Original Article

J Korean Soc Radiol 2020;81(3):632-643
https://doi.org/10.3348/jksr.2020.81.3.632
pISSN 1738-2637 / eISSN 2288-2928

대한영상의학회지

Correlation between MR Image-Based Radiomics Features and Risk Scores Associated with Gene Expression Profiles in Breast Cancer

유방암에서 자기공명영상 근거 영상표현형과 유전자 발현 프로파일 근거 위험도의 관계

Ga Ram Kim, MD[1], You Jin Ku, MD[2], Jun Ho Kim, MD[1], Eun-Kyung Kim, MD[3]*

[1]Department of Radiology, Inha University Hospital, Inha University School of Medicine, Incheon, Korea
[2]Department of Radiology, International St. Mary's Hospital, Catholic Kwandong University, Incheon, Korea
[3]Department of Radiology, Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

Purpose To investigate the correlation between magnetic resonance (MR) image-based radiomics features and the genomic features of breast cancer by focusing on biomolecular intrinsic subtypes and gene expression profiles based on risk scores.

Materials and Methods We used the publicly available datasets from the Cancer Genome Atlas and the Cancer Imaging Archive to extract the radiomics features of 122 breast cancers on MR images. Furthermore, PAM50 intrinsic subtypes were classified and their risk scores were determined from gene expression profiles. The relationship between radiomics features and biomolecular characteristics was analyzed. A penalized generalized regression analysis was performed to build prediction models.

Results The PAM50 subtype demonstrated a statistically significant association with the maximum 2D diameter (p = 0.0189), degree of correlation (p = 0.0386), and inverse difference moment normalized (p = 0.0337). Among risk score systems, GGI and GENE70 shared 8 correlated radiomic features (p = 0.0008–0.0492) that were statistically significant. Although the maximum 2D diameter was most significantly correlated to both score systems (p = 0.0139, and p = 0.0008), the overall degree of correlation of the prediction models was weak with the highest correlation coefficient of GENE70 being 0.2171.

Conclusion Maximum 2D diameter, degree of correlation, and inverse difference moment nor-

Received February 12, 2019
Revised June 27, 2019
Accepted September 14, 2019

***Corresponding author**
Eun-Kyung Kim, MD
Department of Radiology, Severance Hospital, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.

Tel 82-2-2228-7400
Fax 82-2-393-3035
E-mail ekkim@yuhs.ac

ORCID iDs
Ga Ram Kim
https://orcid.org/0000-0002-4481-5792
You Jin Ku
https://orcid.org/0000-0002-6727-0289
Jun Ho Kim
https://orcid.org/0000-0002-7913-7506
Eun-Kyung Kim
https://orcid.org/0000-0002-3368-5013

malized demonstrated significant relationships with the PAM50 intrinsic subtypes along with gene expression profile-based risk scores such as GENE70, despite weak correlations.

Index terms Breast Neoplasms; Magnetic Resonance Imaging; Gene Expression Profiling

INTRODUCTION

Gene expression profiling by high-throughput technologies has provided deeper insight into the complex biomolecular nature of breast cancer (1-7). Some investigators have discovered gene expression profiles that can be used to classify intrinsic subtypes to predict prognosis and treatment response and this knowledge has helped to individualize treatment strategies for breast cancer patients (8). However, gene expression profiling is not yet readily applicable in daily practice.

Recent advances in the computer-aided quantitative analysis of radiologic images (so called radiomics) have enabled us to go beyond the detection and diagnosis of cancer to image-based cancer phenotyping. With radiomics, we cannot only correlate images to pathologic tumor or node stage, nuclear grade and molecular subtype, but also gather further information on prognosis and treatment response. This can be done by converting images to high-throughput quantitative data and subsequently analyzing the statistical relationships between radiomics features and clinicopathologic factors. Radiogenomics refers to the study of mathematical relationships between radiomics features and genomic features (9); the Cancer Genome Atlas (TCGA) of the National Cancer Institute (10) and its imaging counterpart, the Cancer Imaging Archive (TCIA) (11) facilitate cross-disciplinary research to find relationships between imaging phenotypes and genomic subtypes.

PAM50 is a well-known gene assay for breast cancer; it was developed as a 50-gene quantitative real time polymerase chain reaction assay that identifies and categorizes intrinsic molecular subtypes of breast cancer into the luminal A, luminal B, human epidermal growth factor receptor 2 (HER2)-enriched, basal-like, and normal-like phenotypes from RNA isolated from formalin-fixed, paraffin-embedded tissue. The PAM50 assay was also used to develop a prognostic score for risk of relapse based on the relative distance to the centroid of each subtype; a proliferation score based on a gene subset related to cell cycle progression; and composite scores that include tumor size with molecular phenotypes (1). The PAM50-based risk score was found to be significant in tumors less than 5 cm in size that were estrogen receptor (ER)-positive, HER2-negative and lymph node-negative (12-15).

If radiomics features, such as those of MRI which is widely used in the preoperative evaluation of breast cancer, can predict the genomic features of breast cancer, we could readily acquire information that can be used to tailor treatment for individuals even within routine clinical practice. Therefore, the purpose of this study is to investigate the relationship between MR image-based radiomics features and genomic features of breast cancer by focusing on biomolecular intrinsic subtypes and gene expression profiles based on risk scores.

MATERIALS AND METHODS

DATA DOWNLOAD

We had assess to only de-identified data and the approval of the Institutional Review Board was unnecessary. Clinical and genomic data for the patients were downloaded from TCGA from the Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov) along with the simultaneous MR images from TCIA (https://www.cancerimagingarchive.net/). After matching, finally 122 patients with simultaneous gene expression data and appropriate MR images were enrolled in this study. Among the included MR images, 91 cases were obtained with a GE 1.5 T MRI scanner (GE Medical Systems, Milwaukee, WI, USA), 13 with a Siemens 1.5 T MRI scanner (Siemens, Berlin, Germany), and 15 with a Phillips 1.5 T MRI scanner (Philips Medical Systems, Bothell, WA, USA). Three cases were acquired on a Phillips 3.0 T scanner (Philips Medical Systems). All MR images acquired with GE scanners were obtained using a standard double breast coil and a gadolinium-based contrast agent.

MR IMAGE REVIEW, FEATURE EXTRACTION AND INTEROBSERVER AGREEMENT

The obtained MR images were reviewed independently by two breast radiologists with more than 8 years of experience in breast imaging. Each radiologist reviewed all 122 cases and determined a representative slice for every patient. In case of discordant findings, the two radiologists discussed and reached a consensus. Region of interests (ROIs) were then drawn in a semiautomatic manner using MIPAV (https://mipav.cit.nih.gov). Radiomic features were extracted using pyradiomics (https://github.com/Radiomics/pyradiomics). A total of 100 features of 7 categories were extracted from each ROI. The 7 categories were first order statistics (18 features), shape-based features (8 features), gray-level co-occurrence matrix (23 features), gray-level run-length matrix (16 features), gray-level size zone matrix (16 features), neighboring gray-tone difference matrix (5 features), and gray-level dependence matrix (14 features).

Interobserver agreement for the radiomic features extracted from the ROIs was evaluated with the intraclass correlation coefficient (ICC) and 95% confidence interval (CI). The 'irr' R package was used for ICC analysis (R version 3.5.1, http://www.R-project.org).

DATA AND STATISTICAL ANALYSES

The purpose of this study was twofold. The first purpose was to investigate radiomic characteristics of each biomolecular subtype based on genomic characteristics and the second purpose was to determine the correlation between radiomics features extracted from MRI and gene expression profile-based recurrence (or prognosis) risk score systems. To define radiomic characteristics, we used 100 radiomics features extracted from MRI using the pyradiomics package. The Kruskal-Wallis test was performed to identify differences in radiomics features among individual biomolecular subtypes.

To investigate the correlation between radiomic features and gene expression profile-based recurrence (or prognosis) risk score systems, we first performed Spearman's correlation test for individual radiomics features and 11 risk score systems. Then, we identified statistically

significant features, and used these features to establish prediction models using penalized generalized regression with the least absolute shrinkage and selection operator (LASSO). Radiomics features were selected with LASSO using the 'glmnet' R package to study the correlation between risk scores determined with gene expression profiles and mathematical models built on radiomics features. Cross validation was performed in a leave-one-out manner.

Genomic analyses were performed using 1092 gene expression profiles by RNA sequencing. The normalized RSEM data were downloaded using 'TCGABiolinks' package in R. The PAM50 intrinsic subtypes were classified as described in a previous report (6) and risk scores were determined using the 'genefu' package in R. The risk score systems included single gene based prognosis prediction (ESR1, ERBB2, and AURKA) (16), EndoPredict (17), GENIUS (2), GGI (6), OncotypeDx (8), TamR (4), GENE70 (7), PIK3CA gene signature (5), and ROR-S (1). The analyses were run on a set of 1092 cases and the results of the 122 enrolled patients were used for further analysis. The Kruskal-Wallis test was performed to analyze the relationship between intrinsic subtypes and radiomics features.

All statistical analyses were performed with R version 3.5.1. R packages including 'TCGABiolinks', 'genefu', 'glmnet', and 'irr' were used to extract TCGA data, calculate the risk score of the intrinsic subtype, and perform penalized generalized regression with the LASSO, and ICC analysis, respectively.

RESULTS

PATIENT CHARACTERISTICS

From TCGA, 122 patients with normalized RNA sequencing data available for gene expression and simultaneous TCIA MR images were enrolled. The median age of the patients was 55 years (range from 29 to 83 years). Infiltrating ductal carcinoma was the most frequent pathologic type of breast cancer observed in this study population, and followed by lobular carcinoma. The luminal A type was the most dominant molecular subtype (Table 1).

INTEROBSERVER AGREEMENT BETWEEN THE TWO RADIOLOGISTS

The interobserver agreement for feature extraction between the two radiologists was acceptable (ICC 95% CI, 0.768–1.000). The agreement was the highest for gray-level size zone matrix features and the lowest for first order features (Table 2).

RELATIONSHIP BETWEEN RADIOMICS FEATURES AND GENE EXPRESSION PROFILE-BASED FEATURES

The PAM50 subtype was significantly correlated to three radiomic features, which were the maximum 2D diameter (p = 0.0189), correlation (p = 0.0386) and inverse difference moment normalized (p = 0.0337). In univariate analysis, overall shape features seemed to be more related to risk scores than texture features contrary to the intrinsic subtype. GGI and GENE70 showed significantly more related radiomics features than the other risk score systems. ERBB2, GENIUS, and PIK3CA were not significantly related with radiomic features (Table 3). Among the risk score systems, GGI was significantly correlated to 2 shape features [elongation (p = 0.0199), and max 2D diameter column (p = 0.0139)], 2 gray-level dependence matrix

Table 1. Patient Characteristics

Characteristics	Number
Pathology	
Infiltrating duct and lobular carcinoma	1
Infiltrating duct carcinoma, not otherwise specified	102
Infiltrating duct mixed with other types of carcinoma	1
Lobular carcinoma, not otherwise specified	16
Medullary carcinoma, not otherwise specified	1
Pleomorphic carcinoma	1
Stage	
Stage I	22
Stage Ia	8
Stage II	1
Stage IIa	50
Stage IIb	24
Stage IIIa	10
Stage IIIc	7
Age, years	
Median (range)	55 (29–83)
Race	
Asian	1
African American	20
Caucasian	101
PAM50 subtype	
Basal	17
HER2	8
Luminal A	79
Luminal B	18

HER2 = human epidermal growth factor receptor 2

Table 2. Interobserver Agreement

Feature Class	Intraclass Correlation Coefficient, 95% Confidence Interval
Shape	0.771–1.000
Gray-level dependence matrix	0.799–1.000
Gray-level co-occurrence matrix	0.817–1.000
First order statistics	0.768–1.000
Gray-level run-length matrix	0.787–1.000
Gray-level size-zone matrix	0.820–1.000
Neighboring gray-tone difference matrix	0.818–1.000

features [small dependence low gray-level emphasis (p = 0.0261), and low gray-level emphasis (p = 0.331)], 2 first order features [total energy (p = 0.0412), and 10 percentile (p = 0.0214)], 2 gray-level run-length matrix features [short-run low gray-level emphasis (p = 0.0244), and low gray-level run emphasis (p = 0.0320)], and a gray-level size-zone matrix feature–[(small-area

Table 3. Number of Radiomic Features that Demonstrated a Statistically Significant ($p < 0.05$) Association with Intrinsic Subtypes or Risk Scores

Feature Class	Shape	Gray-Level Dependence Matrix	Gray-Level Co-occurrence Matrix	First Order Statistics	Gray-Level Run-Length Matrix	Gray-Level Size-Zone Matrix	Neighboring Gray-Tone Difference Matrix
Intrinsic Subtype							
PAM50	1	0	2	0	0	0	0
CNV*	0	0	0	1	0	0	0
Mutation*	0	3	6	7	3	2	1
DNA methylation*	0	2	1	4	3	2	0
mRNA*	1	0	0	0	0	0	0
miRNA*	4	2	10	8	1	7	4
lncRNA*	0	4	0	0	0	5	0
Protein*	1	0	0	0	0	0	0
PARADIGM*	2	0	0	0	0	0	0
Risk Score†							
AURKA	2	1	0	0	0	0	3
ESR1	1	0	0	0	0	0	0
ERBB2	0	0	0	0	0	0	0
GGI	2	2	0	2	2	1	0
GENIUS	0	0	0	0	0	0	0
EndoPredict	1	0	0	0	0	0	0
OncotypeDx	3	0	0	0	0	0	0
TamR	1	1	0	1	0	0	0
GENE70	3	3	0	3	2	1	0
PIK3CA	0	0	0	0	0	0	0
ROR-S	2	0	0	0	0	0	0

*Clustering results from a previous study (9).
†'For research' and 'NOT for clinical' scores determined based on gene expression profiles.

Table 4. Penalized Generalized Regression

Risk Scores*	Adjusted R^2	*p*-Value
AURKA	0.1998	< 0.001
ESR1	0.152	< 0.001
ERBB2	-	not significant
GGI	0.1835	< 0.001
GENIUS	-	not significant
EndoPredict	0.1118	0.00693
OncotypeDx	0.1474	0.00167
TamR	0.1991	< 0.001
GENE70	0.2171	< 0.001
PIK3CA	-	not significant
ROR-S	0.1903	< 0.001

*'For research' and 'NOT for clinical' scores determined based on gene expression profiles.

low gray-level emphasis (p = 0.0491)]. GENE70 score was significantly correlated to 3 shape features [elongation (p = 0.0122), max 2D diameter column (p = 0.0008), and surface area (p = 0.0377)], a gray-level dependence matrix feature [small dependence low gray-level emphasis (p = 0.0169)], 3 first order features [total energy (p = 0.0251), mean (p = 0.0492), and 10 percentile (p = 0.0101)], 2 gray-level run-length matrix features [short run low gray-level emphasis (p = 0.0350), and low gray-level run emphasis (p = 0.0478)], and a gray-level size-zone matrix feature [small area low gray-level emphasis (p = 0.0392)].

On the basis of significantly correlated features, prediction models were established. Most risk score prediction modes showed statistically significant p values except ERBB2, GENIUS, and PIK3CA (Table 4). However, overall correlation was weak as the adjusted R^2 values were low (below 0.3) with the adjusted R^2 value of GENE70 being the highest at 0.2171.

UNSUPERVISED HIERARCHICAL CLUSTERING

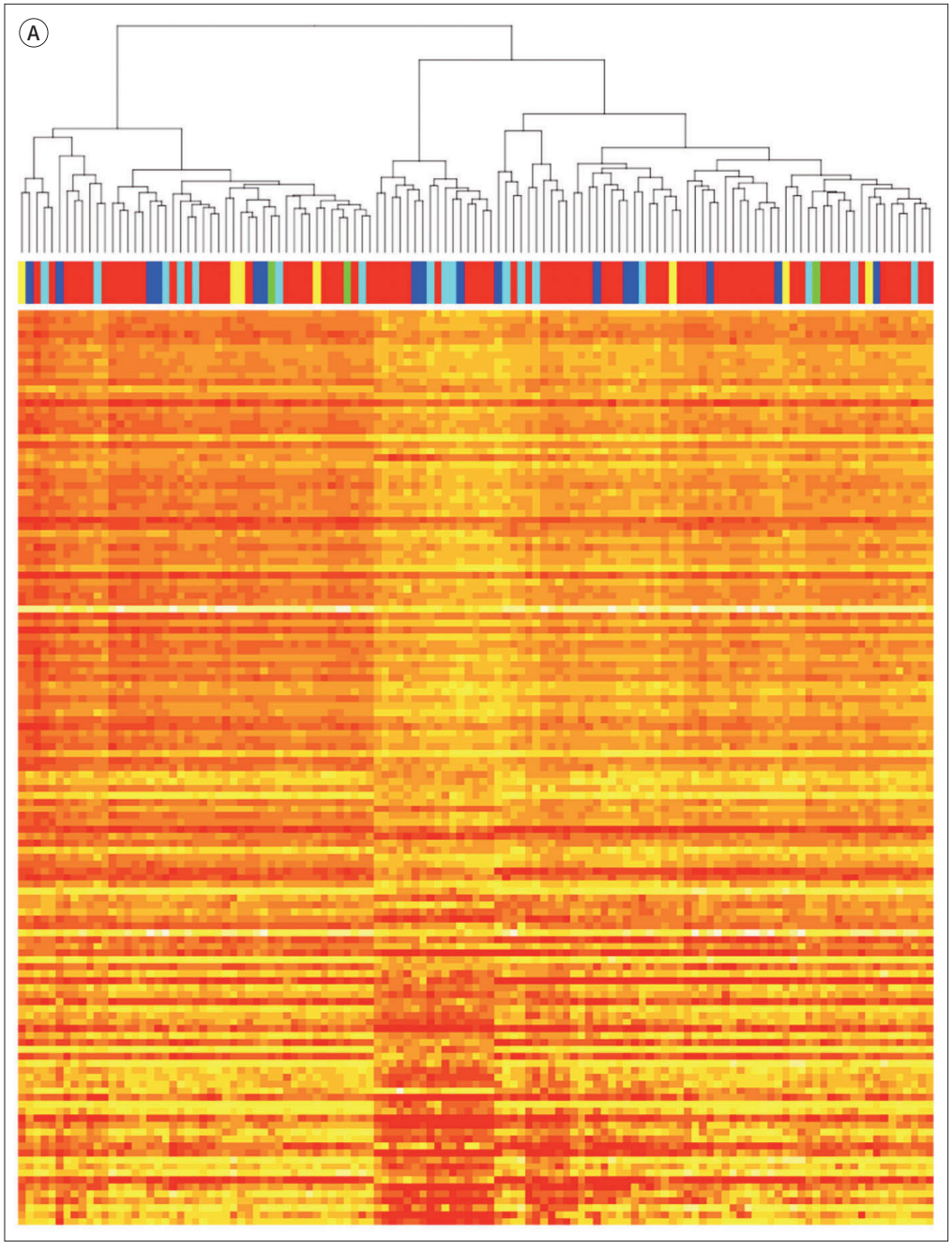
Differentially expressed genes according to PAM50 classification were extracted using the Kruskal-Wallis test. A p value less than 10^{-9} was considered statistically significant and 133 genes whose expression was significantly different among the PAM50 classifications were selected. An unsupervised hierarchical clustering analysis did not correlate well with PAM50 classification (Fig. 1A). Differentially extracted radiomic features according to PAM50 classification were selected using the Kruskal-Wallis test. Only 4 features with p values less than 10^{-1} were selected. An unsupervised hierarchical clustering failed to show significant correlation with PAM50 classification (Fig. 1B).

DISCUSSION

With recent advances in computational biology, gene expression profiles allow more useful information to be collected regarding prognosis than conventional clinicopathological studies. Especially for breast cancer, there are risk score systems based on multi-gene expression profiles that provide more information to predict recurrence and treatment response than traditional clinical and histopathological factors (1, 2, 4-8, 16, 17). Based on these risk scores, treatment strategies can be tailored to each individual patient. Studying the relationships between gene expression profiles and image phenotypes may provide valuable opportunities to develop robust tools for tailored treatment. Eventually, we will be able to obtain information regarding intrinsic subtypes and risk scores based on biomolecular characteristics in an automatized manner with software embedded in imaging machines.

In our study, the interobserver agreement between the two radiologists for feature-computerized extraction by drawing the ROIs of 122 MR lesions was comparably high (ICC 95% CI, 0.768–1.000). Qualitative assessments made by humans will naturally lead to interobserver variations. The interobserver variability of three radiologists for 294 breast MR lesions was substantial for mass internal enhancement (k = 0.62) and moderate for peritumoral edema (k = 0.46) with the k agreement in a past study (18). On the other hand, the interobserver reproducibility of two radiologists for computerized extraction of texture features by drawing the ROIs of 50 breast ultrasound (US) lesions was said to be high in another study, with a somewhat lower ICC than ours (ICC 95%, 0.691–1.000) (19). We could increase interobserver agree-

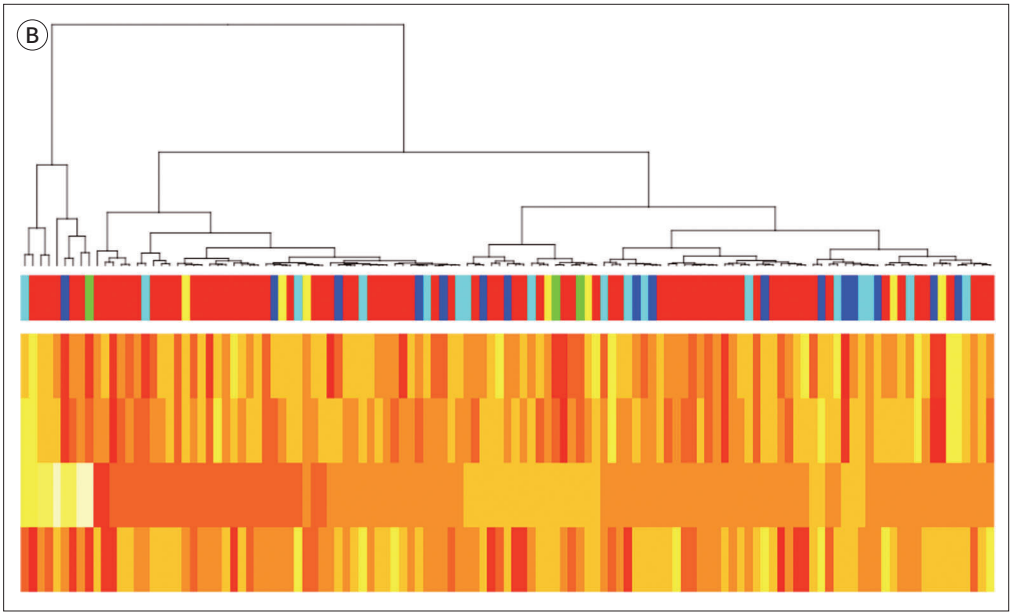
Fig. 1. Unsupervised hierarchical clustering analysis of the enrolled cases with differently expressed genes (A) and differently extracted radiomic features (B). In both the (Top) dendrograms, (Mid) the color bars indicate PAM50 classification (red: luminal A, cyan: luminal B, yellow: human epidermal growth factor receptor 2, green: normograde, and blue: basal), and (Bottom) the heatmaps of gene expression (A), and radiomic features (B).



ment with semi-automatized techniques to draw ROIs. As interobserver variation originates from human judgement, we can expect automatized segmentation of tumors to eliminate this variation in the future.

We found the PAM50 intrinsic subtype to be significantly related to shape and texture (gray-level co-occurrence matrix) features. The significant shape feature was the "maximum

Fig. 1. Unsupervised hierarchical clustering analysis of the enrolled cases with differently expressed genes (A) and differently extracted radiomic features (B). In both the (Top) dendrograms, (Mid) the color bars indicate PAM50 classification (red: luminal A, cyan: luminal B, yellow: human epidermal growth factor receptor 2, green: normograde, and blue: basal), and (Bottom) the heatmaps of gene expression (A), and radiomic features (B).



2D diameter" which reflected the size of the ROI. Previous studies reported ER negative (ER-) and triple negative (TN) subtypes to be related to larger tumors (20, 21). These subtypes are known to have higher microvessel density as well as higher proliferation activity (22-24). The significant texture features were 'correlation' and 'inverse difference moment normalized', which reflects texture heterogeneity. In a previous report by Waugh et al. (25), texture heterogeneity was significantly increased in HER2-enriched and TN subtypes. We also found that the number of radiomics features showed significant correlation with risk scores based on gene expression profiles. Especially, heterogeneity texture features were consistently related to risk scores with statistical significance. These radiomics features quantitatively measure the heterogeneous nature of enhancement within the ROI. Breast cancer has heterogeneous genomic characteristics with multiple driver mutations, the degree of which are known to be related to treatment resistance and poor prognosis (26). Thus, a non-invasive quantitative measurement of heterogeneity may be useful for determining optimal treatment strategies.

Among the risk score systems analyzed, those with relatively fewer signature genes, tended to have none or few significantly related radiomics features. This finding indicates that radiomics features may not reflect a single gene or individual signaling pathway, but rather overall patterns of gene expression. Zhu et al. (27) made the same speculation after observing that radiomics features were not correlated to mutations or copy number profiles in their study. Although the number of radiomics features was significantly correlated to risk scores, generalized regression analyses failed to build strong prediction models. The correlation coefficient of the GENE70 model was the highest at 0.2171. This indicates that radiomics features cannot be used to predict prognostic risk for clinical use at this time and that further study is needed to develop mathematical models to predict biomolecular risk scores.

This study has some limitations. First, the study enrolled a relatively small number of patients because they were collected from a limited source of data sets (TCGA and TCIA). Another limitation was the uneven quality of MR images. Most of the archives images were obtained on outdated machines without standardized protocols. Also, there might have been variability arising when radiomics features were extracted because the two radiologists drew the ROIs and MR images were obtained with machines manufactured by three different companies. Lastly, some of the risk scores were calculated for research purposes with an algorithm-based method and these calculation methods were different from the original methods for risk scores, which were not from clinical tests. Thus, there might be discrepancies between 'research purpose' risk scores and 'clinical purpose' risk scores. Despite these limitations, the results of this study suggest that image-based biomolecular phenotypes have the potential to predict the prognosis of breast cancer.

In conclusion, the radiomics features of maximum 2D diameter, correlation and inverse difference moment normalized showed significant relationships with biomolecular characteristics, PAM50 intrinsic subtypes and gene expression profile-based risk scores such as GENE70, although the correlations were weak. Thus, further studies are necessary to develop adequate prediction models using MR image-based phenotypes.

Author Contributions

Conceptualization, K.G.R.; data curation, K.G.R.; formal analysis, K.G.R.; funding acquisition, K.G.R.; investigation, K.G.R., K.Y.J.; methodology, K.G.R.; project administration, K.E.; resources, K.G.R., K.Y.J., K.J.H.; supervision, K.E.; visualization, K.G.R.; writing—original draft, K.G.R.; and writing—review & editing, K.G.R., K.E.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Acknowledgments

This work was supported by an Inha University Hospital Research Grant.

REFERENCES

1. 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
2. Haibe-Kains B, Desmedt C, Rothé F, Piccart M, Sotiriou C, Bontempi G. A fuzzy gene expression-based computational approach improves breast cancer prognostication. *Genome Biol* 2010;11:R18
3. Kim GR, Ku YJ, Cho SG, Kim SJ, Min BS. Associations between gene expression profiles of invasive breast cancer and Breast Imaging Reporting and Data System MRI lexicon. *Ann Surg Treat Res* 2017;93:18-26
4. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics* 2008;9:239
5. Loi S, Haibe-Kains B, Majjaj S, Lallemand F, Durbecq V, Larsimont D, et al. PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. *Proc Natl Acad Sci U S A* 2010;107:10208-10213
6. 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;98:262-272
7. 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

8. 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. *N Engl J Med* 2004;351:2817-2826
9. Kim JH. Imaging informatics: a new horizon for radiology in the era of artificial interlligence, big data, and data science. *J Korean Soc Radiol* 2019;80:176-201
10. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 2012;490:61-70
11. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 2013;26:1045-1057
12. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol* 2016;34:1134-1150
13. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol* 2013;31:2783-2790
14. Caan BJ, Sweeney C, Habel LA, Kwan ML, Kroenke CH, Weltzien EK, et al. Intrinsic subtypes from the PAM50 gene expression assay in a population-based breast cancer survivor cohort: prognostication of short- and long-term outcomes. *Cancer Epidemiol Biomarkers Prev* 2014;23:725-734
15. Sestak I, Cuzick J, Dowsett M, Lopez-Knowles E, Filipits M, Dubsky P, et al. Prediction of late distant recurrence after 5 years of endocrine treatment: a combined analysis of patients from the Austrian breast and colorectal cancer study group 8 and arimidex, tamoxifen alone or in combination randomized trials using the PAM50 risk of recurrence score. *J Clin Oncol* 2015;33:916-922
16. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clin Cancer Res* 2008;14:5158-5165
17. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res* 2011;17:6012-6020
18. Song SE, Shin SU, Moon HG, Ryu HS, Kim K, Moon WK. MR imaging features associated with distant metastasis-free survival of patients with invasive breast cancer: a case-control study. *Breast Cancer Res Treat* 2017;162:559-569
19. Lee SE, Han K, Kwak JY, Lee E, Kim EK. Radiomics of US texture features in differential diagnosis between triple-negative breast cancer and fibroadenoma. *Sci Rep* 2018;8:13546
20. Agner SC, Rosen MA, Englander S, Tomaszewski JE, Feldman MD, Zhang P, et al. Computerized image analysis for identifying triple-negative breast cancers and differentiating them from other molecular subtypes of breast cancer on dynamic contrast-enhanced MR images: a feasibility study. *Radiology* 2014;272:91-99
21. Youk JH, Son EJ, Chung J, Kim JA, Kim EK. Triple-negative invasive breast cancer on dynamic contrast-enhanced and diffusion-weighted MR imaging: comparison with other breast cancer subtypes. *Eur Radiol* 2012;22:1724-1734
22. Badve S, Dabbs DJ, Schnitt SJ, Baehner FL, Decker T, Eusebi V, et al. Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. *Mod Pathol* 2011;24:157-167
23. Fuckar D, Dekanić A, Stifter S, Mustać E, Krstulja M, Dobrila F, et al. VEGF expression is associated with negative estrogen receptor status in patients with breast cancer. *Int J Surg Pathol* 2006;14:49-55
24. Koukourakis MI, Manolas C, Minopoulos G, Giatromanolaki A, Sivridis E. Angiogenesis relates to estrogen receptor negativity, c-erbB-2 overexpression and early relapse in node-negative ductal carcinoma of the breast. *Int J Surg Pathol* 2003;11:29-34
25. Waugh SA, Purdie CA, Jordan LB, Vinnicombe S, Lerski RA, Martin P, et al. Magnetic resonance imaging texture analysis classification of primary breast cancer. *Eur Radiol* 2016;26:322-330
26. Koren S, Bentires-Alj M. Breast tumor heterogeneity: source of fitness, hurdle for therapy. *Mol Cell* 2015;60:537-546
27. Zhu Y, Li H, Guo W, Drukker K, Lan L, Giger ML, et al. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. *Sci Rep* 2015;5:17787

J Korean Soc Radiol 2020;81(3):632-643

대한영상의학회지

유방암에서 자기공명영상 근거 영상표현형과 유전자 발현 프로파일 근거 위험도의 관계

김가람[1] · 구유진[2] · 김준호[1] · 김은경[3]*

목적 자기공명영상 근거 영상표현형과 생체분자학적 아형, 유전자 발현 프로파일 근거 위험도 등 유방암 유전체 특징의 관계를 분석하고자 하였다.

대상과 방법 The Cancer Genome Atlas와 and the Cancer Imaging Archive에 공개된 자료를 이용하였다. 122개의 유방암의 자기공명영상에서 영상표현형이 추출되었다. 유전자 발현 프로파일에 따라 PAM50아형을 분류하고 위험도를 지정하였다. 영상표현형과 생체분자학적 특징의 관계를 분석하였다. 예측모델을 알아보기 위해 penalized generalized regression analysis를 이용하였다.

결과 PAM50아형은 maximum 2D diameter (p = 0.0189), degree of correlation (p = 0.0386), 그리고 inverse difference moment normalized (p = 0.0337)와 유의하게 관련이 있었다. 위험도 시스템 중에 GGI와 GENE70이 통계적으로 유의하게 8개의 영상표현형 특징을 서로 공유하였다(p = 0.0008~0.0492). Maximum 2D diameter가 두 위험도 시스템에서 가장 유의하게 관련있는 특징이었으나(p = 0.0139, p = 0.0008) 예측모델의 전반적인 연관 정도는 약했고 가장 높은 연관계수는 GENE70이 0.2171이었다.

결론 영상표현형 중에 maximum 2D diameter, degree of correlation, 그리고 inverse difference moment normalized가 PAM50 아형 그리고 GENE70과 같은 유전자 발현 프로파일 근거 위험도와 그 연관도는 약하였으나 유의한 관련을 보였다.

[1]인하대학교 의과대학 부속병원 영상의학과,
[2]가톨릭관동대학교 국제성모병원 영상의학과,
[3]연세대학교 의과대학 세브란스병원 방사선의과학연구소 영상의학교실

https://doi.org/10.3348/jksr.2020.81.3.632