

# Hybrid Method for Prediction of Metastasis in Breast Cancer Patients Using Gene Expression Signals

Alireza Mehri Dehnavi<sup>1</sup>, Mohammad Reza Sehhati<sup>1</sup>, Hossein Rabbani<sup>1,2</sup>

<sup>1</sup>Department of Biomedical Engineering, and <sup>2</sup>Medical Image and Signal Processing Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

Submission: 17-03-2013 Accepted: 12-04-2013

## ABSTRACT

Using primary tumor gene expression has been shown to have the ability of finding metastasis-driving gene markers for prediction of breast cancer recurrence (BCR). However, there are some difficulties associated with analysis of microarray data, which led to poor predictive power and inconsistency of previously introduced gene signatures. In this study, a hybrid method was proposed for identifying more predictive gene signatures from microarray datasets. Initially, the parameters of a Rough-Set (RS) theory based feature selection method were tuned to construct a customized gene extraction algorithm. Afterward, using RS gene selection method the most informative genes selected from six independent breast cancer datasets. Then, combined set of these six signature sets, containing 114 genes, was evaluated for prediction of BCR. In final, a meta-signature, containing 18 genes, selected from the combination of datasets and its prediction accuracy compared to the combined signature. The results of 10-fold cross-validation test showed acceptable misclassification error rate (MCR) over 1338 cases of breast cancer patients. In comparison to a recent similar work, our approach reached more than 5% reduction in MCR using a fewer number of genes for prediction. The results also demonstrated 7% improvement in average accuracy in six utilized datasets, using the combined set of 114 genes in comparison with 18-genes meta-signature. In this study, a more informative gene signature was selected for prediction of BCR using a RS based gene extraction algorithm. To conclude, combining different signatures demonstrated more stable prediction over independent datasets.

**Key words:** Breast cancer recurrence prediction, gene expression signature, meta-signature, rough-set theory

## INTRODUCTION

The engaged patients with breast cancer recurrence (BCR) can range from 20% to 90% upon diagnosing and treating in different stages of the cancer.<sup>[1]</sup> Whatever certain treatments can reduce cancer recurrence risk, BCR prediction can help us to prevent overtreatment.<sup>[2]</sup> In the last decades, poor prediction of BCR using clinical factors has been prompted researchers to identify cancer markers through examination of genome-wide expression profiles.<sup>[3]</sup> Numerous studies have performed for extracting a combination of genes from messenger ribonucleic acid (mRNA) microarrays as biomarkers or cancer related genes. The entire have used expression level of these genes for prediction of relapse or distant metastases in breast cancer.<sup>[4-7]</sup> However, low accuracy of the published methods,<sup>[8,9]</sup> or limitations of the methods to clinically specific types of breast tumors,<sup>[9]</sup> were encouraged the researchers to identify more general, robust and accurate prognostic markers. Meanwhile, they considered combining previously introduced gene

signatures,<sup>[8]</sup> and meta-analysis of all available datasets,<sup>[5]</sup> to obtain a unified set of biomarkers.

In this paper, a hybrid method was built to predict BCR by finding informative gene sets regarding recurrence event based on Rough-Set (RS) theory.<sup>[10]</sup> Then, an appropriated supervised classifier was applied to the discovered genes. The RS method is preferred for gene selection in this study because of its proved power for dealing with vagueness in microarray data.<sup>[11,12]</sup> This method also preserves the meaning of the original feature sets and it has interpretability advantage comparing to traditional transform-based feature selection techniques.<sup>[13]</sup> A key requirement for successful usage of RS is appropriate data discretization. Therefore, all expression signals were discretized in three levels based on previously successful applications.<sup>[14]</sup> Recently, different techniques have been applied to RS algorithm to relief its inadequacies. However, these methods are data dependent and should be modified for gene selection from microarray data. In this regard, various criteria like fuzzy entropy,<sup>[15-17]</sup>

### Address for correspondence:

Dr. Alireza Mehridehnavi, Department of Biomedical Engineering, Medical Image & Signal Processing Research Center, Isfahan University of Medical Sciences, Isfahan, Iran. E-mail: mehri@med.mui.ac.ir

dependency and consistency,<sup>[11]</sup> were investigated and the final RS method was constructed using appropriate parameters for generalized dependency function based on theta-eta model of RS domain.<sup>[11]</sup> For this purpose, different gene sets have been extracted from Wang dataset,<sup>[3]</sup> using different parameters. Thereafter, the optimized algorithm applied on different datasets for extracting the indicator genes.

There is a small overlap between published gene signatures for BCR prediction, which reflects a small chance to reach a general and robust indicator, set to judge about all experimental data from different studies. So we decided to combine the extracted genes of different datasets together. This approach has less prediction performance in comparison to apply extracted genes for specific dataset, but overall performance among all datasets will be increased. Among published gene-expression-based prediction methods, we have chosen the work of Li *et al.*,<sup>[7]</sup> that has one of the best recent claims in robustness and accuracy for overall comparison with our algorithm.

The remainder of this paper is organized as follows: In Section 2, Utilized datasets introduced and applied normalization procedure described. Subsequently details of gene selection procedure and classification of samples reported in this section. Experimental results are reported in Section 3. In Section 4, we discussed on the obtained results, then concluded our work and outlined some future work directions.

## MATERIALS AND METHODS

In this section, an overview of the whole study will be described by introducing the utilized data in this work, and details of the proposed approach step by step.

### Breast Tumor Datasets

We utilized human breast cancer microarray datasets of six studies with the same platform including 1338 samples that are publicly available from gene expression omnibus (GEO) database and will be referenced later in this paper by their GEO series code (GSE xxx) as illustrated in Table 1. Affymetrix gene-chip human genome U133 array (HG-U133A) set, also known as GEO platform 96 (GPL96), microarrays were taken into account because of frequent usage and abundance of samples. This selection helps us to overcome many problems such as non-overlapped transcripts, different precision and varying relative scales, and distinctive dynamic ranges of gene expression among different platforms. At the first step, the presented transcripts on the microarray have been pruned. In this regard, the genes that were not present significantly across all samples,<sup>[3]</sup> and samples that were censored and had no information about relapse or metastasis have been removed from all datasets. In final,

Table 1: Summary of breast cancer microarray datasets

Dataset	#Samples	#High-risk	#Low risk	Source
GSE2034	286	95	169	Wang <i>et al.</i> , 2005 <sup>[3]</sup>
GSE7390	198	54	100	Desmedt <i>et al.</i> , 2007 <sup>[18]</sup>
GSE2990	244	65	123	Loi <i>et al.</i> , 2007 <sup>[19]</sup>
GSE4922	249	70	139	Ivshina <i>et al.</i> , 2006 <sup>[20]</sup>
GSE6532	236	37	158	Miller <i>et al.</i> , 2005 <sup>[21]</sup>
GSE3494	125	30	67	Sotiriou <i>et al.</i> , 2006 <sup>[22]</sup>
Metadata	1338	470	868	

Table samples were classified into two groups, high and low risk, according to the time to metastasis using a threshold of 5 years; # – Number of; GSE – GEO series

metadata has been constructed by combining all utilized datasets.

### Microarray Data Normalization

The most important pre-processing step is a normalization of expression signals. In this regard expression values were  $\log_2$  transformed, then the base signal ( $\log_2(600)$  or  $\log_2(500)$ ), have been subtracted from all data. After this, maximum positive signal mapped to +1 and minimum negative signal mapped to -1. Before determining the differentially expressed genes as the next step, we discretized the expression values at three levels. This takes advantage of lightening the undesirable effect of noisy data and produces more reliable results when working with different datasets.<sup>[23]</sup> Following the procedure of fuzzy discretization technique,<sup>[14]</sup> the performance of fuzzy RS gene selecting step improved as described at the next subsection.

### Selection of Prognoses Indicator Genes

RS theory is a new fashion to deal with uncertainty and incompleteness so that it has become a favorable technique for feature selection; especially in microarray data analysis.<sup>[11-13]</sup> Using RS theory for gene selection we can screen informative and related genes to our desired output. So we can find a set of genes that their expressions among diverse samples have significant predictive power regarding recurrence event. Among various RS based feature selection methods that proposed before,<sup>[11-13]</sup> we generated multiple subset of genes by applying various sub-methods with different parameters on the whole gene set. After primary investigation of specific parameters of the algorithm,<sup>[11]</sup> these parameters were tuned to gain the best classification performance on Wang data set. It should be noted that the 10-fold cross-validation (CV) test was used for tuning the algorithm's parameters.

The customized RS gene selection algorithm applied to different datasets independently and demonstrated different gene signatures with different accuracies. Working in high dimension and sensitivity of feature selection method to sample cohorts are significant issues in this procedure

that resulted in extracting different signature genes from different datasets. A set of 18 genes also selected from normalized combination of all samples from all datasets. Li *et al.*, considered voting by different sets and claimed that simply combining the sets did not improve the results. However, as it was discussed, we observed acceptable improvement in robustness with combining extracted gene sets from different datasets.

### Classification by Neuro-fuzzy Inference System

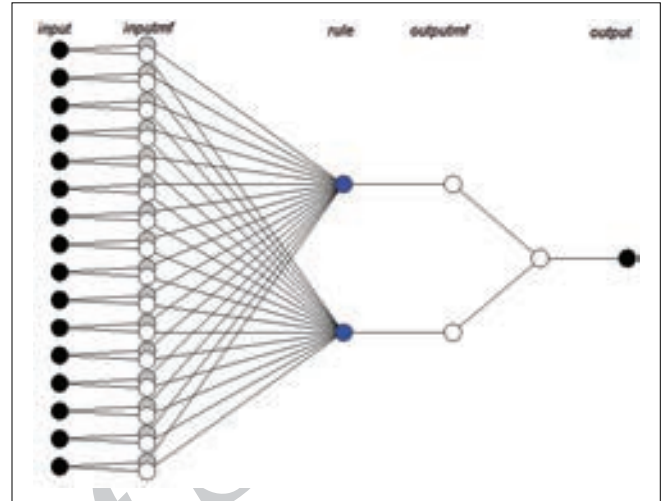
Choosing appropriate classifier has also great importance that affects the total performance of the proposed algorithm for BCR prediction. Numerous supervised classifiers have been used for classifying cancer-based gene expression data.<sup>[24]</sup> However, less attention have been paid on FISs, which have been successfully applied in many different areas.<sup>[17]</sup> They have been suggested as a commanding technique for dealing with noisy data with complex interactions.<sup>[17]</sup> In this regard, the adaptive neuro-fuzzy inference system (ANFIS) that was first proposed by Jang,<sup>[25]</sup> was used for identifying parameters of a FIS. The combination of RS theory and ANFIS called hybrid method in this work.

We utilized the MATLAB fuzzy toolbox for generating a sugeno FIS with its default settings which previously proposed by Cetisli.<sup>[26]</sup> The obtained ANFIS model structure was shown in Supplementary Figure 1. In this regard, the grid partitioning procedure was used for subdividing the input space and generating the rules. Hybrid learning algorithm, which is a combination of the back-propagation gradient descent procedure and least-squares method, was applied for training the FIS membership function parameters. The performance evaluation plot for training ANFIS on meta-signature for 100 epochs of training was shown in Supplementary Figure 2.

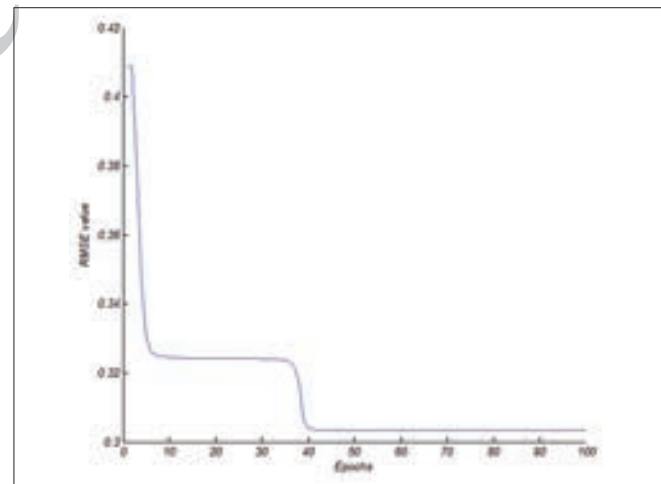
## RESULTS

We utilized microarray datasets from six human breast cancer studies for gene set extraction using RS feature selection algorithm with tuned parameters on Wang dataset. Then, a 10-fold CV test was performed for evaluation of the gene sets, which extracted from all datasets, using a

neuro-fuzzy classifier. It should be noted that 10-fold CV test also used in preparing of RS feature selection method and mapping steps of ANFIS model. Table 2 reported the MCR, which obtained for prediction of recurrence over the validation dataset, using the selected signature from reference dataset. Number of extracted genes from each



**Supplementary Figure 1:** The adaptive neuro-fuzzy inference system model structure. In this structure there are 16 Gaussian membership functions for input nodes and two Gaussian MFs for middle and output nodes. Settings for fuzzy inference system are: And = "prod"; Or = "probor"; Defuzzifier = "wtaver"; Implication = "prod"; Aggregation = "sum"



**Supplementary Figure 2:** Performance evaluation in 100 epochs of training

**Table 2:** Misclassification error rate of 10-fold independent cross-validation in six breast cancer studies

Reference dataset	Number of genes	Validation datasets					
		GSE2034	GSE3494	GSE2990	GSE4922	GSE6532	GSE7390
GSE2034	24	0.22	0.21	0.48	0.41	0.44	0.49
GSE3494	23	0.40	0.25	0.39	0.40	0.39	0.46
GSE2990	12	0.38	0.26	0.29	0.37	0.36	0.50
GSE4922	20	0.33	0.30	0.36	0.36	0.49	0.43
GSE6532	14	0.39	0.26	0.42	0.40	0.34	0.50
GSE7390	21	0.40	0.40	0.40	0.37	0.40	0.22

GSE – GEO series

datasets was different as it was shown at the second column of Table 2. The number of extracted biomarkers by RS method depends on the threshold of dependency in the algorithm that obtained by primary optimization of classification error rate on Wang dataset. In this way, a local maximum in accuracy obtained using the selected number of genes. In connection with this, adapting looser thresholds or increase the number of selected features will help us to reach to more general results with more overlap among other datasets. However, by increasing the dimension, according to the low number of available samples, the over fitting problem would override the classification results.<sup>[27]</sup>

We also evaluated the classification accuracies after constructing two new signature sets. The first one, containing 114 genes, was constructed by combining all of six gene signatures together. The second gene signature, containing 18 genes, was constructed by applying RS feature selection algorithm to the metadata. Table 3 showed the MCR results of 10-fold CV test for the combined and meta-signature over all datasets. As it is shown in Table 3, the best result was obtained for Sotiriou dataset at MCR = 0.27 using combined gene signature.

We also applied the introduced gene sets by Li *et al.*,<sup>[7]</sup> called national research council (NRC) sets, to the fuzzy classifier to compare the performance of their gene sets with resultant signatures of this work. It should be noted that there is no common gene between selected genes in two studies. Because the Wang dataset have been used by Li *et al.*,<sup>[7]</sup> as training data and they gave the best results on this dataset we chose this dataset for a fair comparison. Table 4 reported the 10-fold CV MCR using nine different gene signatures, which introduced by Li *et al.*, and evaluated over the Wang dataset. As it is shown in Table 4, the best classification accuracy was reached at MCR = 0.29 by the NRC7 signature, which extracted from negative estrogen receptor (ER-) samples. Li *et al.*,<sup>[7]</sup> reached at 87% of accuracy for classification on low-risk patients.

However, in this work the overall MCR on all samples were evaluated and reported in the results. Two rows of Table 4 corresponds to the two different normalization methods, one used by Li *et al.*, [second row of Table 4], and the other was used to obtain the presented results in Tables 2 and 3. We also applied the nearest shrunken centroid classification method that have been used by Li *et al.*, using our extracted genes from Wang dataset. After this, we encountered almost with the same overall MCR (<0.32) after performing leave-one-out CV test with all six resultant gene sets of the proposed approach.

We also considered the effect of ER – status on prediction accuracy by separating samples according to ER status. In concordance with other studies accuracy will be improved up to 6%, when ER status taken into account. The selected genes from different samples and related gene ontology analysis were described in the Supplementary File 1.

In order to show the association between introduced gene sets and survival, the samples are grouped according to the output of classifier that applied to the expression of selected genes. After that, the two groups, with known time of relapse or censoring are contrasted by a Kaplan–Meier plot. Figure 1 shows the probability of survival versus time from primary detection of tumor, for two groups of patients classified by the proposed method. Figure 1a shows survival curve obtained from 286 samples of Wang dataset using total 114 combined signature gene. We also plotted the survival curve for combined data of all datasets containing 1338 samples in Figure 1b. It can be seen in Figure 1 that two groups are very distinctive for both datasets with a little decreasing in the significant of distinction between samples with poor and good outcomes in metadata. We used Wilcoxon rank sum test for calculating *P* value. Survival curves for other five datasets [Supplementary Figures 3-8] imply the power of combined gene signatures in significant distinction between patients with distant metastasis and relapse-free ones, which were classified by the proposed method.

**Table 3: Misclassification error rate of 10-fold independent cross-validation in six breast cancer studies with combined set and meta-selected genes**

Feature set	Number of genes	Validation dataset (GEO series code)					
		GSE2034	GSE3494	GSE2990	GSE4922	GSE6532	GSE7390
Combined	114	0.28	0.27	0.32	0.34	0.33	0.33
Meta	18	0.40	0.29	0.39	0.40	0.39	0.46

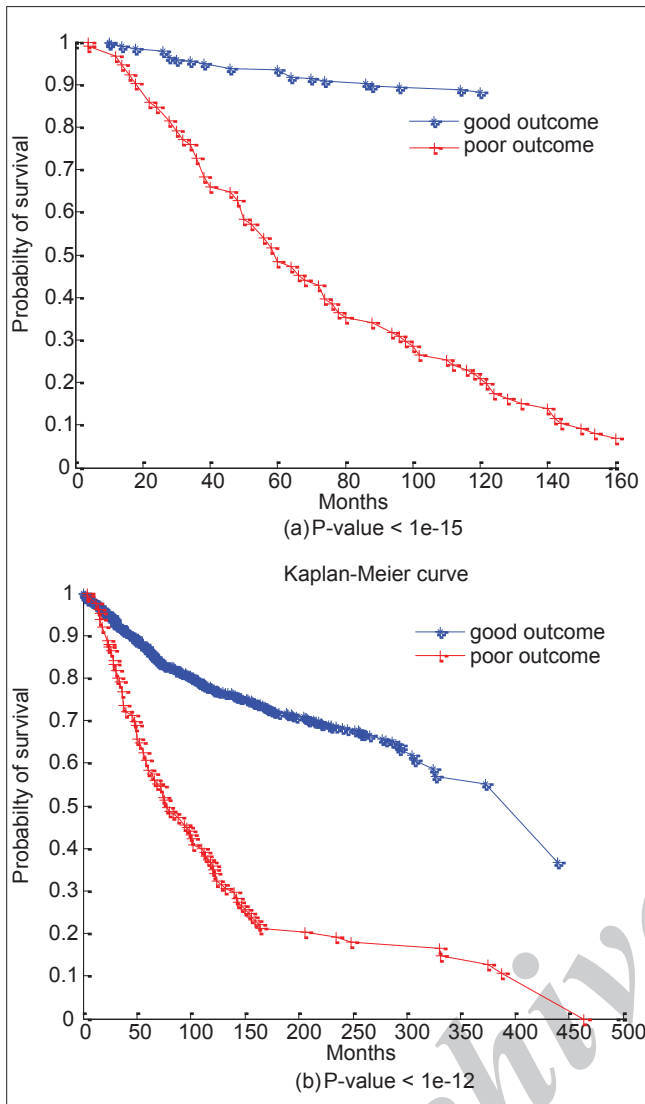
GEO – Gene expression omnibus; GSE – GEO series

**Table 4: Misclassification error rate of 10-fold cross-validation in Wang dataset (GSE2034) with Li gene sets (NRCx)**

Normalization/gene signature	NRC1 (ER+)	NRC2 (ER+)	NRC3 (ER+)	NRC4 (ER+)	NRC5 (ER+)	NRC6 (ER+)	NRC7 (ER-)	NRC8 (ER-)	NRC9 (ER-)
Normalization 1	0.38	0.35	0.33	0.41	0.40	0.39	0.29	0.33	0.33
Normalization 2	0.37	0.36	0.31	0.40	0.40	0.35	0.31	0.30	0.5

NRCx – National research council; GSE – GEO series; ER – Estrogen receptor

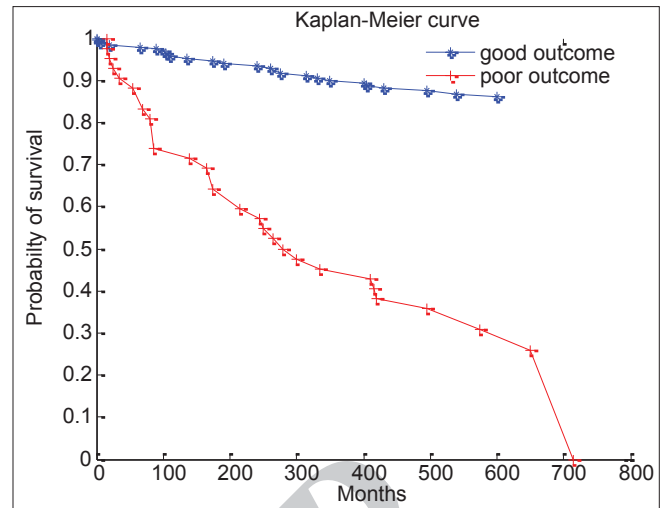




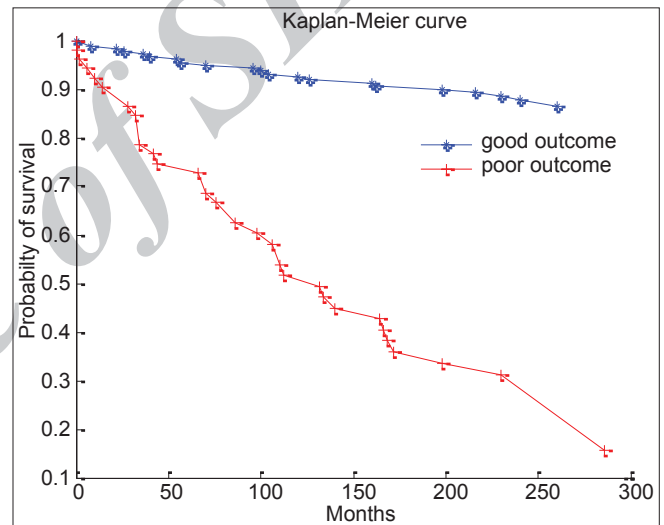
**Figure 1:** Kaplan-Meier relapse-free survival curves. (a) Classified samples on Wang dataset with 114-genes combined signature; (b) Classified samples on combined dataset of 1338 samples with 114-genes combined signature

## DISCUSSION

In this paper, selection of constructive genes and choosing an appropriate classifier helped us to present a prosperous method for prediction of breast cancer relapse. In this regard, a RS based feature selection method was proposed. Then a fuzzy classifier was used for classification of metastatic and non-metastatic cohorts in six independent breast cancer microarray datasets. According to Table 2 our approach reached acceptable MCR in testing cross accuracy of the introduced gene signatures. We also combined the extracted genes by RS algorithm and compared the results by extracted features from metadata. Table 3 demonstrated that combining extracted signatures from independent datasets reached at 7% improvement in average accuracy in comparison to 18-genes meta-signature, which selected from metadata.



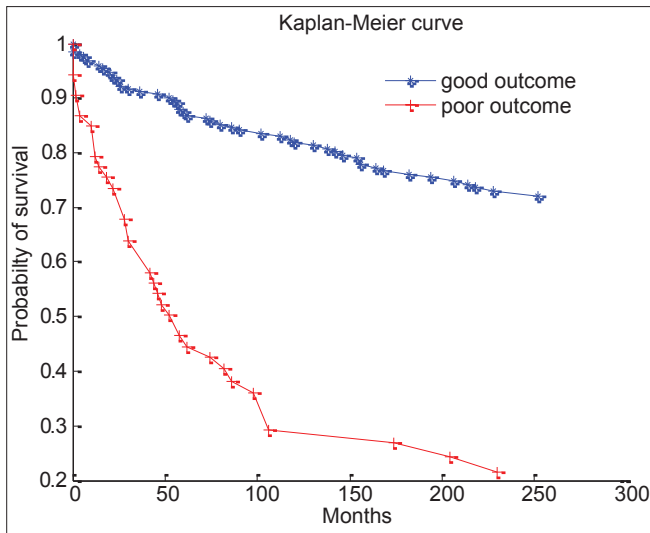
**Supplementary Figure 3:** KM plot for classified samples of dataset GSE3494 using 23 genes extracted from this dataset ( $P = 3 \text{ e-}05$ )



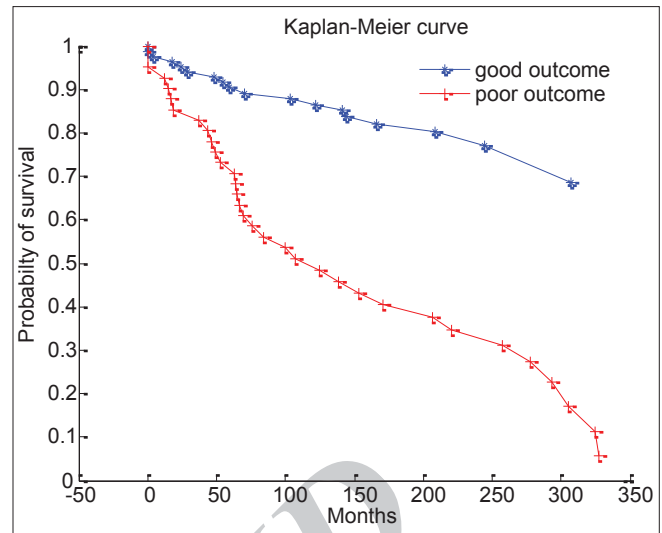
**Supplementary Figure 4:** KM plot for classified samples of dataset GSE3494 using 114 genes from six datasets ( $P = 1 \text{ e-}07$ )

Recently, Li *et al.*,<sup>[7]</sup> introduced nine gene signature at equal size 30 and reported good results for classification of low risk patients. According to Table 4, the best classification accuracy was reached at MCR = 0.29 by the NRC7 signature, which extracted from negative ER – samples. Comparing the obtained results in Tables 2 and 3 and 4 demonstrated the better accuracy of our approach. To be specific, MCR after 10-CV on Wang dataset was 0.22 for our single gene set and 0.28 for the combined set that is better than the minimum value of 0.29, which obtained by NRC sets.

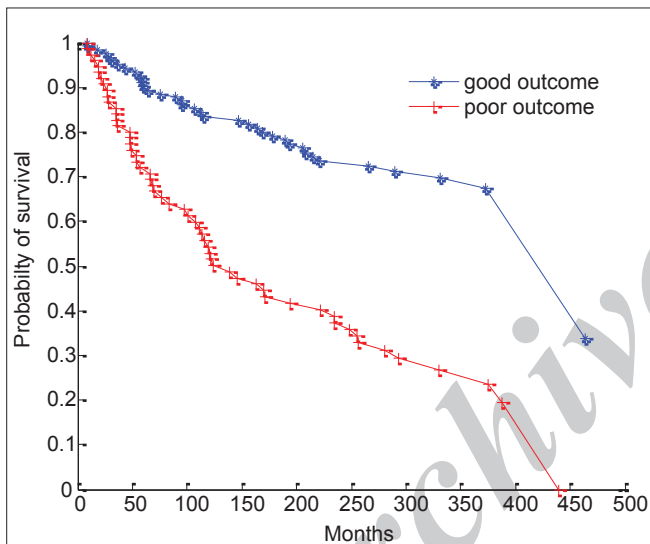
Wei and Li<sup>[24]</sup> reported the comparison results of the average MCR of their proposed classifier and nine commonly used procedures based on 10-fold CV for three breast cancer data sets. Reported results by Wei and Li show that the best MCR among these ten different classifiers was 0.29 for Wang



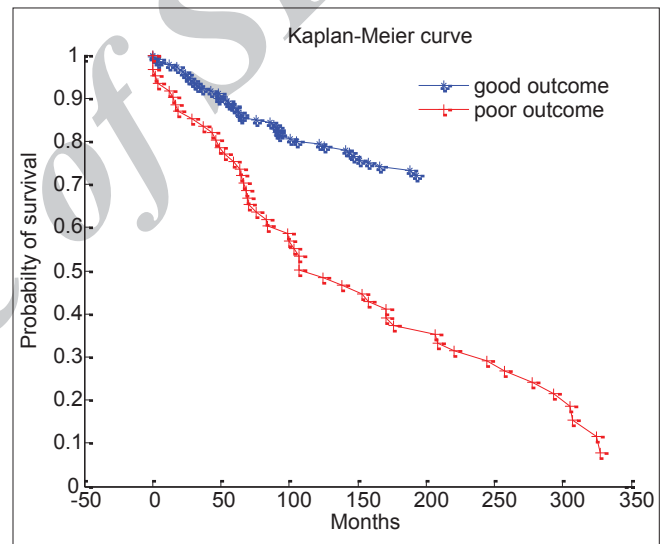
**Supplementary Figure 5:** KM plot for classified samples of dataset GSE4922 using 114 genes from six datasets ( $P = 3 \text{ e-}09$ )



**Supplementary Figure 6:** KM plot for classified samples of dataset GSE2990 using 114 genes from six datasets ( $P = 0.01$ )



**Supplementary Figure 7:** KM plot for classified samples of dataset GSE7390 using 114 genes from six datasets ( $P = 6 \text{ e-}06$ )



**Supplementary Figure 8:** KM plot for classified samples of dataset GSE6532 using 114 genes from six datasets ( $P = 0.09$ )

dataset, which is lower in accuracy from that we reached by the proposed hybrid method.

Rely on literature and the presented results, using any algorithm for extracting informative genes from microarray, diverse set of genes were selected from different datasets, which all of the led to nearly acceptable classification performance. Because of a wide variety and heterogeneity in microarray experiments we cannot have rational reasoning about extracted genes and introduce extracted genes from a specific study as general biomarkers. According to the represented results on survival analysis [Figure 1] and CV test on different combination of gene signatures [Tables 2 and 3] there is a trade-off between robustness and accuracy of prediction. In this regard, combining gene signatures led to bigger gene signature with more robust prediction

results. However, adding extra genes to the signature set, which have no useful information for classification, decreased the prediction accuracy. Moreover, when we cannot find a unique gene set, only a general model can help us to interpret the selected biomarkers, biologically. In this regard, model-based algorithms and their limitations should be considered in the future works.

## REFERENCES

1. Mayer M. Treatment and outcomes for high-risk and metastatic breast cancer in California: An inquiry into disparities and research needs. *Cal Breast Cancer Res* 2003. Available from: <http://www.cbcrp.org/publications/papers/Mayer/index.php>.
2. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.

3. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;365:671-9.
4. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* 2005;436:518-24.
5. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, *et al.* Meta-analysis of gene expression profiles in breast cancer: Toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res* 2008; 10:R65.
6. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, *et al.* Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat* 2012;132:523-35.
7. Li J, Lenferink AE, Deng Y, Collins C, Cui Q, Purissima EO, *et al.* Identification of high-quality cancer prognostic markers and metastasis network modules. *Nat Commun* 2010;1:34.
8. Zhao X, Rødland EA, Sørli T, Naume B, Langerød A, Frigessi A, *et al.* Combining gene signatures improves prediction of breast cancer survival. *PLoS One* 2011;6:e17845.
9. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, *et al.* Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 2006;355:560-9.
10. Pawlak Z. *Rough Sets; Theoretical Aspects of Reasoning About Data.* Dordrecht: Kluwer Academic Publishers; 1991.
11. Mayer B. *Bioinformatics for Omics Data: Methods and Protocols.* London: Humana Press; 2011.
12. Qinghua H, Daren Y, Jinfu L, Congxin W. Neighborhood rough set based heterogeneous feature subset selection. *Inf Sci* 2008;178:3577-94.
13. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res* 2003;3:1157-82.
14. Sokhansanj BA, Fitch JP, Quong JN, Quong AA. Linear fuzzy gene network models obtained from microarray data by exhaustive search. *BMC Bioinformatics* 2004;5:108.
15. Feifei XU, Jingsheng L, Lai W. Improved mutual information-based gene selection with fuzzy rough sets. *J Comput Inf Syst* 2011;7: 3166-73.
16. Xu D, Keller JM, Popescu M, Bondugula. *Applications of Fuzzy Logic in Bioinformatics.* London: Imperial College Press; 2008.
17. Huerta EB, Duval B, Hao JK. Fuzzy logic for elimination of redundant information of microarray data. *Genomics Proteomics Bioinformatics* 2008;6:61-73.
18. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, *et al.* Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* 2007;13:3207-14.
19. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, *et al.* Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* 2007;25:1239-46.
20. Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, *et al.* Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res* 2006;66:10292-301.
21. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, *et al.* An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* 2005;102:13550-5.
22. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, *et al.* Gene expression profiling in breast cancer: Understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006;98:262-72.
23. Steinhoff C, Vingron M. Normalization and quantification of differential expression in gene expression microarrays. *Brief Bioinform* 2006;7:166-77.
24. Wei Z, Li H. Nonparametric pathway-based regression models for analysis of genomic data. *Biostatistics* 2007;8:265-84.
25. Jang JS. ANFIS: Adaptive-network-based fuzzy inference systems. *IEEE Trans Syst Man Cybern* 1993;23:665-85.
26. Cetisli B. Development of an adaptive neuro-fuzzy classifier using linguistic hedges: Part 1. *J Expert Syst Appl* 2010;37:6093-101.
27. Mehridehnavi A, Ziaei L. Minimal gene selection for classification and diagnosis prediction based on gene expression profile. *Adv Biomed Res* 2013;2:26.

**How to cite this article:** Dehnavi AM, Sehhati MR, Rabbani H. Hybrid method for prediction of metastasis in breast cancer patients using gene expression signals. *J Med Sign Sens* 2012;3:79-86.

**Source of Support:** Nil, **Conflict of Interest:** None declared

## BIOGRAPHIES



**Alireza Mehri Dehnavi** was born in Isfahan province at 1961. He had educated in Electronic Engineering at Isfahan University of Technology at 1988. He had finished Master of Engineering in Measurement and Instrumentation at Indian Institute of Technology Roorkee (IIT Roorkee) in India at 1992. He has finished his PhD in Medical Engineering at Liverpool University in UK at 1996. He is an Associate Professor of Medical Engineering at Medical Physics and Engineering Department in Medical School of Isfahan University of Medical Sciences. He is currently visiting at School of Optometry and Visual Science at University of Waterloo in Canada. His research interests are medical optics, devices and signal processing

**E-mail:** mehri@med.mui.ac.ir



**Mohammad Reza Sehhati** was born in Isfahan, Iran, in 1981. He received the B.S and M.S degree in Biomedical Engineering from Shahed University and University of Tehran, Tehran, Iran, respectively. He is

now Ph.D Candidate of Biomedical Engineering in Isfahan University of Medical Sciences, Isfahan, Iran. His research interests include Bioinformatics, Machine Learning, Data Mining, Image Processing, and Hospital Information Systems

**E-mail:** sehhati@resident.mui.ac.ir



**Hossein Rabbani** is an Associate Professor at Isfahan University of Medical Sciences, in Biomedical Engineering Department also Medical Image & Signal Processing Research Center. Involved research topics include medical image/volume processing, noise reduction and estimation problem, image enhancement, blind deconvolution, video restoration, probability models of sparse domain's coefficients especially complex wavelet coefficients. He is a member of IEEE, Signal Processing Society, Engineering in Medicine and Biology Society, and Circuits and Systems Society

**E-mail:** h\_rabbani@med.mui.ac.ir

Archive of SID