



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Breast cancer molecular subtype
prediction using radiomics signature on
synthetic mammography from digital
breast tomosynthesis

Jinwoo Son

Department of Medicine

The Graduate School, Yonsei University

Breast cancer molecular subtype
prediction using radiomics signature on
synthetic mammography from digital
breast tomosynthesis

Jinwoo Son

Department of Medicine

The Graduate School, Yonsei University

Breast cancer molecular subtype
prediction using radiomics signature on
synthetic mammography from digital
breast tomosynthesis

Directed by Professor Eun-Kyung Kim

Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

Jinwoo Son

June 2020

This certifies that the Doctoral
Dissertation of Jinwoo Son is approved.

Thesis Supervisor : Eun-Kyung Kim

Thesis Committee Member#1: Joon Seok Lim

Thesis Committee Member#2: Dosik Hwang

Thesis Committee Member#3: Ja Seung Koo

Thesis Committee Member#4: Hwiyoung Kim

The Graduate School
Yonsei University

June 2020

ACKNOWLEDGEMENTS

Thank you to my supervisor, Prof. Eun-Kyung Kim, for providing guidance and feedback throughout this project. I also would like to acknowledge Prof. Sungwon Kim for guiding and supporting me in the radiomics analysis. Thanks to my wife Gowoon, for always supporting me and putting up with me being sat in the hospital for hours.

<TABLE OF CONTENTS>

| | |
|---|----|
| ABSTRACT | 1 |
| I. INTRODUCTION | 3 |
| II. MATERIALS AND METHODS | 4 |
| 1. Patients selection | 4 |
| 2. Pathologic examination | 6 |
| 3. Image acquisition | 6 |
| 4. Tumor segmentation | 6 |
| 5. Radiomics feature extraction & selection | 8 |
| 6. Molecular subtype classification | 9 |
| 7. Clinical feature assessment | 9 |
| 8. Statistics analysis | 10 |
| III. RESULTS | 10 |
| 1. Radiomics features and prediction performance | 10 |
| 2. Comparison with digital mammography | 13 |
| 3. Considering the impact of margin | 14 |
| 4. Comparison of prediction performance between clinical and radiomics model | 15 |
| 5. Correlation between radiomics signature and BI-RADS features | 18 |
| IV. DISCUSSION | 19 |
| V. CONCLUSION | 24 |
| REFERENCES | 26 |
| APPENDICES | 30 |
| ABSTRACT(IN KOREAN) | 36 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1. Patients selection | 6 |
| Figure 2. Segmentation example 1 | 9 |
| Figure 3. Segmentation example 2 | 9 |
| Figure 4. Interobserver reproducibility of radiomics features .. | 13 |
| Figure 5. ROC curves of the synthetic mammography and digital mammography in the validation cohort | 16 |
| Figure 6. ROC curves of the original ROI, ROI with inner portion and the ROI with marginal portion | 17 |
| Figure 7. ROC curve of the clinical model and combined model in the validation cohort for TN vs non-TN | 19 |

LIST OF TABLES

| | |
|---|----|
| Table 1. Classification performance of radiomics models in validation cohort | 14 |
| Table 2. Comparison of AUC between radiomics models (P value) | 15 |
| Table 3. Comparison of classification performances between synthetic mammography and digital mammography in validation cohort | 15 |
| Table 4. Comparison of classification performances among ROIs | |

| | |
|---|----|
| in validation cohort | 15 |
| Table 5. Univariate and multivariate logistic regression of clinical model and combined model for TN subtype of breast cancer | 18 |
| Table 6. AUC value of clinical and combined model in validation cohort | 20 |

ABSTRACT

Breast cancer molecular subtype prediction using radiomics signature on synthetic mammography from digital breast tomosynthesis

Jinwoo Son

*Department of Medicine
The Graduate School, Yonsei University*

(Directed by Professor Eun-Kyung Kim)

Purpose: To predict molecular subtype of breast cancer using radiomics signature extracted from synthetic mammography reconstructed from digital breast tomosynthesis (DBT).

Materials and Methods: From December 2015 to September 2016, 365 patients with pathologically confirmed invasive breast cancer with preoperative DBT were included. Among the 294 patients from December 2015 to July 2016, 150 patients were selected [50 consecutive patients for each molecular subtype (luminal A+B, luminal; HER2-positive, HER2; triple-negative, TN)], and assigned to the training set. A temporally independent validation cohort consisted of 71 patients with breast cancer between August 2016 and September 2016 (50 luminal, 9 HER2, and 12 TN). Total of 129 radiomics features was extracted from the craniocaudal (CC) and mediolateral oblique (MLO) view of the synthetic and digital mammography. In addition, the radiomics features were separately obtained from the marginal and inner portions of the lesion. The performances of binary radiomics classifications for each subtype were measured using the area under the receiver operating characteristic curve (AUC). The radiomics model

was built using the elastic-net with ten-fold cross-validation and validated in the independent validation cohort. The clinical model included patient's age and BI-RADS based image features.

Results: In the validation cohort, the radiomics models yielded an AUC of 0.838 for TN, 0.556 for HER2, and 0.645 for luminal subtypes. With the optimal cut-off value of radiomics signature, sensitivity, and specificity of the models in the validation cohort were 83.3% and 79.7% for TN, 11.1% and 79.0% for HER2, 44.0% and 66.7% for luminal subtypes, respectively. There were no significant differences between synthetic mammography and digital mammography in predicting the three subtypes ($p = 0.812$ for TN, 0.268 for HER2 and 0.833 for luminal). There were no significant differences when comparing the original ROI with the ROIs with only inner and marginal portions ($p = 0.084$ and 0.051). In multivariate analysis of radiomics signature and clinical features in the classification task of TN versus non-TN, radiomics signature was the only independent predictor for predicting TN subtype. In addition, the combination of radiomics signature and clinical features showed significantly higher AUC value than the clinical features only in distinguishing TN subtype ($p = 0.045$).

Conclusion: The radiomics signature derived from the synthetic mammography from DBT showed high performance in distinguishing between TN and non-TN breast cancer. The radiomics signature from the synthetic 2D mammography of the DBT may serve as a biomarker to distinguish TN subtype of breast cancer and may affect the direction of treatment.

Key words: breast cancer, molecular subtype, radiomics, digital breast tomosynthesis, synthetic mammography

Breast cancer molecular subtype prediction using radiomics signature on synthetic mammography from digital breast tomosynthesis

Jinwoo Son

*Department of Medicine
The Graduate School, Yonsei University*

(Directed by Professor Eun-Kyung Kim)

I. INTRODUCTION

Breast cancer is the most common cancer diagnosed in women, and the second leading cause of cancer related deaths¹. Early diagnosis of breast cancer and prediction of prognosis are the key goals of clinical practice and research.

Depending on the expression level of the receptor on cancer cells, breast cancer can be divided into various subtypes, such as are luminal, human epidermal growth factor receptor 2 (HER2)-enriched, and triple-negative (TN) subtypes^{2,3}. Among these, the TN subtype is more aggressive and difficult to treat^{2,4,5}. Therefore, it accounts for a large portion of breast cancer deaths after diagnosis⁶. Patients with TN breast cancer derive no benefit from endocrine therapy or trastuzumab, because they lack the appropriate targets for these drugs. On the other hand, TN breast cancer is known to respond well to neoadjuvant chemotherapy, and it is known to improve prognosis in patients with good response⁷⁻⁹.

Several reports have found that image findings on mammography, ultrasonography, or MRI are related to the molecular subtype of breast cancer¹⁰⁻¹². Recently, several attempts have been made to predict molecular subtypes of breast cancer by radiomics approach. Radiomics refers to the transformation of image data into computer-based, high-dimensional data. These data reflects not only the characteristics of the tissue but is also

associated with gene expression¹³. A few studies have shown that radiomics features obtained from magnetic resonance imaging (MRI) are associated with molecular subtype of breast cancer¹⁴⁻¹⁶.

The use of digital breast tomosynthesis (DBT) has been increased, and several studies have reported that using DBT addition to digital mammography can increase the detection rate in breast cancer screening rather than digital mammography alone^{17,18}. The limitation of using digital mammography plus DBT for screening was the increase in the radiation dose¹⁹. To overcome this, a method was developed for reconstructing synthetic mammography images from information acquired during a DBT data acquisition. There are increasing evidences that synthetic mammography can replace digital mammography^{20,21}. However, to our knowledge, there is no research of radiomics on synthetic mammography from DBT for molecular subtyping.

The purpose of this study was to investigate whether radiomics features obtained from synthetic mammography image reconstructed from DBT could distinguish molecular subtype of breast cancer.

II. MATERIALS AND METHODS

1. Patients selection

This retrospective study was approved by the institutional review board at our center. The need for informed consent was waived. From December 2015 to September 2016, 691 patients who were diagnosed with invasive breast cancer who underwent preoperative DBT were included. Exclusion criteria were: (1) patients who received chemotherapy before underwent DBT (n = 114), (2) patients who received surgical excision or vacuum assisted biopsy (n = 41), (3) asymmetries that only visible on single view (n = 40), (4) diffuse infiltrative lesion involving whole breast (n=10), (5) lesions that partially masked by a marker (n = 15), (6) lesions were not fully included in the synthetic mammography (n = 34), and (7) lesions not clearly delineated

on both synthetic and digital mammography (n = 75).

Finally, 365 patients were included in the study. Because there are remarkable differences in incidence among molecular subtypes²², the same number of patients was assigned to each group to avoid inappropriate feature selection due to class imbalance and to improve the performance of classification^{23,24}. Among the 294 patients from December 2015 to July 2016, 50 consecutive patients were selected for each molecular subtype and assigned to the training set. Accordingly, a total of 150 patients were finally included in the training set. For the validation cohort, temporally independent 71 patients from August 2016 to September 2016 were assigned. Validation cohort consisted of 50 luminal subtype, 9 HER2 subtype and 12 TN subtype (Figure 1).

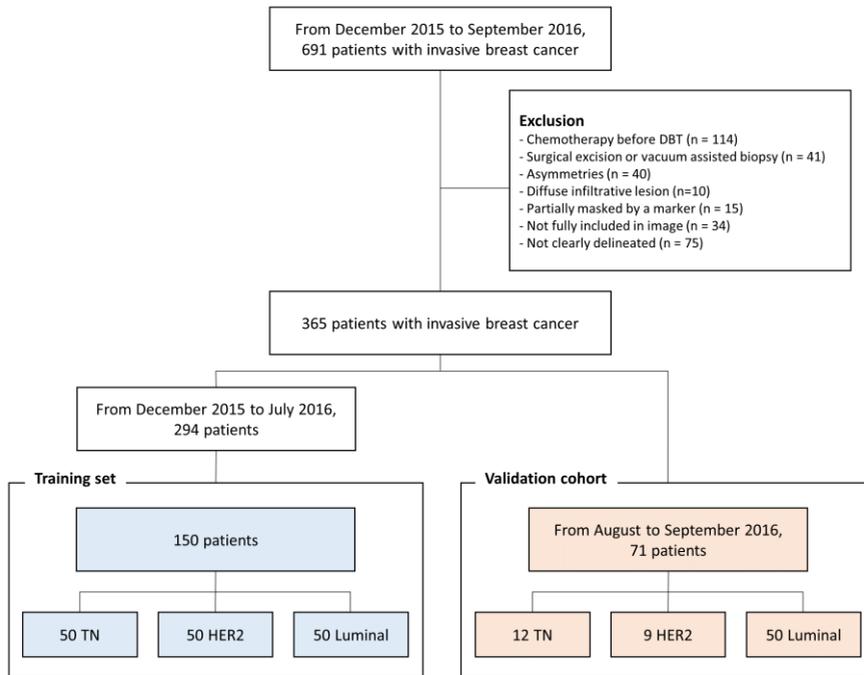


Figure 1. Patients selection.

2. Pathologic examination

A pathologic report of all breast cancers included the expression level of estrogen receptor (ER), progesterone receptor (PR) and HER2. The breast cancers were classified as “Luminal”, “HER2 (HER2-enriched)” or “TN (triple negative)” according to the ASCO/CAP guidelines²⁵.

3. Image acquisition

DBT was performed using a mammography machine (Selenia Dimensions System; Hologic, Belford, Conn) with bilateral craniocaudal (CC) and mediolateral oblique (MLO) views. The X-ray tube rotated over a 15° arc with the breast compressed. After scanning, the projection images and the data from the frames were combined to create 3D DBT images and synthesized 2D mammography images were concurrently processed. At the same time, digital mammography with the same field of view as DBT was taken with the same machine.

All synthetic and digital mammography images underwent following preprocessing before radiomics analysis. Each pixel was resampled to 0.1 x 0.1 mm in size. On the resampling process, the linear interpolation method was used for mammography images and the nearest neighbors method was used for masks. Z-normalization was applied for pixel intensities covering the breast.

4. Tumor segmentation

The 2D region of interest (ROI) covering the tumor on synthetic mammography was manually segmented (Figure 2) by one resident radiologist with 3-year of experience (reader 1) using the “mipav” software (<https://mipav.cit.nih.gov>). Then, drawn ROIs were checked in detail and confirmed by a breast radiologist with 25-year of subspecialty experience (reader 2). Disagreements about the ROI were resolved by consensus-based

discussion. Another breast radiologist with 1-year of subspecialty experience (reader 3) independently placed ROI on images for randomly selected 40 patients from training set for evaluate interobserver reproducibility. All readers were blinded to the molecular subtype or pathologic report of the breast cancer.

From the original ROI, an ROI covering only the inner portion and an ROI covering only the marginal portion were obtained (Figure 3). The ROI with inner portion, excluding the marginal area of the lesion, was obtained by applying 3 mm of erosion from the original ROI. The ROI with marginal portion was obtained by subtracting the ROI with the erosion from the original ROI.

For the evaluation of digital mammography, the same ROI drawn in synthetic mammography was applied. In our hospital, the DBT and digital mammography were taken concurrently with the same field of view (FOV).

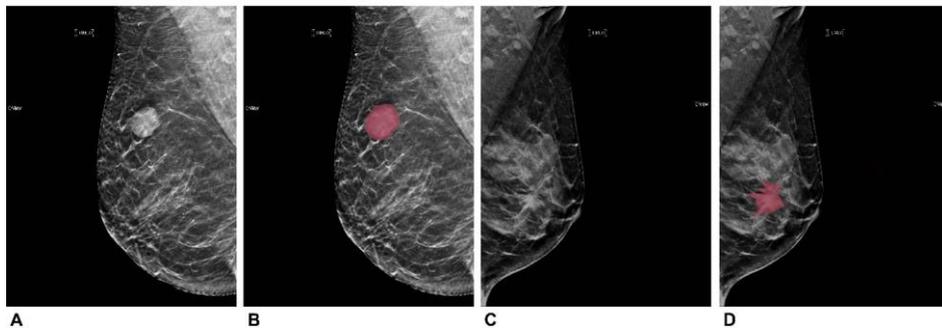


Figure 2. Segmentation example 1. Example of tumor segmentation on synthetic mammography. **A. B.** The synthetic mediolateral oblique view from 58 years old female diagnosed with triple negative subtype of breast cancer. The lesion was shown as circumscribed and round shaped mass with high density. **C. D.** The synthetic mediolateral oblique view from 47 years old female diagnosed with luminal subtype of breast cancer. The lesion was shown as spiculated mass with architectural distortion.

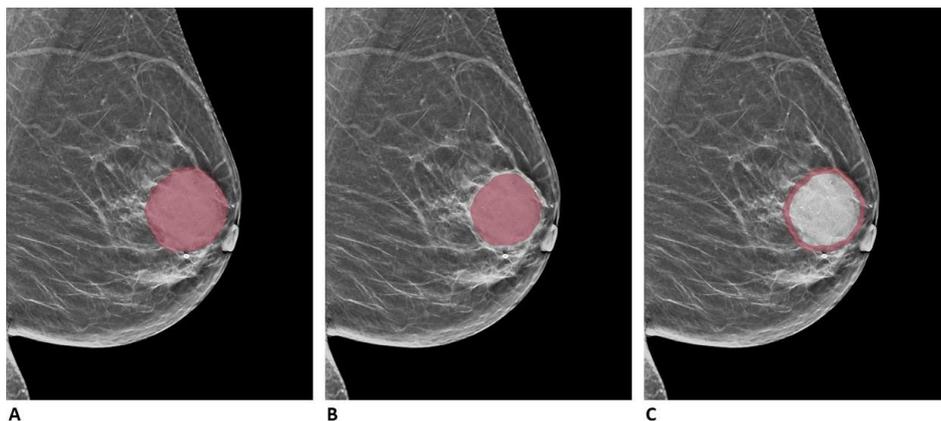


Figure 3. Segmentation example 2. Example of tumor segmentation on synthetic mammography from 57 years old female diagnosed with HER2 subtype of breast cancer. **A.** original region of interest (ROI), **B.** ROI with inner portion, **C.** ROI with marginal portion.

5. Radiomics feature extraction & selection

The calculation of the radiomics feature was based on segmented ROIs and was performed using open source software “pyradiomics” (<https://pyradiomics.readthedocs.io>, version 2.1.2)²⁶. The radiomics feature categories were as follows: (1) first order; 18 features, (2) GLCM; 22 features, (3) GLRLM; 16 features, (4) GLSZM; 16 features. A full list of the specific features contained in each category is described in the supplementary material (Appendix E1). Total of 72 radiomics features for each view were obtained.

Elastic-net approach was used to select the appropriate features and to build a radiomics model. Elastic-net is a logistic regression model which is combination of ridge regression and least absolute shrinkage and selection operator (LASSO)^{27,28}. Parameter tuning process of elastic-net was performed through ten-fold cross validation. For tuning coefficient λ and α , the criterion of minimum standard deviation and maximum AUC was followed, respectively.

Feature selection and modeling process were done using R software (version 3.5.1; <http://www.Rproject.org>) and “glmnet” package (version 2.0-16).

6. Molecular subtype classification

We performed three binary classifications in molecular subtype prediction. This is for intuitive results while avoiding statistical complexity^{29,30}. In order to overcome the unbalance of the number of lesions belonging to each subset in modeling process, we applied the synthetic minority oversampling technique (SMOTE) method. SMOTE is a commonly used oversampling method that was proposed to improve random oversampling^{31,32}. After the modeling process, selected features were extracted and their linear combination formed the radiomics signature of each lesion.

The modeling process was repeated for features from the CC view only (CC model), features from the MLO view only (MLO model), and concatenated features from both views (CC + MLO model).

7. Clinical feature assessment

For all breast cancer lesions, two radiologists (reader 1 and reader 2) evaluated the lesions on synthetic mammography image based on Breast Imaging Reporting and Data System (BI-RADS). Consensus was achieved when the observation of the two radiologists were different. Together with these radiological features based on BI-RADS, the patient's age and lesion size were combined to form clinical features.

8. Statistics analysis

Statistical comparisons for the continuous values were performed by using Student's t-test. Comparisons for categorical variables were performed

by using Pearson's Chi-square test or Fisher's exact test. Univariate and multivariate logistic regression analyses for clinical features were used to identify independent predictors of breast cancer molecular subtype. A “combined model” was built by performing a multivariate logistic regression that included both the radiomics signature and independent variable from multivariate analysis of clinical features. Two-sided $P < .05$ was considered indicative of a statistically significant difference. The classification performances were evaluated based on the Receiver Operating Characteristic (ROC) curve and Area under curve (AUC) in the validation cohort. Comparison of two ROC curves was performed using Delong's test. The correlation between the radiomics signature and BI-RADS features was obtained using Pearson correlation coefficients. Interobserver reproducibility was assessed by intraclass correlation coefficients (ICCs). An $ICC > 0.75$ was considered as a good agreement.

III. RESULTS

A total of 150 patients (TN = 50, HER2 = 50, Luminal = 50) were assigned to the training set and 71 patients (TN = 12, HER2 = 9, Luminal = 50) were assigned to the validation cohort.

1. Radiomics features and prediction performance

Among the all radiomics features, 71 features in MLO view and 58 features in CC view showed good interobserver reproducibility (Figure 4). Finally, a total of 129 features were included in the analysis.

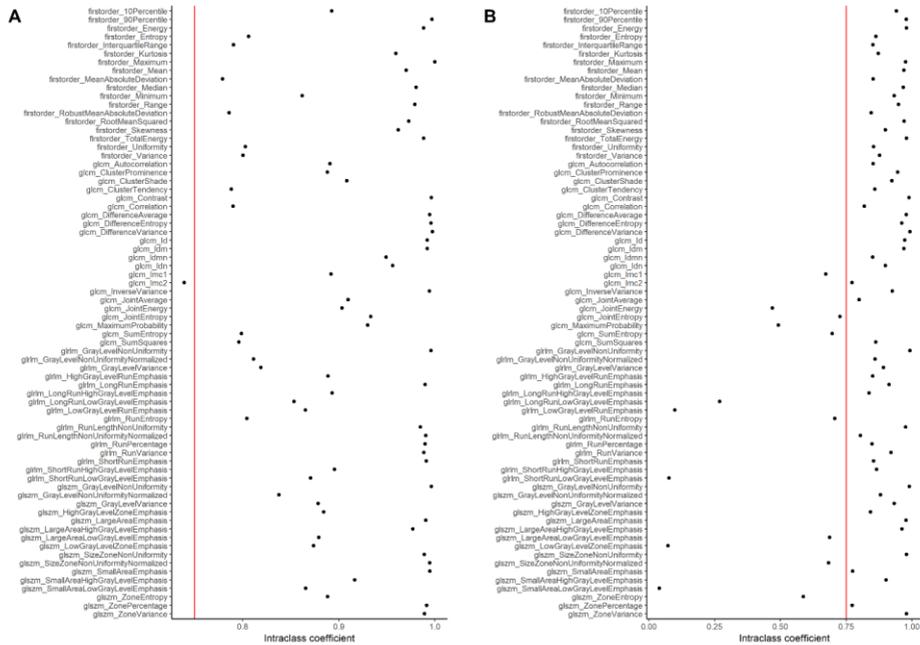


Figure 4. Interobserver reproducibility of radiomics features. **A.** MLO view and **B.** CC view.

When concatenating (CC + MLO model) all features, 20 features were selected for TN vs non-TN, 18 for HER2 vs non-HER2, and 66 features for luminal vs non-luminal. List of selected features are in the supplementary material (Appendix E2). When including only features obtained from CC view (CC model), 6 features were selected for TN vs non-TN, 34 for HER2 vs non-HER2, and 43 features for luminal vs non-luminal. In case of MLO view (MLO model), 17 features were selected for TN vs non-TN, 34 for HER2 vs non-HER2, and 42 features for luminal vs non-luminal. In the training set, the CC+ MLO model yielded an AUC of 0.834 for TN, 0.842 for HER2, and 0.941 for luminal subtypes.

In the validation cohort, the CC + MLO model yielded an AUC of 0.838 for TN, 0.556 for HER2, and 0.645 for luminal subtypes. With the optimal

cut-off value of radiomics signature in this model, sensitivity, and specificity of the models in the validation cohort were 83.3% and 79.7% for TN, 11.1% and 79.0% for HER2, 44.0% and 66.7% for luminal subtypes, respectively (Table 1).

Comparing AUC of CC + MLO, CC and MLO models for three binary classifications, there were no statistically significant differences (Table 2).

Table 1. Classification performance of radiomics models in validation cohort.

| | | TN vs Non-TN | HER2 vs Non-HER2 | Luminal vs Non-luminal |
|-------------------|-------------|--------------|---------------------|---------------------------|
| CC model | AUC | 0.819 | 0.520 | 0.659 |
| | Accuracy | 0.817 | 0.761 | 0.563 |
| | Sensitivity | 0.750 | 0.222 | 0.440 |
| | Specificity | 0.831 | 0.839 | 0.867 |
| MLO model | AUC | 0.791 | 0.645 | 0.627 |
| | Accuracy | 0.718 | 0.747 | 0.521 |
| | Sensitivity | 0.917 | 0.111 | 0.480 |
| | Specificity | 0.678 | 0.839 | 0.619 |
| CC + MLO model | AUC | 0.838 | 0.556 | 0.645 |
| | Accuracy | 0.803 | 0.704 | 0.507 |
| | Sensitivity | 0.833 | 0.111 | 0.440 |
| | Specificity | 0.797 | 0.790 | 0.667 |

Table 2. Comparison of AUC between radiomics models.

| | TN vs Non-TN | HER2 vs Non-HER2 | Luminal vs Non-luminal |
|--------------------------------|-----------------|---------------------|---------------------------|
| CC model vs MLO model | 0.6460 | 0.3545 | 0.5445 |
| CC model vs CC + MLO model | 0.5257 | 0.6940 | 0.7422 |
| MLO model vs CC + MLO model | 0.2500 | 0.1713 | 0.6470 |

The value of each column represents the P value using Delong's test.

2. Comparison between synthetic and digital mammography

Applying the same radiomics analysis as synthetic mammography on digital mammography, classification performance of synthetic mammography was compared with digital mammography in the validation cohort. There were no significant differences between synthetic mammography and digital mammography in predicting the three subtypes (Table 3, Figure 5). The number of features selected for the TN subtype was 20 in synthetic mammography and 48 in digital mammography (Appendix E3).

Table 3. Comparison of classification performances between synthetic mammography and digital mammography in validation cohort.

| | | Synthetic mammography | Digital mammography | P value |
|---------------------------|-------------|--------------------------|------------------------|---------|
| TN vs Non-TN | Sensitivity | 0.833 | 0.750 | 0.812 |
| | Specificity | 0.797 | 0.678 | |
| | Accuracy | 0.803 | 0.690 | |
| | AUC | 0.838 | 0.828 | |
| HER2 vs Non-HER2 | Sensitivity | 0.111 | 0.111 | 0.268 |
| | Specificity | 0.790 | 0.887 | |
| | Accuracy | 0.704 | 0.789 | |
| | AUC | 0.556 | 0.518 | |
| Luminal vs Non-luminal | Sensitivity | 0.440 | 0.420 | 0.833 |
| | Specificity | 0.667 | 0.762 | |
| | Accuracy | 0.507 | 0.521 | |
| | AUC | 0.645 | 0.652 | |

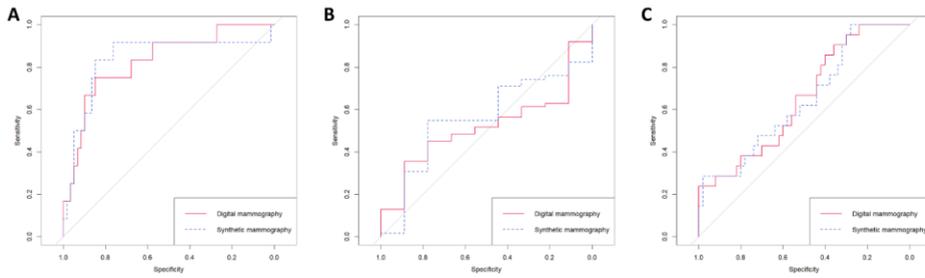


Figure 5. ROC curves of the synthetic mammography and digital mammography in the validation cohort. **A.** TN vs. non-TN, **B.** HER2 vs. non-HER2 and **C.** luminal vs. non-luminal.

3. Considering the impact of margin

On synthetic mammography, three different ROIs were applied for the analysis to determine the effect of different parts of the lesion on the outcome. In predicting the TN subtype, compared to the original ROI, the ROIs with only inner and marginal portions showed lower AUC value than original ROI (Table 4, Figure 6). However, there were no statistically significant differences ($p = 0.084$ and 0.051). For the original ROI, the AUC was 0.838 for distinguishing TN subtype, but for the ROI with inner portion and the ROI with marginal portion, it decreased to 0.672 and 0.754. There was no significant difference between ROI with inner portion and ROI with marginal portion ($p = 0.345$). There were no statistically significant differences in other molecular subtypes.

Table 4. Comparison of classification performances among ROIs in validation cohort.

| | | Original | Inner | Margin | P value | | |
|---------------------------|-------------|----------|----------|----------|---------|--------|--------|
| | | ROI (A) | only (B) | only (C) | A vs B | A vs C | B vs C |
| TN vs Non-TN | Sensitivity | 0.833 | 0.417 | 0.750 | 0.084 | 0.051 | 0.345 |
| | Specificity | 0.797 | 0.881 | 0.762 | | | |
| | Accuracy | 0.803 | 0.803 | 0.760 | | | |
| | AUC | 0.838 | 0.672 | 0.754 | | | |
| HER2 vs Non-HER2 | Sensitivity | 0.111 | 0.222 | 0.222 | 0.209 | 0.883 | 0.229 |
| | Specificity | 0.790 | 0.935 | 0.822 | | | |
| | Accuracy | 0.704 | 0.845 | 0.746 | | | |
| | AUC | 0.556 | 0.556 | 0.554 | | | |
| Luminal vs Non-luminal | Sensitivity | 0.440 | 0.500 | 0.400 | 0.288 | 0.152 | 0.867 |
| | Specificity | 0.667 | 0.524 | 0.667 | | | |
| | Accuracy | 0.507 | 0.507 | 0.479 | | | |
| | AUC | 0.645 | 0.571 | 0.581 | | | |

A. original ROI, B. ROI with inner portion and C. ROI with marginal portion.

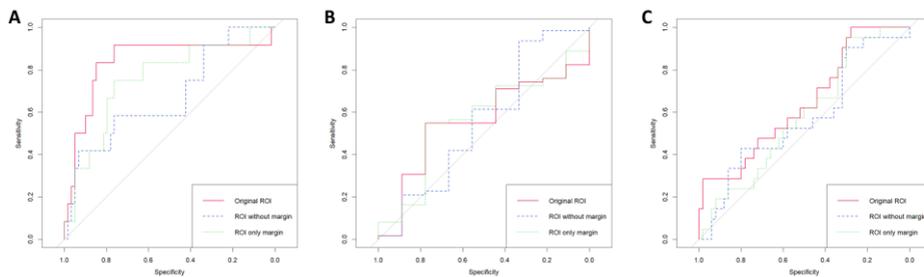


Figure 6. ROC curves of the original ROI, ROI with inner portion and the ROI with marginal portion. **A.** TN vs. non-TN, **B.** HER2 vs. non-HER2 and **C.** luminal vs. non-luminal.

4. Comparison of prediction performance between clinical and radiomics model

We compared the predictive performance of clinical model with radiomics model and combined model. In the TN subtype, univariate analysis of the clinical features showed that round shape, high density and architectural distortion were statistically significant features. In multivariate analysis of the clinical model, round shape and high density were identified as independent factors for predicting TN subtype (Table 5).

Table 5. Univariate and multivariate logistic regression of clinical model and combined model for TN subtype of breast cancer.

| Feature | Univariate analysis | | | | Multivariate analysis | | With radiomics signature | |
|----------------------------|---------------------|------------------|---------|-----------------------|-----------------------|----------------------|--------------------------|----------------------|
| | TN | Non-TN | P value | Odds ratio | P value | Odds ratio | P value | Odds ratio |
| Age | 54.08 ±10.48 | 54.18 ± 9.87 | 0.954 | 0.999 (0.965, 1.034) | | | | |
| Size | 33.98 ±17.45 | 33.35 ± 19.07 | 0.844 | 1.002 (0.983, 1.020) | | | | |
| Breast composition | | | | | | | | |
| Dense | 40 | 71 | Ref | 1 | | | | |
| Fatty | 10 | 29 | 0.239 | 0.612 (0.260, 1.351) | | | | |
| Gross feature | | | | | | | | |
| Mass only | 29 | 46 | Ref | 1 | | | | |
| Mass + calcification | 21 | 47 | 0.330 | 0.709 (0.351, 1.413) | | | | |
| Calcification only | 0 | 7 | 0.986 | NA | | | | |
| Shape | | | | | | | | |
| Oval | 3 | 4 | 0.409 | 1.925 (0.362, 9.235) | | | | |
| Round | 17 | 12 | 0.003 | 3.636 (1.567, 8.696) | 0.016 | 3.028 (1.233, 7.681) | 0.335 | 1.695 (0.575, 4.998) |
| Irregular | 30 | 77 | Ref | 1 | | | | |
| Mass margin | | | | | | | | |
| Obscured | 10 | 18 | 0.986 | 0.992 (0.393, 2.414) | | | | |
| Microlobulated | 7 | 9 | 0.555 | 1.389 (0.452, 4.134) | | | | |
| Indistinct | 28 | 50 | Ref | 1 | | | | |
| Spiculated | 5 | 16 | 0.301 | 0.558 (0.168, 1.598) | | | | |
| Mass density | | | | | | | | |
| Low | 3 | 7 | 0.861 | 1.137 (0.230, 4.478) | | | | |
| Equal | 23 | 61 | Ref | 1 | | | | |
| High | 24 | 24 | 0.013 | 2.546 (1.223, 5.372) | 0.018 | 2.542 (1.180, 5.573) | 0.370 | 1.525 (0.598, 3.834) |
| Architectural distortion | 5 | 25 | 0.036 | 0.333 (0.107, 0.869) | 0.107 | 0.403 (0.121, 1.143) | 0.419 | 0.575 (0.138, 2.084) |
| Calcification morphology | | | | | | | | |
| Benign | 1 | 1 | 0.451 | 3.000 (0.112, 80.288) | | | | |
| Amorphous | 2 | 2 | 0.300 | 3.000 (0.329, 27.556) | | | | |
| Coarse heterogeneous | 3 | 5 | 0.468 | 1.800 (0.327, 8.635) | | | | |
| Fine pleomorphic | 11 | 33 | Ref | 1 | | | | |
| Fine linear branching | 4 | 14 | 0.817 | 0.857 (0.209, 3.009) | | | | |
| Calcification distribution | | | | | | | | |
| Diffuse | 0 | 1 | 0.992 | NA | | | | |
| Regional | 2 | 2 | 0.328 | 2.786 (0.310, 25.086) | | | | |
| Grouped | 4 | 13 | 0.813 | 0.857 (0.214, 2.907) | | | | |
| Linear | 1 | 0 | 0.991 | NA | | | | |
| Segmental | 14 | 39 | Ref | 1 | | | | |
| Radiomics signature | | | <0.001 | 1781 (190, 23225) | | | <0.001 | 828 (78, 12147) |

Multivariate analysis of independent clinical features with radiomics signature revealed that radiomics signature was the only statistically significant variable. The combined model yield AUC value of 0.868 in the

validation cohort (Table 6). In the ROC analysis, the performance of the combined model was statistically significantly higher than the clinical model ($p = 0.045$, Figure 7).

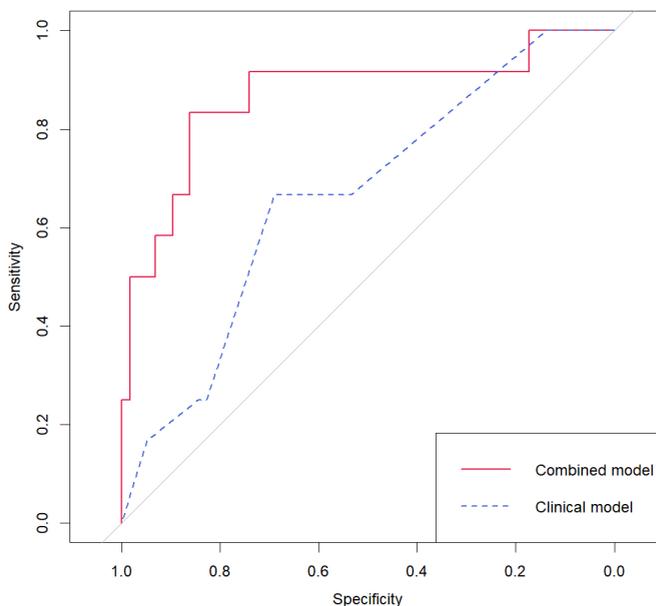


Figure 7. ROC curve of the clinical model and combined model in the validation cohort for TN vs non-TN. The AUC of the combined model was 0.868 and that of the clinical model was 0.665. The two ROC curves showed significant difference ($p = 0.045$).

The results of univariate and multivariate analysis for HER2 and luminal subtypes are presented in supplementary materials (Appendix E4 and E5). The HER2 and luminal subtypes did not show significant differences when comparing the performance of the clinical, radiomics and combined models in the validation cohort (Table 6).

Table 6. AUC value of clinical and combined model in validation cohort.

| | Radiomics model (A) | Clinical model (B) | Combined model (C) | P value | | |
|---------|------------------------|-----------------------|-----------------------|---------|--------|--------|
| | | | | A vs B | A vs C | B vs C |
| TN | 0.838 | 0.665 | 0.868 | 0.048 | 0.766 | 0.045 |
| HER2 | 0.556 | 0.501 | 0.582 | 0.724 | 0.849 | 0.159 |
| Luminal | 0.645 | 0.690 | 0.677 | 0.710 | 0.736 | 0.952 |

5. Correlation between radiomics signature and BI-RADS features

The correlations between radiomics signature and BI-RADS feature in each molecular subtype of breast cancer were shown in Figure 4 in the order of correlation coefficient. For TN subtype, round shape and high density showed high positive correlation with radiomics signature. Architectural distortion and segmental distribution of microcalcification showed negative correlation. For HER2 subtype, segmental distribution of microcalcification, mass with microcalcification and fine linear microcalcification showed positive correlation with radiomics signature, and gross feature of mass showed negative correlation. For luminal subtype, fatty breast composition and spiculated margin showed positive correlation, and obscured margin and dense breast composition showed negative correlation.

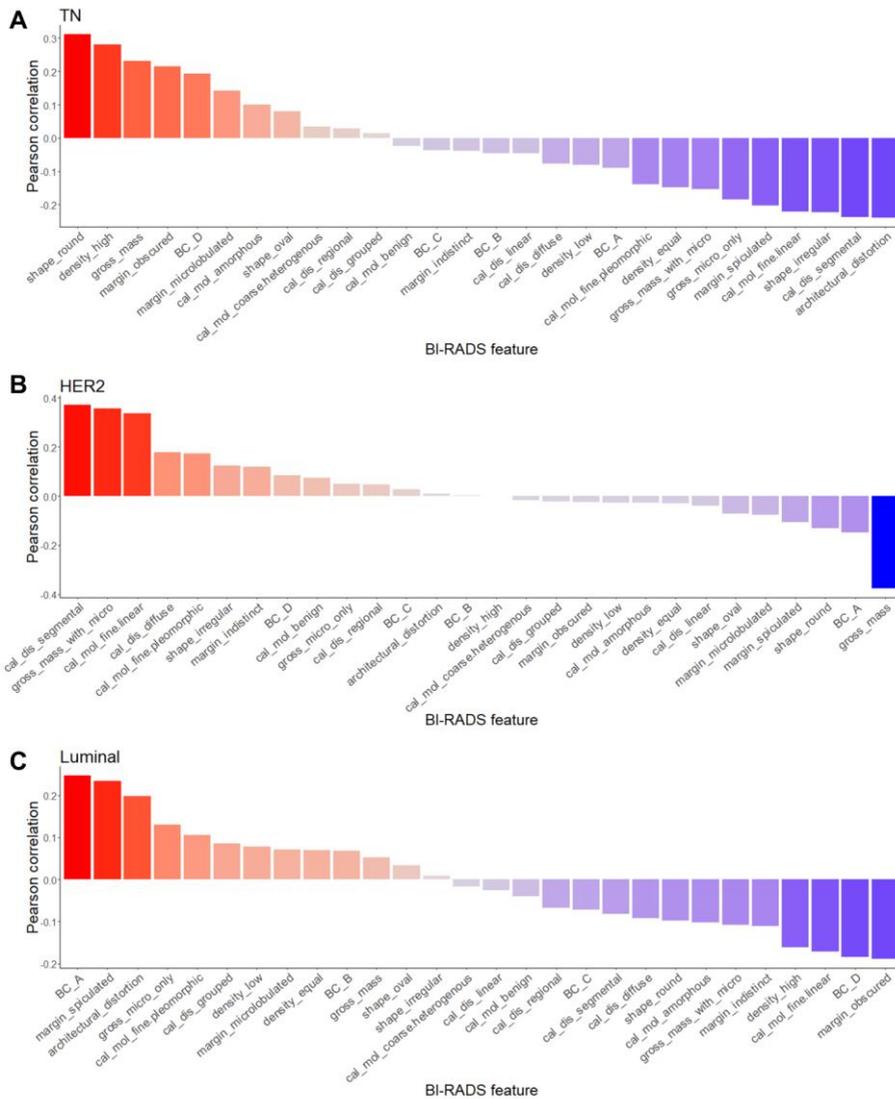


Figure 4. Correlation between radiomics signature and BI-RADS features in the **A.** TN, **B.** HER2 and **C.** luminal subtype of breast cancer.

IV. DISCUSSION

This study has revealed that TN subtype of breast cancer can be distinguished by radiomics analysis in synthetic mammography reconstructed from DBT. The radiomics model showed good performance in identifying TN subtype in the temporally independent validation cohort. In addition, the combined model—combination of clinical model and radiomics signature—showed significantly higher performance compared to the clinical model only. This means that the radiomics signature has an added value compared to the clinical model which is consist of age, tumor size and qualitative image findings.

The combination of DBT and digital mammography in the screening of breast cancer has been reported to show higher sensitivity than digital mammography alone^{17,18}. However, taking mammography and DBT at the same time results in high radiation doses, thus efforts have been made to replace digital mammography with synthetic mammography from DBT¹⁹. Since synthetic mammography with DBT have been reported to have comparable sensitivity with the digital mammography, the use of DBT alone as screening modality has been attempted in North America^{20,21}. As the role of the DBT increases, research into the possibility of applying radiomics in the DBT is being actively conducted.

A study demonstrated that radiomics can be used to discriminate cancerous breasts from patients with dense breasts and negative mammography in DBT³³. Another study showed that Ki-67 expression could be predicted using radiomics in DBT³⁴. Although these studies reported preliminary results, they suggest that the radiomics methodology could be applicable to DBT, similar to mammography. In this study, using radiomics analysis of synthetic mammography from DBT, we could discriminate TN subtype breast cancer with high performance. The TN subtype was known to require different treatment methods due to the absence of targeted agent and have a poorer prognosis than

other subtypes. If the screening modality, DBT, can provide information about the TN subtype using radiomics, it will be helpful for clinician to establish a treatment plan. In addition, radiomics can provide radiologists more confidence in diagnosing TN subtypes using DBT.

Previous studies to predict molecular subtypes of breast cancer using radiomics were focused on MRI¹⁴⁻¹⁶ because of its high soft tissue contrast and visualization of the tumor perfusion dynamics. One study reported an overall accuracy of 71.2% in subtyping using only the radiomics features of MRI, and 89.2% when combined with pathological features¹⁴. Another study reported an AUC of 0.65-0.89 for each subtype in subtyping using MRI data in TCGA/TCIA¹⁵. However, these studies have a limitation in that they only perform internal validation using leave-one-out method without an independent validation set. Due to differences in biological characteristics and treatments of TN subtypes, there have been a study of radiomics to distinguish only TN subtype. Radiomics analysis of both tumor and background parenchymal enhancement has been shown to increase AUC from 0.782 to 0.878 in predicting TN subtype¹⁶. Although the performance of MRI-based radiomics is reported high, the importance of mammography-based radiomics remains valid, because MRI is expensive modality and can only be available in tertiary medical center. Meanwhile, mammography is a first-line imaging modality for cancer screening and is applicable to all breast cancer patients.

Recent pioneering studies suggested the possibility of predicting molecular subtypes by analyzing digital mammography with radiomics^{30,35}. Ma et al. showed that TN, HER2, and luminal subtypes can be distinguished with relatively high performance, and among them, discrimination of TN shows the best performance³⁰. Zhang et al. also reported high performance in distinguishing TN from non-TN using radiomics in digital mammography³⁵. However, these studies have limitations in that they did not evaluated in an independent validation set. In addition, these studies were performed on digital

mammography and it is not known whether the same performance can be guaranteed in synthetic mammography from DBT. The present study showed that the analysis of synthetic mammography with radiomics could predict TN subtype with high performance and validated the performance in an independent validation cohort. Comparing with digital mammography, synthetic mammography showed similar degree of classification performances. This suggests that synthetic mammography is comparable to digital mammography in the field of radiomics analysis.

Previous studies compared digital mammography and synthetic mammography in qualitative aspect^{36,37}, but there was no study on whether they were comparable in quantitative analysis such as radiomics. In the present study, although similar performance was acquired in both modalities, there was a difference in the number of selected features. Despite the difference, the features selected from the two modalities showed considerable overlap. This means that the images were inherently related to the molecular subtype regardless of modality. Studies based on recent reconstruction algorithms for synthetic mammography have shown that synthetic mammography is superior to digital mammography in depicting some features, including increased conspicuity of calcifications, increased definition of spiculated margins, and better visualization of architectural distortions^{37,38}. This is probably because synthetic mammography has the advantage of obtaining more information from multiple projection views based on DBT, which may be a clue to explain that synthetic mammography shows the same performance with fewer features in this study.

Because radiomics extract the features inherent in an image, it is considered to be correlated with the qualitative evaluation by the reader. Several studies have reported that some mammographic findings are associated with the molecular subtype of breast cancer^{39,40}. TN subtype of breast cancer has been reported to be associated with round or oval mass and circumscribed margin³⁹

or oval shaped hyperdense mass⁴⁰. Consistent with these studies, in the present study, the round shape and high density showed a high positive correlation with the radiomics signature for predicting TN subtype. HER2 subtype was reported to have an indistinct margin with suspicious microcalcification³⁹ and luminal subtypes was reported to have a spiculated margin and architectural distortion³⁹. Similar correlations were found between radiomics signature and image findings in the present study. This means that radiomics signature well reflected the known mammographic findings associated with each molecular subtype. Conversely, this result means that new image findings can be found intuitively through the morphological features represented by the combination of features revealed through radiomics analysis.

Both the texture inside the lesion and the margin of the lesion are related to the molecular subtype^{39,40}. In this study, we extracted the radiomics feature independently from the inner portion of the lesion and from the marginal portion of the lesion. In both cases, the performances for distinguishing the TN subtype were decreased compared to the original ROI, but there was no statistically significant difference. From this result, it is concluded that both the texture and the border of the lesion are significant in differentiating the molecular subtype. However, there was a limitation of the ROI manipulation method using simple erosion for a complex lesion on mammography. More rigorous method to separately extract the features from texture and margin of the lesion would be needed in the future research.

In the task of distinguishing HER2 and luminal subtype of breast cancer, the radiomics models failed to show sufficient performances. In addition, there was no added value of combining clinical model with radiomics model. In predicting HER2 and luminal subtypes breast cancer, the radiomics model appeared to be overfitted to the training set and showed inferior performance in the validation cohort. This means that, unlike TN type, radiomics failed to extract general characteristics suitable for HER2 and luminal types. The

previous study has reported that there is much in common with the phenotype between HER2 and luminal types⁴¹. Microcalcifications and the form of non-masses are well-known common morphologic characteristics of the two subtypes. Therefore, this result may be a methodological limitation of radiomics, but may actually be due to a classification limitation based on the morphology between HER2 and luminal subtypes.

In this study, we excluded the radiomics features belonging to the shape class which is one of the commonly used classes in radiomics analysis. This is to prevent the researcher's intention from being reflected in the results. Features belonging to the shape class are determined by the shape of the ROI itself²⁶. The manually drawn ROI may vary in shape depending on the researcher's subjectivity. As already mentioned above, the shape of the lesion is related to the molecular subtype. Therefore, if shape features were included in the analysis, it can be expected that classification performance would be improved. However, semiautomatic segmentation is necessary to apply it to mammography while maintaining objectivity⁴². In mammography, it is difficult to distinguish the lesion from the surrounding parenchyma, so there are limitations in applying semiautomatic segmentation method to mammography.

There are several limitations in this study. First, there were inherent limitations of retrospective study design. Second, relatively large number of radiomics features were included in the final model. This makes it difficult to interpret the meaning of each radiomics feature. By showing the relationship between radiomics signature and mammographic features, we verified that mammographic findings were reflected in molecular subtype prediction by radiomics analysis. Third, the extraction of radiomics features were based on the manually drawn ROIs. To overcome this, features with poor interobserver reproducibility were excluded from the analysis. Another limitation was that we only included lesions that were clearly delineated in synthetic mammography. Because the lesion contrast of synthetic mammography was limited compared

to the original DBT image, a relatively large number of lesions were excluded from the analysis.

V. CONCLUSION

In conclusion, this study showed a significant relationship between the radiomics signature based on synthetic mammography reconstructed from DBT images and the molecular subtype of breast cancer. The radiomics signature from synthetic mammography was able to distinguish TN subtypes of breast cancer with high accuracy. Since DBT is an image modality that can be widely used in mostly all patients, it can be used as a potential biomarker that can help clinical diagnosis and treatment in many breast cancer patients.

REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9-29.
2. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama* 2006;295:2492-502.
3. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 2011;22:1736-47.
4. Lam SW, Jimenez CR, Boven E. Breast cancer classification by proteomic technologies: current state of knowledge. *Cancer Treat Rev* 2014;40:129-38.
5. Huber KE, Carey LA, Wazer DE. Breast cancer molecular subtypes in patients with locally advanced disease: impact on prognosis, patterns of recurrence, and response to therapy. *Semin Radiat Oncol* 2009;19:204-10.
6. Metzger-Filho O, Sun Z, Viale G, Price KN, Crivellari D, Snyder RD, et al. Patterns of Recurrence and outcome according to breast cancer subtypes in lymph node-negative disease: results from international breast cancer study group trials VIII and IX. *J Clin Oncol* 2013;31:3083-90.
7. Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *Journal of clinical oncology* 2008;26:1275-81.
8. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *New England journal of medicine* 2010;363:1938-48.
9. Silver DP, Richardson AL, Eklund AC, Wang ZC, Szallasi Z, Li Q, et al. Efficacy of neoadjuvant Cisplatin in triple-negative breast cancer. *Journal of clinical oncology* 2010;28:1145.
10. Wu M, Ma J. Association Between Imaging Characteristics and Different Molecular Subtypes of Breast Cancer. *Acad Radiol* 2017;24:426-34.
11. Celebi F, Pilanci KN, Ordu C, Agacayak F, Alco G, Ilgun S, et al. The role of ultrasonographic findings to predict molecular subtype, histologic grade, and hormone receptor status of breast cancer. *Diagn Interv Radiol* 2015;21:448-53.
12. Uematsu T, Kasami M, Yuen S. Triple-negative breast cancer: correlation between MR imaging and pathologic findings. *Radiology* 2009;250:638-47.
13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than

- Pictures, They Are Data. *Radiology* 2016;278:563-77.
14. Sutton EJ, Dashevsky BZ, Oh JH, Veeraraghavan H, Apte AP, Thakur SB, et al. Breast cancer molecular subtype classifier that incorporates MRI features. *J Magn Reson Imaging* 2016;44:122-9.
 15. Li H, Zhu Y, Burnside ES, Huang E, Drukker K, Hoadley KA, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ Breast Cancer* 2016;2.
 16. Wang J, Kato F, Oyama-Manabe N, Li R, Cui Y, Tha KK, et al. Identifying triple-negative breast cancer using background parenchymal enhancement heterogeneity on dynamic contrast-enhanced MRI: a pilot radiomics study. *PloS one* 2015;10:e0143308.
 17. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 2013;267:47-56.
 18. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. *The Lancet Oncology* 2016;17:1105-13.
 19. Olgar T, Kahn T, Gosch D. Average glandular dose in digital mammography and breast tomosynthesis. *Rofo* 2012;184:911-8.
 20. Skaane P, Bandos AI, Eben EB, Jepsen IN, Krager M, Haakenaasen U, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiology* 2014;271:655-63.
 21. Zuckerman SP, Conant EF, Keller BM, Maidment AD, Barufaldi B, Weinstein SP, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. *Radiology* 2016;281:730-6.
 22. Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LAG, et al. US Incidence of Breast Cancer Subtypes Defined by Joint Hormone Receptor and HER2 Status. *JNCI: Journal of the National Cancer Institute* 2014;106.
 23. Wei Q, Dunbrack RL, Jr. The role of balanced training and testing data sets for binary classifiers in bioinformatics. *PLoS One* 2013;8:e67863.
 24. Yasaka K, Akai H, Abe O, Kiryu S. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology* 2018;286:887-96.
 25. Hammond ME, Hayes DF, Wolff AC, Mangu PB, Temin S. American

- society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Oncol Pract* 2010;6:195-7.
26. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017;77:e104-e7.
 27. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)* 2005;67:301-20.
 28. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software* 2010;33:1.
 29. Park HJ, Lee SS, Park B, Yun J, Sung YS, Shim WH, et al. Radiomics Analysis of Gadoteric Acid-enhanced MRI for Staging Liver Fibrosis. *Radiology* 2019;290:380-7.
 30. Ma W, Zhao Y, Ji Y, Guo X, Jian X, Liu P, et al. Breast Cancer Molecular Subtype Prediction by Mammographic Radiomic Features. *Acad Radiol* 2019;26:196-201.
 31. Bowyer KW. SMOTE: Synthetic Minority Over-sampling Technique. *The Journal of artificial intelligence research* 2002;16:321-57.
 32. Blagus R, Lusa LJBB. SMOTE for high-dimensional class-imbalanced data. 2013;14:106.
 33. Tagliafico AS, Valdora F, Mariscotti G, Durando M, Nori J, La Forgia D, et al. An exploratory radiomics analysis on digital breast tomosynthesis in women with mammographically negative dense breasts. *Breast* 2018;40:92-6.
 34. Tagliafico AS, Bignotti B, Rossi F, Matos J, Calabrese M, Valdora F, et al. Breast cancer Ki-67 expression prediction by digital breast tomosynthesis radiomics features. *Eur Radiol Exp* 2019;3:36.
 35. Zhang HX, Sun ZQ, Cheng YG, Mao GQ. A pilot study of radiomics technology based on X-ray mammography in patients with triple-negative breast cancer. *J Xray Sci Technol* 2019;27:485-92.
 36. Choi JS, Han B-K, Ko EY, Ko ES, Hahn SY, Shin JH, et al. Comparison between two-dimensional synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammography for the detection of T1 breast cancer. *European Radiology* 2016;26:2538-46.
 37. Mariscotti G, Durando M, Houssami N, Fasciano M, Tagliafico A, Bosco D, et al. Comparison of synthetic mammography, reconstructed from digital breast tomosynthesis, and digital mammography: evaluation of lesion conspicuity and BI-RADS assessment categories. *Breast Cancer Research and Treatment* 2017;166:765-73.
 38. Ratanaprasatporn L, Chikarmane SA, Giess CS. Strengths and

- Weaknesses of Synthetic Mammography in Screening. *RadioGraphics* 2017;37:1913-27.
39. Boisserie-Lacroix M, Hurtevent-Labrot G, Ferron S, Lippa N, Bonnefoi H, Mac Grogan G. Correlation between imaging and molecular classification of breast cancers. *Diagn Interv Imaging* 2013;94:1069-80.
 40. Kim MY, Choi N. Mammographic and ultrasonographic features of triple-negative breast cancer: a comparison with other breast cancer subtypes. *Acta Radiol* 2013;54:889-94.
 41. Ko ES, Lee BH, Kim H-A, Noh W-C, Kim MS, Lee S-A. Triple-negative breast cancer: correlation between imaging and pathological findings. *European Radiology* 2010;20:1111-7.
 42. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust Radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One* 2014;9:e102107.

APPENDICES

Appendix E1. List of all features included in this study.

| Class | Features | Class | Features |
|-------------|-----------------------------|-------|----------------------------------|
| First order | 10Percentile | GLRLM | GrayLevelNonUniformity |
| | 90Percentile | | GrayLevelNonUniformityNormalized |
| | Energy | | GrayLevelVariance |
| | Entropy | | HighGrayLevelRunEmphasis |
| | InterquartileRange | | LongRunEmphasis |
| | Kurtosis | | LongRunHighGrayLevelEmphasis |
| | Maximum | | LongRunLowGrayLevelEmphasis |
| | MeanAbsoluteDeviation | | LowGrayLevelRunEmphasis |
| | Mean | | RunEntropy |
| | Median | | RunLengthNonUniformity |
| | Minimum | | RunLengthNonUniformityNormalized |
| | Range | | RunPercentage |
| | RobustMeanAbsoluteDeviation | | RunVariance |
| | RootMeanSquared | | ShortRunEmphasis |
| | Skewness | | ShortRunHighGrayLevelEmphasis |
| | TotalEnergy | | ShortRunLowGrayLevelEmphasis |
| | Uniformity | | |
| Variance | | | |
| GLCM | Autocorrelation | GLSZM | GrayLevelNonUniformity |
| | JointAverage | | GrayLevelNonUniformityNormalized |
| | ClusterProinance | | GrayLevelVariance |
| | ClusterShade | | HighGrayLevelZoneEmphasis |
| | ClusterTendency | | LargeAreaEmphasis |
| | Contrast | | LargeAreaHighGrayLevelEmphasis |
| | Correlation | | LargeAreaLowGrayLevelEmphasis |
| | DifferenceAverage | | LowGrayLevelZoneEmphasis |
| | DifferenceEntropy | | SizeZoneNonUniformity |
| | DifferenceVariance | | SizeZoneNonUniformityNormalized |
| | JointEnergy | | SmallAreaEmphasis |
| | JointEntropy | | SmallAreaHighGrayLevelEmphasis |
| | Imc1 | | SmallAreaLowGrayLevelEmphasis |
| | Imc2 | | ZoneEntropy |
| | Idm | | ZonePercentage |
| | Idmn | | ZoneVariance |
| | Id | | |
| | Idn | | |
| | InverseVariance | | |
| | MaximumProbability | | |
| SumEntropy | | | |
| SumSquares | | | |

Appendix E2. List of selected features in each radiomics model

| Prediction | Selected features |
|------------------------|--|
| TN vs non-TN | MLO_firstorder_10Percentile MLO_firstorder_Entropy MLO_firstorder_MeanAbsoluteDeviation MLO_glcm_SumEntropy MLO_glrIm_GrayLevelNonUniformityNormalized MLO_glrIm_RunEntropy MLO_glszm_GrayLevelNonUniformityNormalized CC_firstorder_10Percentile CC_firstorder_Entropy CC_firstorder_MeanAbsoluteDeviation CC_firstorder_Median CC_firstorder_Minimum CC_firstorder_Uniformity CC_firstorder_Variance CC_glcm_ClusterTendency CC_glcm_SumSquares CC_glrIm_GrayLevelNonUniformityNormalized CC_glrIm_GrayLevelVariance CC_glszm_GrayLevelNonUniformityNormalized CC_glszm_GrayLevelVariance |
| HER2 vs non-HER2 | MLO_firstorder_InterquartileRange MLO_firstorder_Kurtosis MLO_firstorder_Minimum MLO_firstorder_RobustMeanAbsoluteDeviation MLO_glcm_ClusterProminence MLO_glcm_Idn MLO_glszm_GrayLevelNonUniformityNormalized MLO_glszm_LargeAreaHighGrayLevelEmphasis MLO_glszm_LowGrayLevelZoneEmphasis MLO_glszm_SizeZoneNonUniformity MLO_glszm_ZoneEntropy CC_firstorder_Kurtosis CC_firstorder_Minimum CC_glcm_JointAverage CC_glcm_ClusterTendency CC_glcm_Imc2 CC_glszm_LargeAreaEmphasis CC_glszm_SizeZoneNonUniformity |
| Luminal vs non-luminal | MLO_firstorder_10Percentile MLO_firstorder_Entropy MLO_firstorder_InterquartileRange MLO_firstorder_Kurtosis MLO_firstorder_Mean MLO_firstorder_Range MLO_firstorder_RootMeanSquared MLO_firstorder_Skewness MLO_firstorder_Variance MLO_glcm_JointAverage MLO_glcm_ClusterShade MLO_glcm_ClusterTendency MLO_glcm_DifferenceVariance MLO_glcm_Idm MLO_glcm_Idmn MLO_glcm_Id MLO_glcm_InverseVariance MLO_glcm_SumEntropy |

MLO_glcm_SumSquares
MLO_glrlm_GrayLevelNonUniformity
MLO_glrlm_LongRunLowGrayLevelEmphasis
MLO_glrlm_LowGrayLevelRunEmphasis
MLO_glrlm_RunLengthNonUniformityNormalized
MLO_glrlm_RunPercentage
MLO_glrlm_ShortRunEmphasis
MLO_glrlm_ShortRunLowGrayLevelEmphasis
MLO_glszm_HighGrayLevelZoneEmphasis
MLO_glszm_LargeAreaLowGrayLevelEmphasis
MLO_glszm_LowGrayLevelZoneEmphasis
MLO_glszm_SmallAreaHighGrayLevelEmphasis
MLO_glszm_SmallAreaLowGrayLevelEmphasis
MLO_glszm_ZoneEntropy
MLO_glszm_ZonePercentage
CC_firstorder_10Percentile
CC_firstorder_Entropy
CC_firstorder_Kurtosis
CC_firstorder_Maximum
CC_firstorder_MeanAbsoluteDeviation
CC_firstorder_Mean
CC_firstorder_Median
CC_firstorder_Minimum
CC_firstorder_Range
CC_firstorder_RootMeanSquared
CC_firstorder_Skewness
CC_firstorder_Uniformity
CC_firstorder_Variance
CC_glcm_Autocorrelation
CC_glcm_ClusterShade
CC_glcm_ClusterTendency
CC_glcm_Correlation
CC_glcm_Imc2
CC_glcm_SumSquares
CC_glrlm_GrayLevelNonUniformity
CC_glrlm_GrayLevelNonUniformityNormalized
CC_glrlm_GrayLevelVariance
CC_glrlm_HighGrayLevelRunEmphasis
CC_glrlm_RunLengthNonUniformity
CC_glrlm_ShortRunHighGrayLevelEmphasis
CC_glszm_GrayLevelNonUniformity
CC_glszm_HighGrayLevelZoneEmphasis
CC_glszm_LargeAreaEmphasis
CC_glszm_LargeAreaHighGrayLevelEmphasis
CC_glszm_SizeZoneNonUniformity
CC_glszm_SmallAreaEmphasis
CC_glszm_SmallAreaHighGrayLevelEmphasis
CC_glszm_ZoneVariance

Appendix E3. Selected features from synthetic mammography and digital mammography in predicting TN subtype.

| Synthetic mammography (N=20) | Digital mammography (N=48) |
|--|--|
| MLO_firstorder_10Percentile | MLO_firstorder_90Percentile |
| MLO_firstorder_Entropy | MLO_firstorder_Entropy |
| MLO_firstorder_MeanAbsoluteDeviation | MLO_firstorder_Kurtosis |
| MLO_glcM_SumEntropy | MLO_firstorder_Maximum |
| MLO_glrIm_GrayLevelNonUniformityNormalized | MLO_firstorder_MeanAbsoluteDeviation |
| MLO_glrIm_RunEntropy | MLO_firstorder_Mean |
| MLO_glszm_GrayLevelNonUniformityNormalized | MLO_firstorder_Minimum |
| CC_firstorder_10Percentile | MLO_firstorder_RootMeanSquared |
| CC_firstorder_Entropy | MLO_firstorder_Uniformity |
| CC_firstorder_MeanAbsoluteDeviation | MLO_glcM_ClusterShade |
| CC_firstorder_Median | MLO_glcM_Correlation |
| CC_firstorder_Minimum | MLO_glcM_Imc1 |
| CC_firstorder_Uniformity | MLO_glcM_Idmn |
| CC_firstorder_Variance | MLO_glcM_SumEntropy |
| CC_glcM_ClusterTendency | MLO_glrIm_GrayLevelNonUniformityNormalized |
| CC_glcM_SumSquares | MLO_glrIm_LongRunMeanGrayLevelEmphasis |
| CC_glrIm_GrayLevelNonUniformityNormalized | MLO_glrIm_LowGrayLevelRunEmphasis |
| CC_glrIm_GrayLevelVariance | MLO_glrIm_RunEntropy |
| CC_glszm_GrayLevelNonUniformityNormalized | MLO_glszm_GrayLevelNonUniformityNormalized |
| CC_glszm_GrayLevelVariance | MLO_glszm_SizeZoneNonUniformity |
| | CC_firstorder_10Percentile |
| | CC_firstorder_Entropy |
| | CC_firstorder_Kurtosis |
| | CC_firstorder_MeanAbsoluteDeviation |
| | CC_firstorder_Mean |
| | CC_firstorder_Median |
| | CC_firstorder_Minimum |
| | CC_firstorder_RootMeanSquared |
| | CC_firstorder_Skewness |
| | CC_firstorder_Uniformity |
| | CC_firstorder_Variance |
| | CC_glcM_ClusterTendency |
| | CC_glcM_DifferenceEntropy |
| | CC_glcM_DifferenceVariance |
| | CC_glcM_Imc2 |
| | CC_glcM_Idmn |
| | CC_glcM_SumSquares |
| | CC_glrIm_GrayLevelNonUniformityNormalized |
| | CC_glrIm_GrayLevelVariance |
| | CC_glrIm_LongRunEmphasis |
| | CC_glrIm_RunPercentage |
| | CC_glrIm_RunVariance |
| | CC_glrIm_ShortRunEmphasis |
| | CC_glszm_GrayLevelNonUniformityNormalized |
| | CC_glszm_GrayLevelVariance |
| | CC_glszm_LargeAreaEmphasis |
| | CC_glszm_LargeAreaHighGrayLevelEmphasis |
| | CC_glszm_ZoneVariance |

Appendix E4. Univariate and multivariate logistic regression of clinical model and combined model for HER2 subtype of breast cancer.

| Feature | Univariate analysis | | | | Multivariate analysis | | With radiomics signature | |
|-----------------------------------|---------------------|---------------|----------------------|-----------------------|-----------------------|-----------------------|--------------------------|----------------------|
| | HER2 | Non-HER2 | P value | Odds ratio | P value | Odds ratio | P value | Odds ratio |
| Age | 52.70 ± 8.51 | 54.87 ± 10.69 | 0.213 | 0.978 (0.944, 1.012) | | | | |
| Size | 47.78 ± 19.55 | 29.45 ± 16.56 | < 0.001 | 1.037 (1.018, 1.059) | 0.036 | 1.024 (1.002, 1.048) | 0.3815 | 0.987 (0.956, 1.016) |
| Breast composition | | | | | | | | |
| Dense | 41 | 70 | Ref | 1 | | | | |
| Fatty | 9 | 30 | 0.118 | 0.512 (0.211, 1.149) | | | | |
| Gross feature | | | | | | | | |
| Mass only | 16 | 59 | Ref | 1 | | | | |
| Mass + calcification | 30 | 38 | 0.004 | 2.911 (1.418, 6.157) | 0.991 | NA | 0.9934 | NA |
| Calcification only | 4 | 3 | 0.050 | 4.917 (0.990, 27.170) | | | | |
| Shape | | | | | | | | |
| Oval | 1 | 6 | 0.231 | 0.268 (0.014, 1.648) | | | | |
| Round | 4 | 25 | 0.018 | 0.258 (0.072, 0.722) | 0.081 | 0.352 (0.095, 1.047) | 0.0326 | 0.208 (0.042, 0.795) |
| Irregular | 41 | 66 | Ref | 1 | | | | |
| Mass margin | | | | | | | | |
| Obscured | 12 | 16 | 0.515 | 1.339 (0.548, 3.225) | | | | |
| Microlobulated | 3 | 13 | 0.194 | 0.412 (0.089, 1.412) | | | | |
| Indistinct | 28 | 50 | Ref | 1 | | | | |
| Spiculated | 3 | 18 | 0.069 | 0.298 (0.066, 0.977) | | | | |
| Mass density | | | | | | | | |
| Low | 1 | 9 | 0.102 | 0.172 (0.009, 0.977) | | | | |
| Equal | 33 | 51 | Ref | 1 | | | | |
| High | 12 | 37 | 0.084 | 0.501 (0.222, 1.079) | | | | |
| Architectural distortion | | | | | | | | |
| 7 | 23 | 0.198 | 0.545 (0.202, 1.318) | | | | | |
| Calcification morphology | | | | | | | | |
| Benign | 0 | 2 | 0.988 | NA | | | | |
| Amorphous | 1 | 3 | 0.541 | 0.481 (0.023, 4.107) | | | | |
| Coarse heterogeneous | 3 | 5 | 0.857 | 0.867 (0.161, 3.995) | | | | |
| Fine pleomorphic | 18 | 26 | Ref | 1 | | | | |
| Fine linear branching | 13 | 5 | 0.030 | 3.756 (1.192, 13.452) | 0.130 | 2.707 (0.773, 10.604) | 0.3706 | 2.015 (0.447, 9.978) |
| Calcification distribution | | | | | | | | |
| Diffuse | 1 | 0 | 0.995 | NA | | | | |
| Regional | 0 | 4 | 0.989 | NA | | | | |
| Grouped | 5 | 12 | 0.076 | 0.345 (0.098, 1.070) | | | | |
| Linear | 0 | 1 | 0.994 | NA | | | | |
| Segmental | 29 | 24 | Ref | 1 | | | | |
| Radiomics signature | | | <0.001 | 283 (50, 2140) | | | < 0.001 | 616 (61, 9168) |

Appendix E5. Univariate and multivariate logistic regression of clinical model and combined model for luminal subtype of breast cancer.

| Feature | Univariate analysis | | | | Multivariate analysis | | With radiomics signature | |
|-----------------------------------|---------------------|---------------|--------------------|-----------------------|-----------------------|-----------------------|--------------------------|-----------------------|
| | Luminal | Non-Luminal | P value | Odds ratio | P value | Odds ratio | P value | Odds ratio |
| Age | 55.66 ± 10.95 | 53.39 ± 9.52 | ± 0.193 | 1.023 (0.989, 1.059) | | | | |
| Size | 24.92 ± 14.41 | 37.88 ± 18.85 | ± <0.001 | 0.947 (0.919, 0.972) | 0.0017 | 0.946 (0.191, 0.977) | 0.8641 | 0.996 (0.946, 1.044) |
| Breast composition | | | | | | | | |
| Dense | 30 | 81 | Ref | 1 | | | | |
| Fatty | 20 | 19 | 0.007 | 2.842 (1.339, 6.098) | 0.0139 | 3.289 (1.291, 8.708) | 0.3406 | 1.996 (0.472, 8.487) |
| Gross feature | | | | | | | | |
| Mass only | 30 | 45 | Ref | 1 | | | | |
| Mass + calcification | 17 | 51 | 0.058 | 0.500 (0.240, 1.015) | | | | |
| Calcification only | 3 | 4 | 0.883 | 1.125 (0.209, 5.454) | | | | |
| Shape | | | | | | | | |
| Oval | 3 | 4 | 0.621 | 1.479 (0.279, 7.057) | | | | |
| Round | 8 | 21 | 0.537 | 0.751 (0.289, 1.809) | | | | |
| Irregular | 36 | 71 | Ref | 1 | | | | |
| Mass margin | | | | | | | | |
| Obscured | 6 | 22 | 0.487 | 0.694 (0.231, 1.861) | | | | |
| Microlobulated | 6 | 10 | 0.461 | 1.527 (0.471, 4.638) | | | | |
| Indistinct | 22 | 56 | Ref | 1 | | | | |
| Spiculated | 13 | 8 | 0.006 | 4.136 (1.535, 11.794) | 0.2444 | 2.344 (0.544, 10.011) | 0.4703 | 2.243 (0.229, 2.045) |
| Mass density | | | | | | | | |
| Low | 6 | 4 | 0.109 | 3.000 (0.793, 12.560) | | | | |
| Equal | 28 | 56 | Ref | 1 | | | | |
| High | 13 | 36 | 0.413 | 0.722 (0.324, 1.555) | | | | |
| Architectural distortion | 18 | 12 | 0.001 | 4.125 (1.810, 9.724) | 0.0066 | 5.577 (1.663, 20.503) | 0.2661 | 3.250 (0.448, 29.423) |
| Calcification morphology | | | | | | | | |
| Benign | 1 | 1 | 0.649 | 1.933 (0.073, 51.238) | | | | |
| Amorphous | 1 | 3 | 0.714 | 0.644 (0.030, 5.538) | | | | |
| Coarse heterogeneous | 2 | 6 | 0.616 | 0.644 (0.087, 3.204) | | | | |
| Fine pleomorphic | 15 | 29 | Ref | 1 | | | | |
| Fine linear branching | 1 | 17 | 0.044 | 0.114 (0.006, 0.640) | 0.1985 | 0.167 (0.005, 1.688) | 0.4579 | 0.298 (0.006, 4.625) |
| Calcification distribution | | | | | | | | |
| Diffuse | 0 | 1 | 0.992 | NA | | | | |
| Regional | 2 | 2 | 0.169 | 4.300 (0.471, 39.595) | | | | |
| Grouped | 8 | 9 | 0.025 | 3.822 (1.178, 12.664) | 0.4426 | 1.794 (0.403, 8.268) | 0.1657 | 4.310 (0.578, 38.452) |
| Linear | 0 | 1 | 0.992 | NA | | | | |
| Segmental | 10 | 43 | Ref | 1 | | | | |
| Radiomics signature | | | <0.001 | 1536 (221, 17693) | | | < 0.001 | 1673 (170, 3099) |

ABSTRACT(IN KOREAN)

디지털 유방 토모신세시스의 합성 유방촬영술에서 얻은
라디오믹스 시그니처를 이용한 유방암의 분자 아형 예측

<지도교수 김은경>

연세대학교 대학원 의학과

손진우

목적: 디지털 유방 토모신세시스를 기반으로 하는 합성 유방촬영술에서 추출한 라디오믹스 시그니처를 이용하여 유방암의 분자아형을 예측하는 것을 목적으로 한다.

방법: 2015년 12월부터 2016년 9월까지, 침습성 유방암으로 진단받았으며, 수술 전 디지털 유방 토모신세시스 영상이 있는 365명의 환자를 연구에 포함하였다. 2015년 12월부터 2016년 7월까지 294명의 환자들 중 150명의 환자[각각의 분자 아형 (luminal A+B, luminal; HER2-positive, HER2; triple-negative, TN) 별로 순서대로 50명씩]를 선별하여 학습집합에 배정했다. 별도로 2016년 8월부터 2016년 9월까지 71명의 환자들 (luminal 50명, HER2 9명, TN 12명)을 검증집합에 배정하였다. 합성 유방촬영술과 디지털 유방촬영술의 상하위 촬영과 내외사위 촬영에서 총 129개의 라디오믹스 특성을 추출하였다. 추가로 병변의 내부 부분과 둘레 부분에서 각각 독립적으로 라디오믹스 특성을 추출했다. 각각의 분자 아형에 대해 이항 분류를 시행했으며, 분류의 성능은 수용자 작용 특징 곡선(ROC, Receiver Operating Characteristic)과 곡선하면적(AUC, area under the curve)을 통해 평가했다. 라디오믹스 모델은 학습집합에서 elastic-net을 이용하여 구성되었으며, 독립된 검증집합에서 평가되었다. 라디오믹스 모델과 비교하기 위한 임상 모델은

환자의 나이, 병변의 크기, 그리고 BI-RADS에 기반한 영상 소견으로 구성되었다.

결과: 검증집합에서 라디오믹스 모델은 TN에 대해 0.838 그리고 HER2에 대해 0.556, luminal에 대해 0.645의 곡선하면적을 보였다. 최적의 절단값을 적용했을 때, 라디오믹스 시그니처의 민감도와 특이도는 TN에서 83.3%와 79.7%, 그리고 HER2에서 11.1% 와 79.0%, luminal에서 44.0%와 66.7%를 보였다. 합성 유방촬영술과 디지털 유방촬영술을 비교했을 때, 통계적으로 유의미한 차이는 없었다. 병변의 전체를 분석했을 때와 비교하여, 병변의 내부와 둘레 부분만을 독립적으로 분석했을 때에도 유의미한 차이는 없었다. TN 아형을 예측함에 있어서 임상 모델과 라디오믹스 시그니처를 함께 다변량분석 했을 때, 라디오믹스 시그니처는 TN 아형에 대한 유일한 독립 예측 인자였다. 뿐만 아니라 TN 아형을 예측하는데 있어서, 임상 모델과 라디오믹스 시그니처를 결합했을 때, 임상 모델 단독에 비해 유의미하게 높은 곡선하면적을 보였다.

결론: 디지털 유방 토모신세시스의 합성 유방촬영술에서 얻은 라디오믹스 시그니처는 높은 정확도로 TN 아형의 유방암을 예측할 수 있었다. 합성 유방촬영술에서 얻은 라디오믹스 시그니처는 유방암의 TN 아형을 구분하고 향후 치료의 방향에 영향을 미치는 바이오마커의 역할을 할 수 있을 것으로 기대된다.

핵심되는 말: 유방암, 분자 아형, 라디오믹스, 디지털 유방 토모신세시스, 합성 유방촬영술

