration estimations produced with EPIC [39], CIBERSORT [40], TIMER [41], and xCell [42] computational methods [43]. We ended up with a cohort of 1014 breast cancer patients with clinical information and estimates of immune cell content in tumor tissues.

Subsequently, we developed data-driven XAI models using XGBoost and SHAP to enhance the explainability of the breast cancer survivability models based on TME conditions (including both immune cells and tumor-associated cells), understand the underlying reasoning, and expand our knowledge without compromising the predictive accuracy. Moreover, in addition to the global SHAP analysis to determine the order of importance of the TME cells on the patients' survivability rates, we performed local SHAP analysis to identify the inflection points for the TME cells, above or below which the survivability rates may increase. We demonstrated that the local SHAP analysis expanded the potential use of the interpretable AI model to investigate potential immunotherapies to increase patients' survivability rates with enhanced explainability and transparent reasonings.

XGBoost is a variant of a tree-based boosting algorithm. Conceptually, XGBoost learns the functional relationship  $f$  between the features  $x$  and target  $y$  through an iterative process in which the individual trees are sequentially trained on the residuals from the previous tree. Mathematically, the predictions from the trees can be expressed as

$$\hat{y} = \phi(x) = \frac{1}{n} \sum_{k=1}^n f_k(x) \quad (1)$$

where  $\hat{y}$  is the predicted outcome (overall survival and 5-year survival) in breast cancer patients,  $1 \leq k \leq n$ , and  $f_1, f_2, \dots, f_n$  are the functions learned by  $n$  number of trees.

The following regularized objective  $\mathcal{L}(\phi)$  is minimized to learn the set of functions  $f_k$  used in the model

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i - y_i) + \sum_k \Omega(f_k) \quad (2)$$

where  $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2$ .

In Equation (2),  $l$  is the differentiable convex loss function that measures the difference between  $\hat{y}_i$  and  $y_i$ .  $\Omega$  is an extra regularization term that penalizes the growth of more trees in the model to prevent complexity, and thus reduce overfitting.  $\gamma$  is the complexity of each leaf,  $T$  is the number of leaves in a tree,  $\lambda$  is a penalty parameter, and  $\|w\|$  is the vector of scores on the leaves. Note that if the regularization parameter  $\Omega$  is set to zero, the objective falls back to the traditional gradient tree boosting.

SHAP, in contrast, was used to explain the AI models by investigating the relationship and contribution of each feature to the predicted AI-based outcome ( $\hat{y}$ ). SHAP computes the Shapley values that signify the average marginal contribution of each feature value across all possible combinations of features. The features with large absolute Shapley values are deemed impactful. To evaluate the overall feature influence on the predicted outcome, SHAP averages the absolute Shapley values for every feature across the data, sorts them in decreasing order, and plots them. In our work, negative Shapley values associated with the feature instances indicate better chances of  $\geq 5$ -year survival.

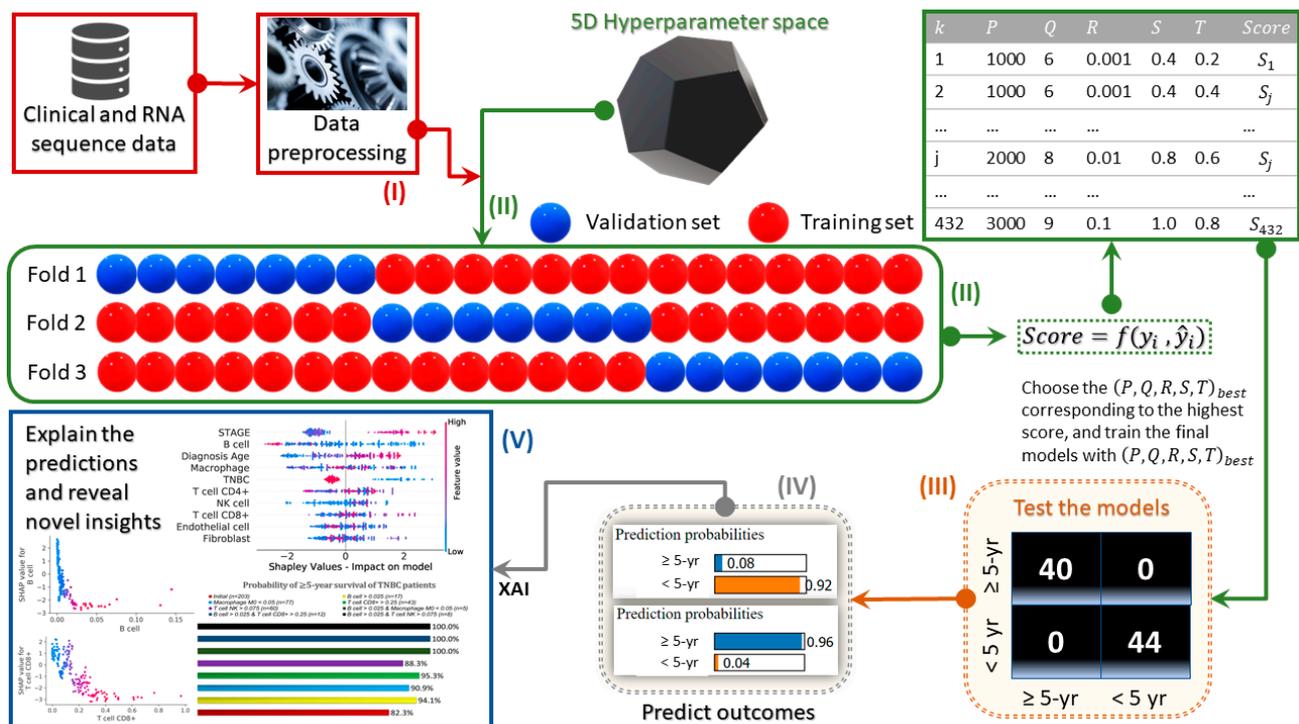
### 3. Results

We developed models with the XAI pipeline, shown schematically in Figure 1, to predict the probability of  $\geq 5$ -year survival of breast cancer patients based on immune cell composition from bulk RNA sequencing (RNA-seq) data estimated using EPIC, CIBERSORT, TIMER, and xCell methods. The XAI models were developed through a multistep process shown in Figure 1, which includes: (I) data preprocessing steps, (II) hyperparameter optimization via a three-fold cross-validation to find the best subset of hyperparameters, (III) testing of the predictive accuracy of the final AI models after being trained with the best subset of hyperparameters, (IV) prediction of the probability of the clinical outcomes, and (V) explanation of the predicted outcomes with a game theory-based XAI model to reveal the contributing factors and respective values of the TME constituents associated with

better outcomes ( $\geq 5$  years of survival) for breast cancer patients. Finally, we quantified the conditional probability of  $\geq 5$ -year survival ( $S_5$ ) for a given certain TME condition ( $C$ ) using

$$P(S_5|C) = 100 \times \frac{P(S_5 \cap C)}{P(C)} \tag{3}$$

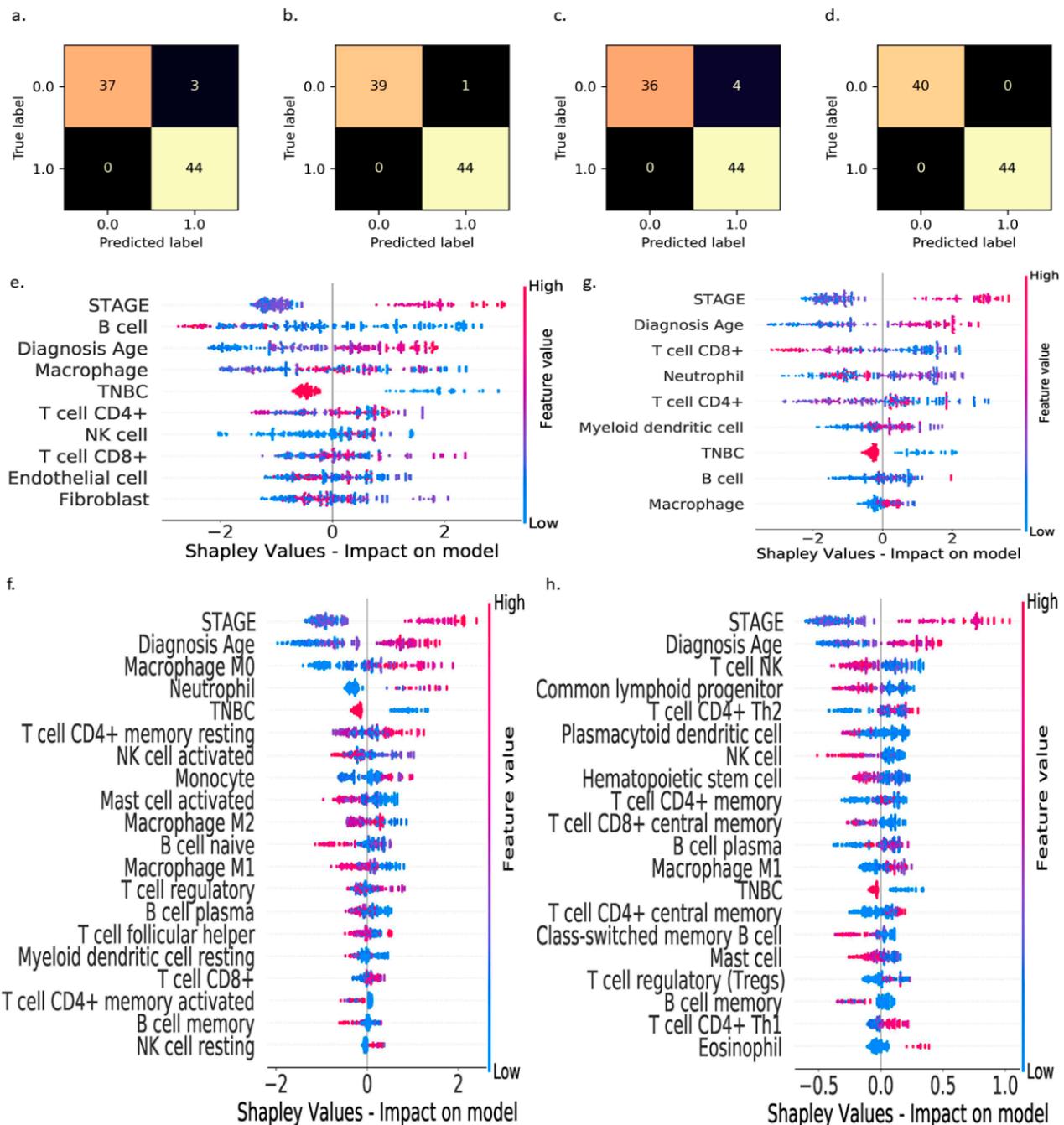
where  $P(S_5 \cap C)$  is the probability that both events  $S_5$  and  $C$  occur simultaneously, and  $P(C)$  is the probability of condition ( $C$ ) to occur.



**Figure 1.** Schematic representation of the modeling pipeline: (I) data preprocessing steps, which include encoding categorical features as integer arrays and data oversampling to convert from an imbalanced to a balanced dataset (i.e., the numbers of surviving and deceased patients were made nearly equal) to avoid bias (i.e., preventing the model from ignoring minority classes); (II) hyperparameter optimization via a 3-fold cross-validation to find the best subset of hyperparameters ( $k$ : hyperparameter subset index number,  $P$ : number of estimators,  $Q$ : maximum depth of each estimator,  $R$ : learning rate,  $S$ : subsample ratio,  $T$ : column sample ratio for each estimator) that improves the models’ ROC–AUC score  $\{f(y_i, \hat{y}_i)\}$  signifying the area under the receiver operating characteristic curve during the  $(432 \times 3)$  iterations over the hyperparameter space (i.e., to enhance the predictive accuracy of the XAI models); (III) testing of the predictive accuracy of the final AI models after being trained with the best subset of hyperparameters; (IV) prediction of the probability of the clinical outcomes with AI models; (V) explanation of the predicted outcomes with a game theory-based XAI model to enhance the interpretability and explainability of the model predictions, identification of critical inflection (turning) points, above or below which the  $\geq 5$ -year survival rates increase, and assessing the conditional probability of  $\geq 5$ -year survival rates from the range of TME factors determined by the inflection points.

With any data-driven XAI model, it is imperative to ensure that the model produces accurate predictions on testing samples that are not used during model training, and the prediction accuracies obtained during model training and testing are comparable to avoid overfitting or underfitting the data. We analyzed and reported the confusion matrices (Figure 2a–d) to gain a better understanding of the models’ performance in predicting the  $\geq 5$ -year survival on the testing data that constitute 25% of the original dataset randomly sampled from the entire preprocessed dataset. These confusion matrices show that our AI models developed with the proposed pipeline produce reliable predictions on the testing data (Figure 2a–d). The accuracy, precision, recall, and F1 score of the proposed AI models

on the testing data are tabularized in Table 1. These results indicate that the custom AI models are capable of accurately predicting (over 90%) the outcome for breast cancer patients based on the data associated with TME conditions in tumors.



**Figure 2.** Predictive accuracies of our custom XAI models in terms of their ability to predict the likelihood of  $\geq 5$ -year survival of breast cancer patients based on estimated immune cell composition from bulk RNA-seq data using EPIC (a), CIBERSORT (b), TIMER (c), and xCell (d) cell type quantification methods. The XAI indicates that the B cells (e), M0 macrophages (f), CD8<sup>+</sup> T cells (g), and NK T cells (h) (estimated using EPIC, CIBERSORT, TIMER, and xCell methods, respectively) are the most important immune cells in the TME features in predicting survivability of breast cancer patients. The features on the y-axis are used in the respective models; their relative positions were determined by their relative importance in making correct predictions. The blue dots represent lower feature values, and the red dots represent higher feature values. In these analyses, Shapley values  $< 0$  represent “likely to survive longer than 5 years after diagnosis” while Shapley values  $> 0$  represent “likely to die within 5 years” after diagnosis.

**Table 1.** Statistical validation of the predictive accuracies of the AI models.

TME Immune Cell Estimation Method	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
EPIC	96.4	93.6	100.0	96.7
CIBERSORT	98.8	97.8	100.0	98.9
TIMER	95.2	91.7	100.0	95.7
xCell	100.0	100.0	100.0	100.0

After evaluating the predictive ability of the models, we used XAI to interpret our AI models' predictions, investigate novel relationships between the TME features and the  $\geq 5$ -year survival status, and identify the critical inflection points, above or below which the  $\geq 5$ -year survival rates improve. We explain the custom XAI models from both global (entire dataset) and local (individual data points) perspectives. The global explanations revealed that the B cells, CD8<sup>+</sup> T cells, M0 macrophages, and NK T cells are the most influential TME factors resulting from the RNA-seq data produced by the EPIC, CIBERSORT, TIMER, and xCell methods in determining the  $\geq 5$ -year survival (Figure 2e–h). The XAI models suggest that higher CD8<sup>+</sup> T cell, NK T cell, and B cell counts, along with low M0 macrophage count, lead to higher survival rates for breast cancer patients. It is worth noting that because different TME features are estimated by the EPIC, CIBERSORT, TIMER, and xCell methods, the relative importance of the TME features varied in the global SHAP analysis for each XAI model in Figure 2e–h.

The EPIC-XAI analysis (Figure 2e) revealed that B cells (antibody-producing machines) play a more critical role than other TME factors in improving the  $\geq 5$ -year survival rates for breast cancer patients. Although the role of B cells was reported to be controversial in cancer immunotherapies [13], Figure 2e reveals that the presence and activation of larger numbers of B cells in the TME could enhance patients' survivability rates and the efficacy of cancer immunotherapies. These findings are particularly important if the cancer is not at an advanced stage, at which a patient would be at higher risk of death according to Shapley analysis in Figure 2e–h. On the other hand, the CIBERSORT-XAI (Figure 2f) analysis revealed that M0 macrophages are the most influential TME factor that lowers the  $\geq 5$ -year survival rates when present in larger numbers. CD8<sup>+</sup> T cells and NK T cells were identified as the most critical tumor-suppressing immune cells by the TIMER-XAI and xCell-XAI analyses, respectively, for the enhanced  $\geq 5$ -year survivability of the patients. We also found that the CD4<sup>+</sup> T cells were identified by the XAI as the third most influential immune cell type on the breast cancer prognosis. These new findings and insights indicate an urgent need to rethink the current cancer immunotherapies that are largely focused on harnessing the antitumor CD8<sup>+</sup> cytotoxic T cell response [12,16].

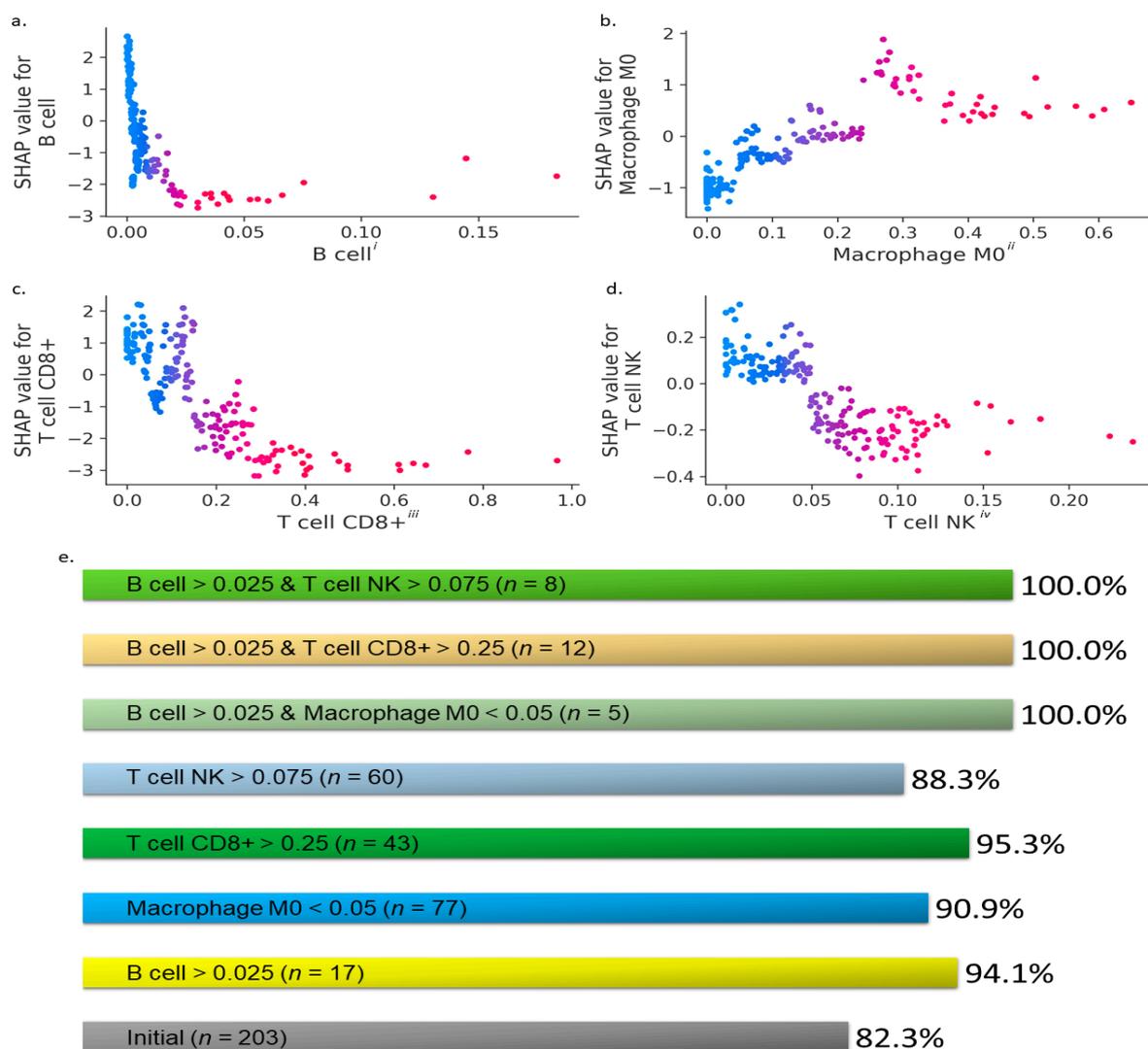
Next, we analyzed the influence of the topmost performing TME features (i.e., B cells, CD8<sup>+</sup> T cells, M0 macrophages, and NK T cells) from the EPIC, CIBERSORT, TIMER, and xCell XAI analyses on the survivability rate in breast cancer patients. In Figure 3, we illustrate the effect of the variations in B cell, CD8<sup>+</sup> T cell, M0 macrophage, and NK T cell estimates on the models' predictions to identify the critical inflection points, above or below which the  $\geq 5$ -year survival rates improve. We found that the B cell fraction  $> 0.025$  (Figure 3a), M0 macrophage fraction  $< 0.05$  (Figure 3b), CD8<sup>+</sup> T cell fraction  $> 0.25$  (Figure 3c), and NK T cell fraction  $> 0.075$  (Figure 3d) are ideal conditions, as characterized by lower SHAP values on the y-axis, for enhanced  $\geq 5$ -year survivability chances for breast cancer patients.

We designed seven different TME conditions based on these inflection points:

- (i) C<sub>1</sub>: B cells  $> 0.025$ ,
- (ii) C<sub>2</sub>: CD8<sup>+</sup> T cells  $> 0.25$ ,
- (iii) C<sub>3</sub>: M0 macrophages  $< 0.05$ ,
- (iv) C<sub>4</sub>: NK T cells  $> 0.075$ ,
- (v) C<sub>5</sub>: B cells  $> 0.025$  and M0 macrophages  $< 0.05$ ,

- (vi)  $C_6$ : B cells > 0.025 and CD8<sup>+</sup> T cells > 0.25, and  
 (vii)  $C_7$ : B cells > 0.025 and NK T cells > 0.075,

which were coupled with Equation (3) to quantify the conditional probability of  $\geq 5$ -year survival ( $S_5$ ) of breast cancer patients (Figure 3e). We found that the initial probability of  $\geq 5$ -year survival ( $S_5$ ) based on the original dataset was 82.3%. In contrast, the probability of  $\geq 5$ -year survival ( $S_5$ ) given conditions  $C_1$  to  $C_7$ , i.e.,  $P(S_5|C_1)$  to  $P(S_5|C_7)$ , increased by up to ~18% (Figure 3e). In other words, it is possible to increase the  $\geq 5$ -year survival chances by ~18% by boosting the B cell and CD8<sup>+</sup> T cell fractions or B cell and NK T cell fractions in the TME above their respective inflection points. The XAI-based revelation of the critical inflection points and the statistical evaluation of certain TME conditions have high potential in deriving novel insights that could have clinical implications for accurate predictions and targeted clinical treatment of patients with breast cancer.



**Figure 3.** XAI results based on data from breast cancer patients who survived  $\geq 5$  years and are still alive, are dead after surviving  $\geq 5$  years, or are dead after surviving <5 years. The local SHAP analysis (a–d) reveals the interaction between B cells (a), M0 macrophages (b), CD8<sup>+</sup> T cells (c), and NK T cells (d) (estimated using EPIC, CIBERSORT, TIMER, and xCell methods) with the  $\geq 5$ -year survival rates, respectively. Units: i—cell fractions relative to all cells in sample, ii—immune cell fractions relative to total immune cell content, iii and iv—cancer type-specific arbitrary units comparable between samples. Lower SHAP values on the y-axis indicate higher chances of  $\geq 5$ -year survival in breast cancer patients. The conditional probabilities of  $\geq 5$ -year survival of all breast cancer patients in various TME conditions from the clinical dataset of patients are shown in (e).

Furthermore, we analyzed the TCGA data to identify the individual TNBC and non-TNBC patients who survived for less than 5 years to verify the effectiveness of the critical TME factors identified by our XAI models on the enhanced survivability rates. We tabulated the individual patients' clinical data and the associated B cell, CD8<sup>+</sup> T cell, NK T cell, and M0 macrophage fractions (Tables 2 and 3). We found that there was only one patient with TNBC (B6-A3ZX) and B cell fraction in the TME above the identified inflection point who survived less than five years after diagnosis. However, this patient was a stage IV TNBC patient, and the effect of higher B cell fraction was likely insignificant on the patient's survival, as the cancer stage was found in Figure 2e to be the more decisive factor than the B cell count on the cancer prognosis. The anomaly in our identified inflection point for B cells is ~2.8% (one in 36 TNBC and NTNBC patients who survived for less than 5 years). Similarly, we found that the anomaly in our identified inflection point for CD8<sup>+</sup> T cells is ~8.3%, i.e., three in 36 TNBC and NTNBC patients (AR-A5QQ, BH-A0C1, BH-A1EY) who survived for less than 5 years. The anomalies associated with the M0 macrophage (BH-A1EW, C8-A3M7, EW-A1P8, BH-A0C1, D8-A1Y1, E2-A14Z, LL-A73Z (stage IV patient)) and NK T cells (A2-A0T2 (stage IV patient), AC-A2QJ, B6-A409, EW-A1P8, BH-A18J, E2-A1LE, LL-A73Z (stage IV patient)) inflection points were relatively higher at ~19.4%, i.e., seven in 36 TNBC and NTNBC patients who survived for less than 5 years.

**Table 2.** Clinical and critical immune cell composition (identified via XAI) from bulk RNA-seq data of TNBC patients that survived for less than 5 years.

TCGA Patient ID	Stage	Age	Months	B Cell	M0 Macrophage	CD8 <sup>+</sup> T Cell	NK T Cell
A1-A0SK	II	54	31.8	0.003	0.325	0.000	0.000
A2-A0CM	II	40	24.8	0.008	0.129	0.025	0.013
A2-A0T2	IV	66	8.4	0.001	0.077	0.098	0.183
A2-A3XY	II	49	35.9	0.013	0.261	0.049	0.072
AC-A2QJ	III	48	14.7	0.000	0.196	0.000	0.146
AR-A5QQ	III	68	10.6	0.017	0.109	0.259	0.061
B6-A3ZX	IV	50	37.9	0.145	0.058	0.110	0.058
B6-A409	III	44	18.8	0.002	0.052	0.000	0.096
BH-A1EW	II	38	55.7	0.004	0.000	0.244	0.047
C8-A3M7	III	60	34.0	0.008	0.000	0.216	0.014
E2-A1LK	III	84	8.7	0.005	0.650	0.000	0.034
EW-A1P8	III	58	7.9	0.003	0.000	0.034	0.101

**Table 3.** Clinical and critical immune cell composition (identified via XAI) from bulk RNA-seq data of non-TNBC patients that survived for less than 5 years.

TCGA Patient ID	Stage	Age	Months	B cell	M0 Macrophage	CD8 <sup>+</sup> T Cell	NK T Cell
A2-A0SV	IV	63	27.1	0.000	0.173	0.016	0.038
A7-A13E	II	62	20.2	0.001	0.504	0.118	0.005
A8-A08J	IV	52	37.1	0.003	0.240	0.014	0.050
AC-A23H	II	90	5.7	0.002	0.419	0.150	0.003
AR-A0TY	II	54	55.9	0.007	0.270	0.030	0.043
BH-A0C1	III	61	46.4	0.002	0.000	0.284	0.034
BH-A18J	IV	56	20.1	0.001	0.375	0.084	0.081
BH-A18P	I	60	30.3	0.003	0.159	0.230	0.049
BH-A18T	II	70	7.4	0.001	0.157	0.000	0.008
BH-A1EV	III	45	12.0	0.001	0.310	0.087	0.053
BH-A1EX	II	67	49.6	0.003	0.061	0.147	0.035
BH-A1EY	II	79	17.7	0.001	0.072	0.250	0.025
BH-A1F8	III	90	25.1	0.005	0.069	0.188	0.000
BH-A1FD	I	68	33.2	0.000	0.130	0.051	0.004
C8-A12Q	III	78	12.7	0.007	0.264	0.177	0.038

Table 3. Cont.

TCGA Patient ID	Stage	Age	Months	B cell	M0 Macrophage	CD8 <sup>+</sup> T Cell	NK T Cell
D8-A1XC	III	85	12.4	0.004	0.266	0.126	0.041
D8-A1Y1	III	80	9.9	0.000	0.000	0.050	0.010
D8-A73W	III	79	12.7	0.002	0.290	0.000	0.072
E2-A14Z	I	64	18.5	0.005	0.018	0.147	0.067
E2-A1LE	III	71	28.9	0.002	0.173	0.219	0.085
E9-A1N6	II	52	22.3	0.000	0.313	0.081	0.038
E9-A1NF	II	60	35.2	0.000	0.280	0.124	0.033
LL-A73Z	IV	55	7.5	0.006	0.045	0.130	0.104
UU-A93S	IV	63	3.8	0.001	0.275	0.038	0.069

#### 4. Discussion

The application of AI models for diagnostic and prognostic assessments has been widely accepted in the context of some cancers [44,45]. The ability of AI models to discover embedded nonlinear patterns within complex multivariate datasets could potentially lead to a better understanding of the complex mechanisms that underlie carcinogenesis and cancer progression [46]. However, recent research indicates that there is a trend toward blind acceptance of black box models that lack transparency and accountability, which could have severe consequences [47]. It is imperative to apply AI models that are inherently interpretable together with XAI methods to produce accurate predictions to better understand the underlying reasoning of the AI approach and discover new interpretable knowledge from large datasets that would otherwise be impossible with traditional statistical techniques [48]. To overcome such problems, which stem from the model's lack of transparency, we used XAI models comprising tree-based ensembles—which are more interpretable than the black box-type deep learning models (e.g., artificial neural networks)—along with game theory-based explanation models to determine prognoses in patients with breast cancer and to disclose valuable information regarding the ideal TME conditions for improved prognosis and treatments.

In the past, AI analysis has focused on early diagnosis of primary cancers, survival prognosis, or risk of relapse [49,50]. Janizek et al. introduced “TreeCombo,” a gradient of the boosted tree-based approach, in combination with Shapley analysis to predict synergy of novel drug combinations [51]. In various cancers, the density of tumor-infiltrating lymphocytes (e.g., B cells, T cells) correlated positively with survival prognosis [52,53]. Using AI-based analysis (through a random forest tree-based classifier) and CIBERSORT, He et al. [54] reported that high immunity (a subset of TNBC according to the authors) was associated with larger numbers of CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, NK cells, and M1 macrophages in the TME, and hence was considered to have more favorable clinical outcomes than other subtypes of TNBC. Similarly, with use of AI-based image analysis, He et al. [55] reported that tumor-infiltrating lymphocyte cells were significantly reduced in the TME in metastatic TNBC, compared with the number of these cells in primary TNBC, and larger numbers of tumor-infiltrating lymphocytes were found to be associated with better prognosis.

Our XAI models indicated that boosting the B cell, CD8<sup>+</sup> T cell, and NK T cell fractions above their inflection points and/or reducing the M0 macrophage fraction to a level below its inflection point in the TME would be conducive for optimal collaboration of the TME features. For example, collaboration between T and B cells to carry out eradication of tumor cells, which may be related to CD4<sup>+</sup> T cells (identified as the third most influential TME features on breast cancer prognosis from the EPIC, CIBERSORT, TIMER, and xCell XAI analyses in Figure 2) causing B cells to proliferate and their progeny to differentiate into antibody-secreting cells [56]. NK T cells contribute to B cell maturation, antibody and cytokine production, and antigen presentation [57]. Subsequently, the B cells mark the tumor cells for destruction, which is carried out by cytotoxic cells such as CD8<sup>+</sup> T cells and NK T cells. Furthermore, this response is likely amplified by T cell receptors arming

the cytotoxic T cells. To the best of our knowledge, the relative importance and novel interactions between the tumor-infiltrating lymphocytes and tumor-associated cells in the TME on the overall and  $\geq 5$ -year survival prognosis of breast cancer patients have not been previously reported, although such immune signatures could have potential clinical implications, especially for TNBC treatment.

In previous studies, CD8<sup>+</sup> T cells were considered to be the crucial effector cells mediating effective antitumor immunity, resulting in better clinical outcomes, whereas intra-tumoral CD4<sup>+</sup> T cells have negative prognostic effects on breast cancer patient outcomes [43]. Hollern et al. reported that CD4<sup>+</sup> T-helper follicular cells, B cells, and the antibodies generated by those B cells play important roles in antitumor response to dual immune checkpoint inhibitors in mouse models [58,59]. Furthermore, previous use of rituximab to deplete B cells demonstrated no real clinical benefits for patients with solid tumors [60,61]. NK T cells have a substantial capacity to produce extensive amounts of cytokines upon stimulation to activate NK cells, regulatory and conventional T cells, and B cells [62]. Antitumor efficacy of T cells was shown in vivo to be dependent on the presence of non-T cells, in which the activated NK T cells rendered T cells resistant to myeloid-derived suppressor cells [63]. Considering their broad cytokine profile and potential immune enhancing and immunosuppressive roles, modulating the NK T cells activity towards immune activation has been considered an immunotherapeutic option [64]. Tumor-associated macrophages promote tumor growth by suppressing immunocompetent cells, including neovascularization and supporting cancer stem cells [65], and could also help the tumor cells escape from elimination and spread to other tissues and organs [66]. Macrophages are re-engineered to modulate their regulatory role, for example, to reduce tumor collagen deposition and promote T cell infiltration into breast tumors [67].

The novel insights derived from our XAI model shed light on the strong impacts of B cells, CD8<sup>+</sup> T cells, and NK T cells, along with M0 macrophages on the survival chances of patients with breast cancer and reveal their critical inflection points for designing innovative strategies to reprogram the TME. The main systemic therapy for metastatic TNBC is chemotherapy, despite the poor prognosis and poor patient survival associated with its use. In such cases, boosting the B cells, CD8<sup>+</sup> T cells, and NK T cells in the TME as a targeted immunotherapy could serve as a better alternative. Our findings underscore the need to rethink the current cancer immunotherapies focused on harnessing only the antitumor CD8<sup>+</sup> cytotoxic T cell response. Increased awareness through use of XAI to understand the dynamics of the TME could lead to more rational and evidence-based therapies, leading to improved outcomes in breast cancer patients.

We recognize that one of the major TCGA data analysis results is that at least 80% of the submitted samples must be composed of tumor cells. In the future, the use of other techniques like radiomic analysis to predict histological outcomes, parenchymal enhancement, and single cell proteomics may provide the data required for AI predictions of tumor behavior and outcomes [68,69].

## 5. Conclusions

Our new XAI model framework with enhanced model interpretability and explainability of the results may reduce the concerns about the transparency and accountability of the use of AI models in medical decisions. Our EPIC, CIBERSORT, TIMER, and xCell XAI analyses revealed that B cells, CD8<sup>+</sup> T cells, NK T cells, and M0 macrophages are the most critical TME features for breast cancer prognosis. Our XAI models further revealed that by boosting the B cell and CD8<sup>+</sup> T cell fractions or B cell and NK T cell fractions in the TME to levels above their inflection points—identified by XAI analysis in this study—the survival rate of breast cancer patients could increase by up to 18%. They could be alternative immunotherapies to conventional breast cancer therapeutics, although these findings require further in vitro and in vivo testing and clinical verifications.

**Author Contributions:** D.C., C.R.-A., H.B., and G.L.-B. conceived the study; H.B. and G.L.-B. supervised the progress; D.C., C.R.-A., H.B., and G.L.-B. edited the manuscript; D.C., C.R.-A., P.A., M.K., C.I., H.B., and G.L.-B. designed and performed the experiments; D.C. and C.I. analyzed the data; and D.C., C.R.-A., P.A., M.K., C.I., H.B., and G.L.-B. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by grants from the National Institutes of Health/National Cancer Institute (5U01CA213759-02, P30CA016672), and the American Cancer Society, National Science Foundation (CHE-1411859), and an endowment grant from the John P. Gaines Foundation. Dr. Cristian Rodriguez-Aguayo and Dr. Paola Amero were supported by the Brain SPORE Career Enhancement Program and NCI grant P50CA127001, as well as by the NIH through the Ovarian SPORE Career Enhancement Program and NCI grant P50CA217685.

**Institutional Review Board Statement:** There are no animal or human data that need approval.

**Informed Consent Statement:** Not applicable, data from humans were downloaded from the cBioPortal for Cancer Genomics publicly available at <https://www.cbioportal.org/> (accessed on 1 February 2021).

**Data Availability Statement:** The data and models will be made available upon request.

**Acknowledgments:** We thank Tamara Locke, Scientific Editor, Research Medical Library at the University of Texas MD Anderson Cancer Center for critical reading of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)]
2. Allaoui, R.; Bergenfelz, C.; Mohlin, S.; Hagerling, C.; Salari, K.; Werb, Z.; Anderson, R.L.; Ethier, S.P.; Jirstrom, K.; Pahlman, S.; et al. Cancer-associated fibroblast-secreted CXCL16 attracts monocytes to promote stroma activation in triple-negative breast cancers. *Nat. Commun.* **2016**, *7*, 13050. [[CrossRef](#)]
3. Liedtke, C.; Mazouni, C.; Hess, K.R.; Andre, F.; Tordai, A.; Mejia, J.A.; Symmans, W.F.; Gonzalez-Angulo, A.M.; Hennessy, B.; Green, M.; et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* **2008**, *26*, 1275–1281. [[CrossRef](#)]
4. Bareche, Y.; Buisseret, L.; Gruosso, T.; Girard, E.; Venet, D.; Dupont, F.; Desmedt, C.; Larsimont, D.; Park, M.; Rothe, F.; et al. Unraveling Triple-Negative Breast Cancer Tumor Microenvironment Heterogeneity: Towards an Optimized Treatment Approach. *J. Natl. Cancer Inst.* **2020**, *112*, 708–719. [[CrossRef](#)]
5. Nagl, L.; Horvath, L.; Pircher, A.; Wolf, D. Tumor Endothelial Cells (TECs) as Potential Immune Directors of the Tumor Microenvironment—New Findings and Future Perspectives. *Front Cell Dev. Biol.* **2020**, *8*, 766. [[CrossRef](#)] [[PubMed](#)]
6. Allinen, M.; Beroukhi, R.; Cai, L.; Brennan, C.; Lahti-Domenici, J.; Huang, H.; Porter, D.; Hu, M.; Chin, L.; Richardson, A.; et al. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* **2004**, *6*, 17–32. [[CrossRef](#)] [[PubMed](#)]
7. Gebremeskel, S.; Clattenburg, D.R.; Slauenwhite, D.; Lobert, L.; Johnston, B. Natural killer T cell activation overcomes immunosuppression to enhance clearance of postsurgical breast cancer metastasis in mice. *Oncoimmunology* **2015**, *4*, e995562. [[CrossRef](#)] [[PubMed](#)]
8. Obeid, E.; Nanda, R.; Fu, Y.X.; Olopade, O.I. The role of tumor-associated macrophages in breast cancer progression (review). *Int. J. Oncol.* **2013**, *43*, 5–12. [[CrossRef](#)]
9. Bianchini, G.; Balko, J.M.; Mayer, I.A.; Sanders, M.E.; Gianni, L. Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **2016**, *13*, 674–690. [[CrossRef](#)]
10. Chokr, N.; Chokr, S. Immune Checkpoint Inhibitors in Triple Negative Breast Cancer: What is the Evidence? *J. Neoplasms.* **2018**, *3*. [[CrossRef](#)]
11. Liu, T.; Han, C.; Wang, S.; Fang, P.; Ma, Z.; Xu, L.; Yin, R. Cancer-associated fibroblasts: An emerging target of anti-cancer immunotherapy. *J. Hematol. Oncol.* **2019**, *12*, 86. [[CrossRef](#)]
12. Li, X.; Gruosso, T.; Zuo, D.; Omeroglu, A.; Meterissian, S.; Guiot, M.C.; Salazar, A.; Park, M.; Levine, H. Infiltration of CD8<sup>+</sup> T cells into tumor cell clusters in triple-negative breast cancer. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3678–3687. [[CrossRef](#)]
13. Shen, M.; Wang, J.; Ren, X. New Insights into Tumor-Infiltrating B Lymphocytes in Breast Cancer: Clinical Impacts and Regulatory Mechanisms. *Front Immunol.* **2018**, *9*, 470. [[CrossRef](#)]
14. Schwartz, M.; Zhang, Y.; Rosenblatt, J.D. B cell regulation of the anti-tumor response and role in carcinogenesis. *J. Immunother. Cancer* **2016**, *4*, 40. [[CrossRef](#)]

15. Cassetta, L.; Kitamura, T. Targeting Tumor-Associated Macrophages as a Potential Strategy to Enhance the Response to Immune Checkpoint Inhibitors. *Front Cell. Dev. Biol.* **2018**, *6*, 38. [[CrossRef](#)]
16. Tay, R.E.; Richardson, E.K.; Toh, H.C. Revisiting the role of CD4<sup>+</sup> T cells in cancer immunotherapy—new insights into old paradigms. *Cancer Gene Ther.* **2021**, *28*, 5–17. [[CrossRef](#)]
17. DeNardo, D.G.; Ruffell, B. Macrophages as regulators of tumour immunity and immunotherapy. *Nat. Rev. Immunol.* **2019**, *19*, 369–382. [[CrossRef](#)] [[PubMed](#)]
18. Camacho, D.M.; Collins, K.M.; Powers, R.K.; Costello, J.C.; Collins, J.J. Next-Generation Machine Learning for Biological Networks. *Cell* **2018**, *173*, 1581–1592. [[CrossRef](#)] [[PubMed](#)]
19. Li, J.; Chen, H.; Wang, Y.; Chen, M.M.; Liang, H. Next-Generation Analytics for Omics Data. *Cancer Cell* **2021**, *39*, 3–6. [[CrossRef](#)] [[PubMed](#)]
20. Hutter, C.; Zenklusen, J.C. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. *Cell* **2018**, *173*, 283–285. [[CrossRef](#)] [[PubMed](#)]
21. Srivastava, S.; Ghosh, S.; Kagan, J.; Mazurchuk, R.; National Cancer Institute's, H.I. The Making of a PreCancer Atlas: Promises, Challenges, and Opportunities. *Trends Cancer* **2018**, *4*, 523–536. [[CrossRef](#)]
22. Dragoni, M.; Donadello, I.; Eccher, C. Explainable AI meets persuasiveness: Translating reasoning results into behavioral change advice. *Artif. Intell. Med.* **2020**, *105*, 101840. [[CrossRef](#)]
23. Lou, S.J.; Hou, M.F.; Chang, H.T.; Chiu, C.C.; Lee, H.H.; Yeh, S.J.; Shi, H.Y. Machine Learning Algorithms to Predict Recurrence within 10 Years after Breast Cancer Surgery: A Prospective Cohort Study. *Cancers* **2020**, *12*, 3817. [[CrossRef](#)] [[PubMed](#)]
24. Richter, A.N.; Khoshgoftaar, T.M. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. *Artif. Intell. Med.* **2018**, *90*, 1–14. [[CrossRef](#)] [[PubMed](#)]
25. Ferroni, P.; Zanzotto, F.M.; Riondino, S.; Scarpato, N.; Guadagni, F.; Roselli, M. Breast Cancer Prognosis Using a Machine Learning Approach. *Cancers* **2019**, *11*, 328. [[CrossRef](#)] [[PubMed](#)]
26. Greshock, J.; Lewi, M.; Hartog, B.; Tendler, C. Harnessing Real-World Evidence for the Development of Novel Cancer Therapies. *Trends Cancer* **2020**, *6*, 907–909. [[CrossRef](#)]
27. Gilvary, C.; Madhukar, N.; Elkhader, J.; Elemento, O. The Missing Pieces of Artificial Intelligence in Medicine. *Trends Pharmacol. Sci.* **2019**, *40*, 555–564. [[CrossRef](#)]
28. Yang, J.H.; Wright, S.N.; Hamblin, M.; McCloskey, D.; Alcantar, M.A.; Schrubbers, L.; Lopatkin, A.J.; Satish, S.; Nili, A.; Palsson, B.O.; et al. A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action. *Cell* **2019**, *177*, 1649–1661. [[CrossRef](#)] [[PubMed](#)]
29. Lamy, J.B.; Sekar, B.; Guezennec, G.; Bouaud, J.; Seroussi, B. Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. *Artif. Intell. Med.* **2019**, *94*, 42–53. [[CrossRef](#)]
30. Tjoa, E.; Guan, C. A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Trans Neural. Netw. Learn. Syst.* **2020**. [[CrossRef](#)]
31. Andreson, C. Ready for Prime Time?: AI Influencing Precision Medicine but May Not Match the Hype. *Clin. OMICs* **2018**, *5*, 44–46. [[CrossRef](#)]
32. Shaywitz, D. AI Doesn't Ask Why—But Physicians and Drug Developers Want to Know. *Forbes*, 9 November 2018.
33. Gu, D.; Su, K.; Zhao, H. A case-based ensemble learning system for explainable breast cancer recurrence prediction. *Artif. Intell. Med.* **2020**, *107*, 101858. [[CrossRef](#)] [[PubMed](#)]
34. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; ACM: New York, NY, USA, 2016; pp. 785–794.
35. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.
36. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [[CrossRef](#)] [[PubMed](#)]
37. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2012**, *2*, 401–404. [[CrossRef](#)]
38. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [[CrossRef](#)] [[PubMed](#)]
39. Racle, J.; de Jonge, K.; Baumgaertner, P.; Speiser, D.E.; Gfeller, D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *eLife* **2017**, *6*. [[CrossRef](#)]
40. Newman, A.M.; Liu, C.L.; Green, M.R.; Gentles, A.J.; Feng, W.; Xu, Y.; Hoang, C.D.; Diehn, M.; Alizadeh, A.A. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **2015**, *12*, 453–457. [[CrossRef](#)] [[PubMed](#)]
41. Li, B.; Severson, E.; Pignon, J.C.; Zhao, H.; Li, T.; Novak, J.; Jiang, P.; Shen, H.; Aster, J.C.; Rodig, S.; et al. Comprehensive analyses of tumor immunity: Implications for cancer immunotherapy. *Genome Biol.* **2016**, *17*, 174. [[CrossRef](#)]
42. Aran, D.; Hu, Z.; Butte, A.J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **2017**, *18*, 220. [[CrossRef](#)] [[PubMed](#)]

43. Sturm, G.; Finotello, F.; Petitprez, F.; Zhang, J.D.; Baumbach, J.; Fridman, W.H.; List, M.; Aneichyk, T. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **2019**, *35*, i436–i445. [[CrossRef](#)]
44. Rakha, E.A.; Reis-Filho, J.S.; Ellis, I.O. Combinatorial biomarker expression in breast cancer. *Breast Cancer Res. Treat.* **2010**, *120*, 293–308. [[CrossRef](#)]
45. Kourou, K.; Exarchos, T.P.; Exarchos, K.P.; Karamouzis, M.V.; Fotiadis, D.I. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 8–17. [[CrossRef](#)]
46. Kawakami, E.; Tabata, J.; Yanaihara, N.; Ishikawa, T.; Koseki, K.; Iida, Y.; Saito, M.; Komazaki, H.; Shapiro, J.S.; Goto, C.; et al. Application of Artificial Intelligence for Preoperative Diagnostic and Prognostic Prediction in Epithelial Ovarian Cancer Based on Blood Biomarkers. *Clin. Cancer Res.* **2019**, *25*, 3006–3015. [[CrossRef](#)]
47. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [[CrossRef](#)]
48. Kernbach, J.M.; Staartjes, V.E. Predicted Prognosis of Pancreatic Cancer Patients by Machine Learning-Letter. *Clin. Cancer Res.* **2020**, *26*, 3891. [[CrossRef](#)]
49. Ichimasa, K.; Kudo, S.E.; Mori, Y.; Misawa, M.; Matsudaira, S.; Kouyama, Y.; Baba, T.; Hidaka, E.; Wakamura, K.; Hayashi, T.; et al. Artificial intelligence may help in predicting the need for additional surgery after endoscopic resection of T1 colorectal cancer. *Endoscopy* **2018**, *50*, 230–240. [[CrossRef](#)] [[PubMed](#)]
50. Ito, N.; Kawahira, H.; Nakashima, H.; Uesato, M.; Miyauchi, H.; Matsubara, H. Endoscopic Diagnostic Support System for cT1b Colorectal Cancer Using Deep Learning. *Oncology* **2019**, *96*, 44–50. [[CrossRef](#)] [[PubMed](#)]
51. Janizek, J.D.; Celik, S.; Lee, S.-I. Explainable machine learning prediction of synergistic drug combinations for precision cancer medicine. *bioRxiv* **2018**. [[CrossRef](#)]
52. Bindea, G.; Mlecnik, B.; Tosolini, M.; Kirilovsky, A.; Waldner, M.; Obenaus, A.C.; Angell, H.; Fredriksen, T.; Lafontaine, L.; Berger, A.; et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **2013**, *39*, 782–795. [[CrossRef](#)]
53. Garnelo, M.; Tan, A.; Her, Z.; Yeong, J.; Lim, C.J.; Chen, J.; Lim, K.H.; Weber, A.; Chow, P.; Chung, A.; et al. Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma. *Gut* **2017**, *66*, 342–351. [[CrossRef](#)]
54. He, Y.; Jiang, Z.; Chen, C.; Wang, X. Classification of triple-negative breast cancers based on Immunogenomic profiling. *J. Exp. Clin. Cancer Res.* **2018**, *37*, 327. [[CrossRef](#)]
55. He, T.F.; Yost, S.E.; Frankel, P.H.; Dagens, A.; Cao, Y.; Wang, R.; Rosario, A.; Tu, T.Y.; Solomon, S.; Schmolze, D.; et al. Multi-panel immunofluorescence analysis of tumor infiltrating lymphocytes in triple negative breast cancer: Evolution of tumor immune profiles and patient prognosis. *PLoS ONE* **2020**, *15*, e0229955. [[CrossRef](#)] [[PubMed](#)]
56. Janeway, C.A., Jr.; Travers, P.; Walport, M.; Shlomchik, M.J. Part V. The Immune System in Health and Disease. In *Immunobiology: The Immune System in Health and Disease*, 5th ed.; Garland Science: New York, NY, USA, 2001.
57. Doherty, D.G.; Melo, A.M.; Moreno-Olivera, A.; Solomos, A.C. Activation and Regulation of B Cell Responses by Invariant Natural Killer T Cells. *Front Immunol.* **2018**, *9*, 1360. [[CrossRef](#)]
58. Huang, Y.; Ma, C.; Zhang, Q.; Ye, J.; Wang, F.; Zhang, Y.; Hunborg, P.; Varvares, M.A.; Hoft, D.F.; Hsueh, E.C.; et al. CD4<sup>+</sup> and CD8<sup>+</sup> T cells have opposing roles in breast cancer progression and outcome. *Oncotarget* **2015**, *6*, 17462–17478. [[CrossRef](#)]
59. Hollern, D.P.; Xu, N.; Thennavan, A.; Glodowski, C.; Garcia-Recio, S.; Mott, K.R.; He, X.; Garay, J.P.; Carey-Ewend, K.; Marron, D.; et al. B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast Cancer. *Cell* **2019**, *179*, 1191–1206.e21. [[CrossRef](#)] [[PubMed](#)]
60. Gunderson, A.J.; Coussens, L.M. B cells and their mediators as targets for therapy in solid tumors. *Exp. Cell Res.* **2013**, *319*, 1644–1649. [[CrossRef](#)] [[PubMed](#)]
61. Lu, Y.; Zhao, Q.; Liao, J.Y.; Song, E.; Xia, Q.; Pan, J.; Li, Y.; Li, J.; Zhou, B.; Ye, Y.; et al. Complement Signals Determine Opposite Effects of B Cells in Chemotherapy-Induced Immunity. *Cell* **2020**, *180*, 1081–1097. [[CrossRef](#)]
62. Godfrey, D.I.; Kronenberg, M. Going both ways: Immune regulation via CD1d-dependent NKT cells. *J. Clin. Invest.* **2004**, *114*, 1379–1388. [[CrossRef](#)]
63. Kmiecik, M.; Basu, D.; Payne, K.K.; Toor, A.; Yacoub, A.; Wang, X.Y.; Smith, L.; Bear, H.D.; Manjili, M.H. Activated NKT cells and NK cells render T cells resistant to myeloid-derived suppressor cells and result in an effective adoptive cellular therapy against breast cancer in the FVBN202 transgenic mouse. *J. Immunol.* **2011**, *187*, 708–717. [[CrossRef](#)] [[PubMed](#)]
64. Favreau, M.; Vanderkerken, K.; Elewaut, D.; Venken, K.; Menu, E. Does an NKT-cell-based immunotherapeutic approach have a future in multiple myeloma? *Oncotarget* **2016**, *7*, 23128–23140. [[CrossRef](#)]
65. Lao, L.; Fan, S.; Song, E. Tumor Associated Macrophages as Therapeutic Targets for Breast Cancer. *Adv. Exp. Med. Biol.* **2017**, *1026*, 331–370. [[CrossRef](#)]
66. Zhou, J.; Tang, Z.; Gao, S.; Li, C.; Feng, Y.; Zhou, X. Tumor-Associated Macrophages: Recent Insights and Therapies. *Front Oncol.* **2020**, *10*, 188. [[CrossRef](#)]
67. Zhang, W.; Liu, L.; Su, H.; Liu, Q.; Shen, J.; Dai, H.; Zheng, W.; Lu, Y.; Zhang, W.; Bei, Y.; et al. Chimeric antigen receptor macrophage therapy for breast tumours mediated by targeting the tumour extracellular matrix. *Br. J. Cancer* **2019**, *121*, 837–845. [[CrossRef](#)] [[PubMed](#)]

- 
68. La Forgia, D.; Fanizzi, A.; Campobasso, F.; Bellotti, R.; Didonna, V.; Lorusso, V.; Moschetta, M.; Massafra, R.; Tamborra, P.; Tangaro, S.; et al. Radiomic Analysis in Contrast-Enhanced Spectral Mammography for Predicting Breast Cancer Histological Outcome. *Diagnostics* **2020**, *10*, 708. [[CrossRef](#)] [[PubMed](#)]
  69. Dilorenzo, G.; Telegrafo, M.; La Forgia, D.; Stabile Ianora, A.A.; Moschetta, M. Breast MRI background parenchymal enhancement as an imaging bridge to molecular cancer sub-type. *Eur. J. Radiol.* **2019**, *113*, 148–152. [[CrossRef](#)] [[PubMed](#)]