International Journal of Innovative Research in Medical Science (IJIRMS)
Volume 02 Issue 07 July 2017, ISSN No. - 2455-8737
Available online at - www.ijirms.in

Open Access Journal | Research Article | DOI: 10.23958/ijirms/vol02-i07/15

A Subclassification of Basal-like Breast Cancer for Prognostic Prediction

Tomoko Tsuruga, Keiko Oda, Hidekazu Yoshie, Ayano Shinagawa, Kimino Minagawa, Toshio Kumai, Ko Sato*, Anna S Sedukhina*

Department of Pharmacogenomics, St. Marianna University, 2-16-1 Sugao, Miyamae, Kawasaki, 2168511, Japan

Abstract:

Gene expression profiling can predict the clinical outcome of breast cancer. However, further classification is needed to improve the prediction of prognosis. For further classification, bioinformatics analysis using The Cancer Genome Atlas (TCGA) dataset was performed to determine the effect of the upregulation of basal-like breast cancer-related genes in the PAM50 dataset on survival. Survival-related genes were further evaluated in a large cohort in the Kaplan-Meier plotter (KMplot) dataset. These analyses revealed that increased expression of EXO1, a basal-like breast cancer-related gene, was linked to longer overall survival in basal-like breast cancer [hazard ratio (HR) = 0.135, p < 0.0001 in TCGA and HR = 0.479, p = 0.036 in KMplot] and remained significant in multivariate analysis. Therefore, we propose that basal-like breast cancer can be further divided into two groups with different prognoses based on EXO1 expression.

Keywords: ***Breast cancer, intrinsic subtype, EXO1, prognosis***

Introduction

Since the introduction of intrinsic molecular subtypes of breast cancer based on gene expression patterns, the implications of this classification have been studied extensively.[1,2] The classification of breast cancer into intrinsic subtypes is helpful to predict the clinical prognosis.[2] Several assay systems based on the use of simplified gene sets were established to translate this classification into clinical practice. The MammaPrint microarray assay provides a prognostic prediction in node-negative early stage breast cancer by analyzing the expression of 70 genes.[3,4] The OncotypeDX assay determines the risk of distant recurrence in node-negative estrogen receptor-positive breast cancer according to the expression of 21 genes.[5] The PAM50, which measures the expression of 50 genes to predict risk of recurrence (ROR)[6], can be used for the classification of breast cancer into intrinsic subtypes in addition to the prediction of ROR.

Precision oncology or biomarker-driven therapy, which describes strategies aimed at targeting a single molecule or pathway rather than gene expression patterns, has attracted attention because of its high efficacy and few adverse effects.[7] Trastuzumab, a precision oncology strategy for the treatment of breast cancer, greatly improved the clinical outcome of Her2-positive breast cancers.[8,9] Therefore, gene expression patterns are currently recognized as prognostic factors rather than predictive factors, and additional classification methods for the precise prediction of prognosis are needed.

Here, the possible subdivision of basal-like breast cancer, a subtype linked to poor clinical outcomes, was investigated by bioinformatics analysis of 30 basal-like breast cancer-related genes from the PAM50.[2] *EXO1*, a gene in the PAM50 gene set, is often overexpressed in breast cancer and linked to shorter survival.[10,11] In the present study, unbiased analysis revealed that increased expression of *EXO1* is linked to longer survival in basal-like breast cancer.

Materials and Methods

Bioinformatics

Gene expression data (z-score) from breast invasive carcinoma samples from The Cancer Genome Atlas (TCGA) (RNA sequencing) were downloaded from cBioPortal (www.cbioportal.org).[12] The Kaplan-Meier plotter (KMplot) data were downloaded from the KMplot site (http://kmplot.com/analysis/).[13] The heatmap was generated with the "Heatplus" package in R using gene expression data (z-score). Hierarchical clustering was performed using the "Heatplus" package in R. The forest plot was constructed using the "metafor" package in R.

Statistics

For survival analysis, a Cox proportional hazard model was used to analyze forest plot and multivariate data. The log-rank test was used for all other survival analyses. The Cox

DOI: 10.23958/ijirms/vol02-i07/15

proportional hazard model was fitted using the coxph function (located in the "survival" package in R). The log-rank test was performed using Graphpad prism. Weighted kappa statistics were performed using a web tool available at *Vassar Stats: Website for Statistical Computation* (http://vassarstats.net).

Table 1: Multivariate analysis of *EXO1* expression on OS

	Unadjusted		Adjusted	
	HR (95%CI)	***p*-value**	**HR (95%CI)**	***p*-value**
EXO1	**0.10 (0.03-0.39)**	***** 9.64E-04**	**0.10 (0.02-0.56)**	**** 8.43E-03**
Age	**1.01 (0.96-1.06)**	**0.76**	**0.96 (0.87-1.05)**	**0.34**
Stage	**1.39 (0.80-2.43)**	**0.25**	**1.15 (0.51-2.58)**	**0.73**

OS: overall survival
HR: hazard ratio
CI: confidence interval
Age and stage were evaluated as a continuous variable.

Abbreviations

TCGA: The Cancer Genome Atlas
KMplot: Kaplan-Meier plotter
HR: hazard ratio
ROR: risk of recurrence
OS: overall survival

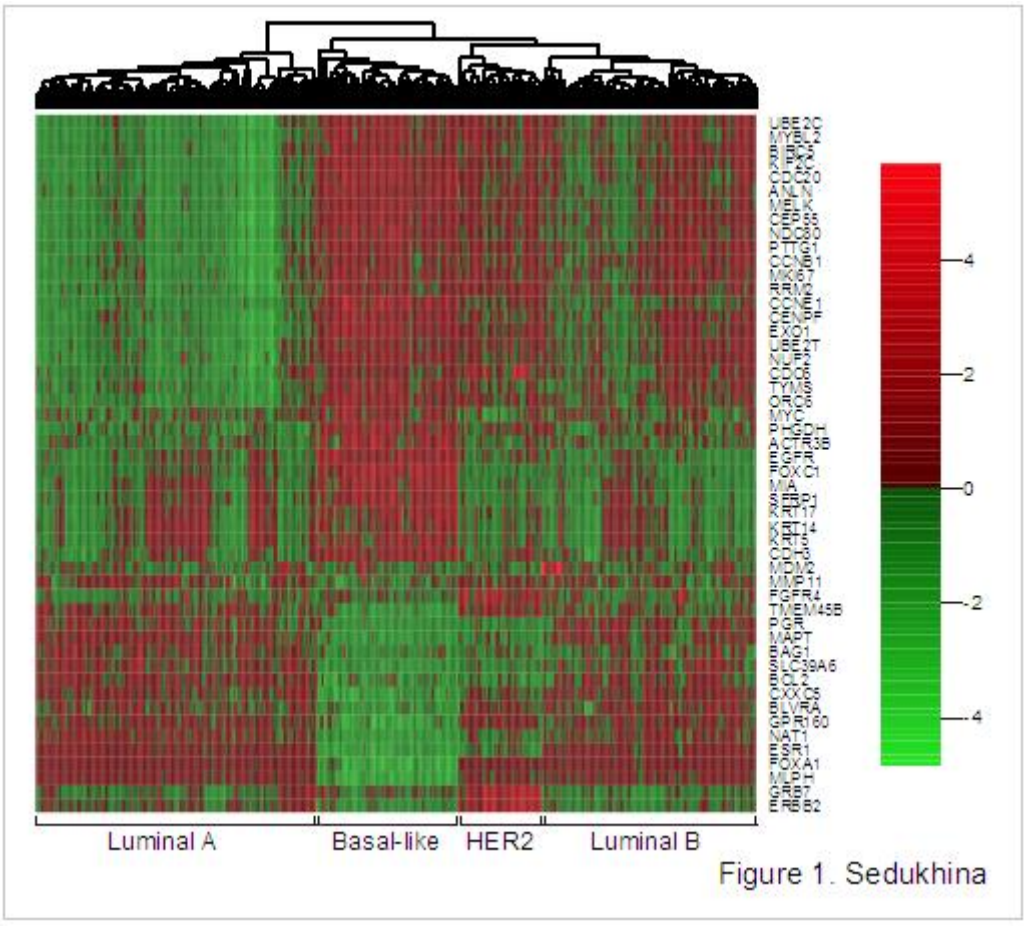

Figure 1: Hierarchical clustering of the TCGA dataset based on the PAM50

The heat map displays the expression pattern of 50 genes in the TCGA dataset. The hierarchical clustering divided the cohort into four subtypes.

DOI: 10.23958/ijirms/vol02-i07/15

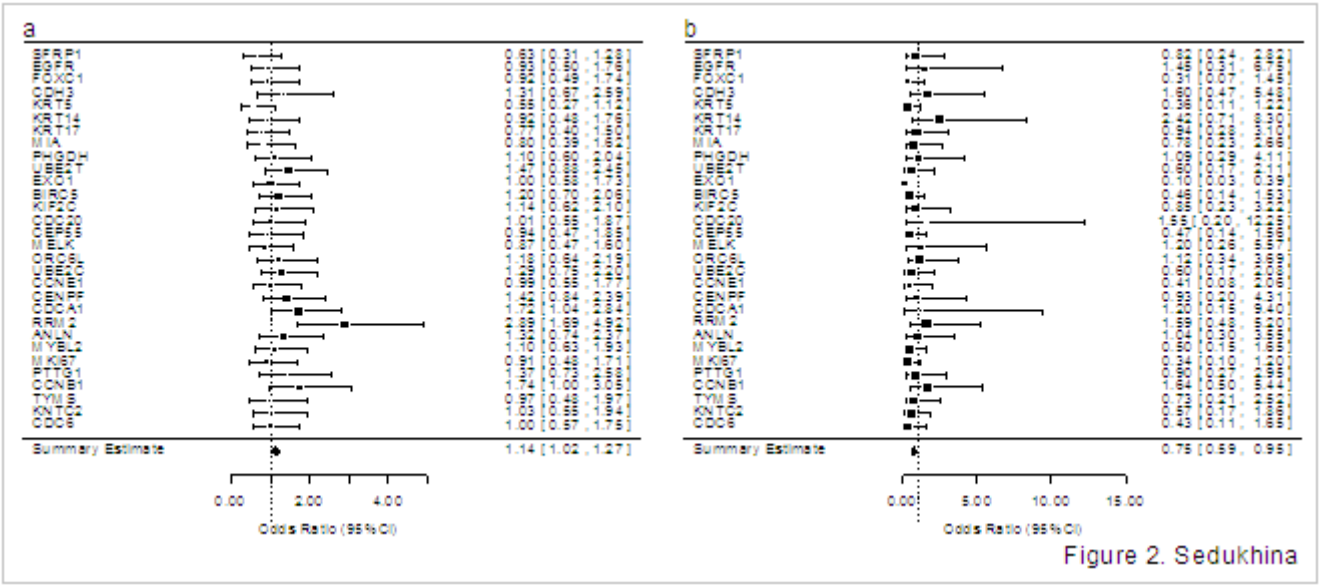

Figure 2: Effect of basal-like breast cancer-defining gene expression on survival

Forest plots show overall survival hazard ratio when genes are overexpressed (z-score ≥1) in all breast cancer cases (A), and in basal-like breast cancer (B).

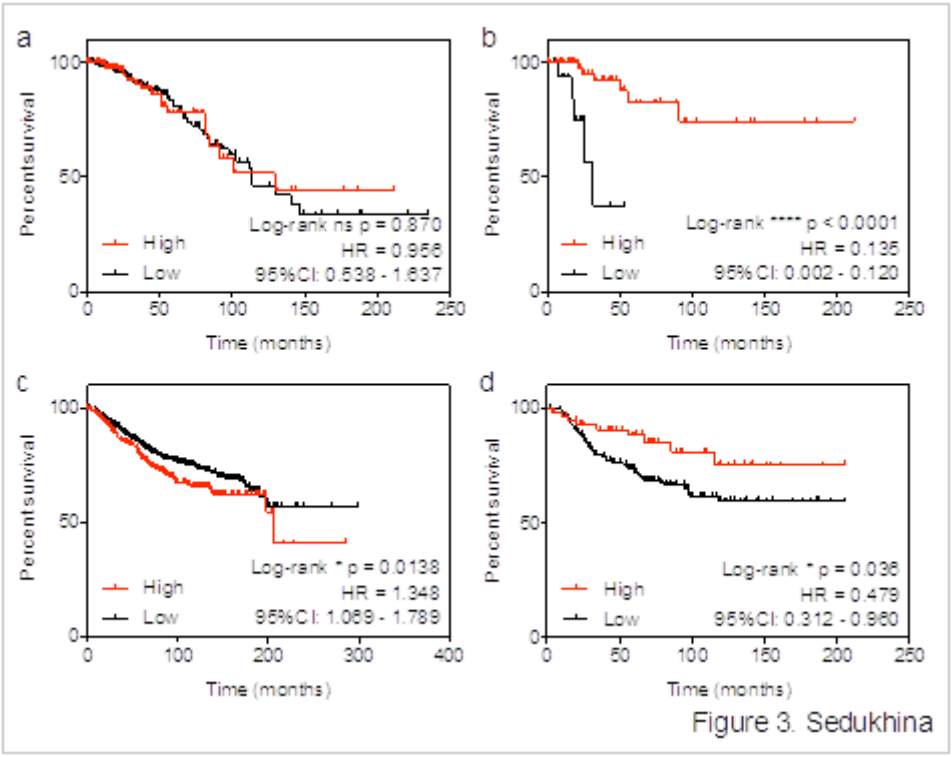

Figure 3: Effect of *EXO1* expression on basal-like breast cancer survival

Overall survival curves of cohorts with or without increased expression of *EXO1* in total breast cancer (A) or basal-like breast cancer (B) in the TCGA dataset and in total breast cancer (C) or basal-like breast cancer (D) in KMplot are shown. The definition of increased expression was based on a z-score ≥1 in the TCGA dataset and upper quartile in KMplot.

Results

Classification based on PAM50

The 511 cases of breast adenocarcinoma in the TCGA dataset were classified into intrinsic subtypes using RNA sequencing data (z-score) from the PAM50 (TCGA Nature 2012).[14] Hierarchical clustering divided the cohort into four groups: luminal A (229 cases), luminal B (124 cases), Her2-enriched (57 cases), and basal-like (97 cases) (Figure 1). The accuracy of this classification was confirmed using a

DOI: 10.23958/ijirms/vol02-i07/15

classification based on microarray data from the same cohort.[14] The concordance was confirmed using the weighted kappa statistic. The results suggested that the present classification was reliable (weighted kappa statistic = 0.8713).

Effect of basal-like breast cancer-related gene expression on survival

Thirty genes (keratin genes: *SFRP1*, *EGFR*, *FOXC1*, *CDH3*, *KRT5*, *KRT14*, *KRT17*, *MIA*, and *PHGDH*; and proliferation-related genes: *UBE2T*, *EXO1*, *BIRC5*, *KIF2C*, *CDC20*, *CEP55*, *MELK*, *ORC6L*, *UBE2C*, *CCNE1*, *CENPF*, *CDCA1*, *RRM2*, *ANLN*, *MYBL2*, *MKI67*, *PTTG1*, *CCNB1*, *TYMS*, *KNTC2*, and *CDC6*) in the PAM50 dataset are highly expressed in basal-like cancer.[15] The overall survival (OS) and hazard ratio (HR) of the cohort with increased expression (z-score ≥1) of the 30 genes were calculated. Increased expression of the 30 genes was linked to shorter OS (HR = 1.13, 95% confidence interval (CI): 1.01–1.26) (Figure 2A). Among the 30 genes, increased expression of three genes (*CDCA1*, *RRM2*, and *CCNB1*) was linked to shorter OS (Figure 2A). The effect of the expression of 50 genes on survival in basal-like breast cancer was also assessed. The results showed that increased expression of *EXO1* was linked to longer OS (Figure 2B).

The effect of *EXO1* expression on survival

EXO1 is a basal-like breast cancer-associated gene, suggesting that it is linked to poor clinical outcome.[15] Studies show that overexpression of *EXO1* is linked to shorter survival in breast cancer.[2] However, the results of the present unbiased bioinformatics approach suggested that increased expression of *EXO1* is linked to longer survival in basal-like breast cancer, as determined using a Cox proportional hazard model (Figure 2B). To confirm these results, we used a different statistical model and multivariate analysis to control for other factors such as age and stage. Although *EXO1* was not related to survival in the total population in the TCGA dataset (HR = 0.870, 95% CI: 0.538–1.637), increased expression of *EXO1* was linked to longer OS in basal-like breast cancer in the log-rank test (HR = 0.135, 95% CI: 0.002–0.120) (Figure 3A and B). These data remained significant in the multivariate analysis after adjusting for age and stage (Table 1). The effect of *EXO1* overexpression on survival was assessed in a marginal large cohort of 1402 cases of breast cancer including 241 cases of basal-like breast cancer in the KMplot dataset. Consistent with previous reports, increased expression of *EXO1* was linked to shorter OS in all breast cancer cases (HR = 1.348, 95% CI: 1.069–1.789) (Figure 3C). The raw clinical and gene expression data were not available in the KMplot dataset. Therefore, the original settings regarding the definition of basal-like breast cancer (ESR1-negative and HER2-negative) and overexpression (upper quartile of the entire population) could not be modified. However, even under different conditions, increased expression of *EXO1* was linked to longer OS in the large cohort of basal-like breast cancer (Figure 3D).

Discussion

In the present study, second-round classification was performed using specific genes in the PAM50 dataset. This analysis further subdivided basal-like breast cancer into two groups based on *EXO1* expression, and the two groups were linked to different clinical outcomes. *EXO1* is an independent poor prognostic factor in breast cancer; however, the present results showed that *EXO1* overexpression was linked to longer survival in basal-like breast cancer.[10,11] Thirty genes in the PAM50 dataset are overexpressed in basal-like breast cancer including *EXO1*, indicating that increased expression of *EXO1* may identify a subset of basal-like breast cancer. As basal-like breast cancer is associated with poor clinical outcome, increased expression of *EXO1* could indicate a poor clinical outcome in the total cohort.[2]

In the present study, the effect of single molecule expression on survival was investigated to identify a system of breast cancer classification beyond the current intrinsic molecular subtype classification. The present results suggest that effective therapies can be identified using precision oncology. Increased expression of *EXO1* was associated with longer survival in basal-like breast cancer, suggesting that decreased expression of *EXO1* is linked to poor prognosis. *EXO1* is an exonuclease involved in DNA end resection at the sites of DNA damage, an initial step of homologous recombination, and the repair machinery for DNA double strand breaks.[16] *EXO1*$^{-/-}$ *S. pombe* displays intact RAD51 foci formation, suggesting that *EXO1* is dispensable for homologous recombination.[17] These data indicate that DNA damaging agents or PARP inhibitors, which specifically kill cells with homologous recombination defects, are not appropriate for the poor prognostic basal-like breast cancer with *EXO1* downregulation. However, a different approach could provide a promising therapeutic option. One group compared gene expression patterns based on *EXO1* expression in breast cancer, and the differentially expressed genes in the *EXO1* overexpressing cohort were further validated using anticancer agents that efficiently suppressed cellular growth.[10] This analysis revealed that PI3K/AKT inhibition is a potential therapeutic option for breast cancer with *EXO1* overexpression. Such procedures can be used to identify novel therapies for the treatment of breast cancer with *EXO1* downregulation.

The present study proposed a novel subdivision of basal-like breast cancer. However, the study was performed using a limited number of genes; a greater number of genes may improve the accuracy of the subdivision. In addition, the

DOI: 10.23958/ijirms/vol02-i07/15

application of this system to clinical practice requires simplified methods, such as immunohistochemistry techniques. Further studies using larger cohorts are required before the clinical application of this method.

Conclusion

Basal-like breast cancer can be further divided into two groups with different prognosis based on *EXO1* expression.

Acknowledgements

This study was supported by grants from Japan Society for the Promotion of Science.

References

[1] Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA et al. Molecular portraits of human breast tumours. Nature 2000; 406:747-752. doi:10.1038/35021093

[2] Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001; 98:10869-10874. doi:10.1073/pnas.191367098 98/19/10869 [pii]

[3] van't Veer LJ, Paik S, Hayes DF Gene expression profiling of breast cancer: a new tumor marker. J Clin Oncol 2005; 23:1631-1635. doi:10.1200/JCO.2005.12.005

[4] van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002; 415:530-536. doi:10.1038/415530a

[5] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. The New England journal of medicine 2004; 351:2817-2826. doi:10.1056/NEJMoa041588

[6] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009; 27:1160-1167. doi:10.1200/JCO.2008.18.1370

[7] Barbieri CE, Chinnaiyan AM, Lerner SP, Swanton C, Rubin MA the Emergence of Precision Urologic Oncology: A Collaborative Review on Biomarker-driven Therapeutics. Eur Urol 2017; 71:237-246. doi:10.1016/j.eururo.2016.08.024

[8] Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 1999; 17:2639-2648. doi:10.1200/JCO.1999.17.9.2639

[9] Dawood S, Broglio K, Buzdar AU, Hortobagyi GN, Giordano SH Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: an institutional-based review. J Clin Oncol 2010; 28:92-98. doi:10.1200/JCO.2008.19.9844

[10] Muthuswami M, Ramesh V, Banerjee S, Viveka Thangaraj S, Periasamy J, Bhaskar Rao D et al. Breast tumors with elevated expression of 1q candidate genes confer poor clinical outcome and sensitivity to Ras/PI3K inhibition. PLoS One 2013; 8:e77553. doi:10.1371/journal.pone.0077553

[11] Vequaud E, Desplanques G, Jezequel P, Juin P, Barille-Nion S Survivin contributes to DNA repair by homologous recombination in breast cancer cells. Breast cancer research and treatment 2016; 155:53-63. doi:10.1007/s10549-015-3657-z

[12] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer discovery 2012; 2:401-404. doi:10.1158/2159-8290.CD-12-0095

[13] Szasz AM, Lanczky A, Nagy A, Forster S, Hark K, Green JE et al. Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. Oncotarget 2016; 7:49322-49333. doi:10.18632/oncotarget.10337

[14] Cancer Genome Atlas N Comprehensive molecular portraits of human breast tumours. Nature 2012; 490:61-70. doi:10.1038/nature11412

[15] Prat A, Pineda E, Adamo B, Galvan P, Fernandez A, Gaba L et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. Breast 2015; 24 Suppl 2:S26-35. doi:10.1016/j.breast.2015.07.008

[16] Zhang F, Shi J, Chen SH, Bian C, Yu X the PIN domain of *EXO1* recognizes poly (ADP-ribose) in DNA damage response. Nucleic acids research 2015; 43:10782-10794. doi:10.1093/nar/gkv939

[17] Tomita K, Matsuura A, Caspari T, Carr AM, Akamatsu Y, Iwasaki H et al. Competition between the Rad50 complex and the Ku heterodimer reveals a role for *Exo1* in processing double-strand breaks but not telomeres. Mol Cell Biol 2003; 23:5186-5197.

DOI: 10.23958/ijirms/vol02-i07/15

**** Corresponding Authors:***

Ko Sato,
Department of Pharmacogenomics, St. Marianna University,
2-16-1 Sugao, Miyamae, Kawasaki, 2168511, Japan
Telephone: +81-44-9778111
Fax: +81-44-9774413

Anna S Sedukhina,
Department of Pharmacogenomics, St. Marianna University,
2-16-1 Sugao, Miyamae, Kawasaki, 2168511, Japan
Telephone: +81-44-9778111
Fax: +81-44-9774413

DOI: 10.23958/ijirms/vol02-i07/15