Neoadjuvant Chemotherapy Treatment Prediction: A Classification Model Based Approach Utilising Pre-treatment DCE-MRI

Martin D Pickles[1], Peter Gibbs[1], Martin Lowry[1], and Lindsay W Turnbull[1]
[1]*Centre for Magnetic Resonance Investigations, Hull York Medical School at University of Hull, Hull, East Yorkshire, United Kingdom*

Target audience: MR researchers, breast radiologists, oncologists, breast clinicians.

Purpose: Patients diagnosed with locally advanced breast cancer (LABC) regularly receive neoadjuvant chemotherapy (NAC) prior to surgery and adjuvant therapies [1]. Patients who achieve pathological complete response (pCR) have a favourable outcome [2]. Five year survival rates for stage III breast cancer patients are reported to be 72% [3]. If an MR biomarker could predict patients outcome (pCR vs. non-pCR) prior to NAC treatment then alternative treatment strategies, more aggressive treatment options for those predicted not to achieve pCR, could be considered in the hope of increasing both the pCR and five year survival rates.
The aim of this work was to develop a classification model to predict pCR, in a cohort of patients undergoing NAC treatment, from pre-treatment dynamic contrast enhanced MRI (DCE-MRI) data.

Classification model theories and challenges: Classification models attempt to predict class labels (e.g. pCR or non-pCR) based on information presented to the model [4]. Importantly, models should not be overfitted and must demonstrate good generalisation i.e. maintain high prediction accuracy when applied to data not used to generate the model [4]. Frequently, such models are developed via a 'training' set and evaluated on a 'test' set which importantly was not used in the development of the model [4]. Following NAC final pathological response varies with a lower proportion of patients demonstrating pCR [5]. This presents a problem for classification algorithms since the two classes are imbalanced. The performance of machine learning algorithms are often evaluated via predictive accuracy, however, this is inappropriate with imbalanced classes [6]. Under-sampling of the majority class has been proposed as a means of balancing the classes. However, with small datasets this increases the risk of model overfitting and a loss of generalisation [4,6]. Whereas the Synthetic Minority Over-sampling TEchnique (SMOTE) [6] can restore balance and synthetically increase the number of cases available to both the training and testing datasets while minimising the risk of model overfitting.

Methods: MR data was obtained from a 3.0T HDx scanner (GE Healthcare) prior to NAC. DCE-MRI data was acquired from a 3D dynamic VIBRANT sequence with a temporal resolution of ~30secs. Pseudo 3D volumes of interest (VOI) were generated via a semi-automated tumour segmentation tool. The signal intensity time course was interrogated in a pixel-by-pixel manner across all dynamic phases to generate empirical model free vascular parameters. To allow an assessment of the tumour vascular heterogeneity the following 1st order statistics were obtained: mean, standard deviation, skew, kurtosis, median, minimum and maximum.
Patients underwent a NAC regime (typically epirubicin and cyclophosphamide followed by docetaxel). Following surgery specimens underwent pathological analysis to determine response class.
SMOTE was utilised to balance the pCR and non-pCR classes. Subsequently, equal numbers of cases from each class were randomly assigned to either the training or testing data sets. To allow the development of a classification algorithm 84 DCE-MRI empirical parameters (12 DCE-MRI parameters x 7 1st order statistics) along with the appropriate class labels for each individual within the training set were presented to a decision tree. The performance of this decision tree was verified via the testing dataset and a confusion matrix was produced.

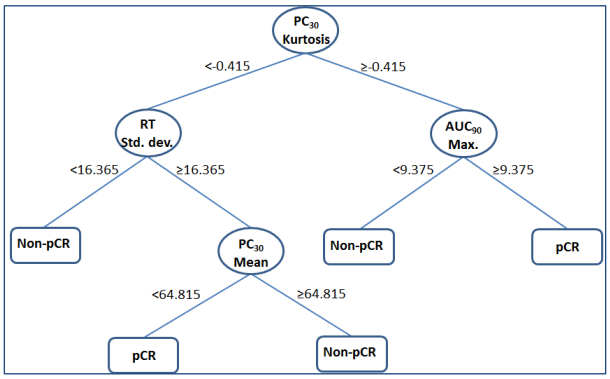

Fig 1. Classification decision tree

Results: Eight-nine patients underwent NAC followed by surgery. Following pathological analysis 11 patients were categorised as pCR and 78 as non-pCR. SMOTE synthetically increased the pCR class to 78 to match the non-pCR class. Figure 1 illustrates the resulting decision tree obtained from the training set. Table 1 details the summary statistics from the train and test decision tree classifiaction. Table 2 presents the confusion matrix from the test dataset. High predictive accuracy was obtained from only 4 DCE-MRI based decision tree nodes (PC30 kurtosis, RT Std. dev., PC30 mean and AUC90 max.).

	Sen.	Spec.	PPV	NPV	Acc.
Train	94.9%	97.4%	97.4%	95.0%	96.2%
Test	89.7%	82.1%	83.3%	88.9%	85.9%

Table 1. Train and test dataset summary statistics

	Pathological response		
	non-pCR	pCR	Totals
Model pCR	7	35	42
Model non-pCR	32	4	36
Totals	39	39	78

Table 2. Test dataset confusion matrix

Discussion: As expected the predictive accuracy of the training set was very high (96.2%) in part reflecting overfitting of the training data. Yet the predictive accuracy of the testing set remained high (85.9%), even though the model was not exposed to this data in its development. This high accuracy in the testing set demonstrates the good generalisation of the final decision tree model. If this model was used purely in isolation to tailor an individual's treatment then 7 patients would have been under-treated, 4 would have been over-treated and 67 would have been correctly stratified.

Conclusions: Prediction of pathological complete response, secondary to NAC treatment, can be made even prior to the initiation of chemotherapy from pre-treatment DCE-MRI parameters via a decision tree classification model with a 86% accuracy.

References:[1]Connolly RM et al. Eur J Pharmacol. 2013;717:58-66. [2]kong X et al. Eur J Cancer. 2011;47:2084-2090. [3]American Cancer Society Available at: http://www.cancer.org/cancer/%20breastcancer/detailedguide/breast-cancer-survival-by-stage. [4]Tan, PN. (2005). Introduction to Data Mining. Addison-Wesley. p145-205. [5]Zambetti M et al. Breast Cancer Res Treat. 2012;132:843–851. [6]Chawla NV et al. J Artif Intell Res. 2002;16:321-357.

Abbreviations: ***PC$_{30}$*** - percentage of the maximum enhancement index recorded 30 seconds from the onset of the enhancement curve. ***RT*** - time (minutes) to reach the half maximum enhancement index point from the start of the uptake curve. ***AUC$_{90}$*** - area under the enhancement factor curve 90 seconds from the onset of enhancement.