# Impact of Variations in Tumor Delineation on Machine Learning-based Breast MRI radiomics

**Sepideh Hatamikia [1, 2]*, Geevarghese George [1], Florian Schwarzhans [1], Amirreza Mahbod [1], Ramona Woitek [1]**

[1] Danube Private University, Krems, Austria
[2] Austrian Center for Medical Innovation and Technology (ACMIT), Wiener Neustadt, Austria

**Poster # 1465**

## Introduction

- Manual delineation of volumes of interest (VOIs) by experts is considered the gold-standard method in radiomics analysis. However, it suffers from inter- and intra-operator variability. A quantitative assessment of the impact of variations in these delineations on the performance of the radiomics predictors is required to develop robust radiomics based prediction models.

## Aim of the study

- To develop radiomics models for the prediction of pCR to neoadjuvant chemotherapy in patients with two different breast cancer subtypes (Triple negative (TNBC) and human epidermal growth receptor 2 positive positive (HER2+)) based on contrast-enhanced magnetic resonance imaging acquired prior to treatment (baseline MRI scans) [1].
- To evaluate robustness of these models against different VOIs modification.

## Dataset

- We used the multi Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and moLecular Analysis (I-SPY1 TRIAL) breast MRI dataset.
- This is an open-access dataset that includes contrast-enhanced MRI and tissue-based biomarkers to predict pathological complete response (pCR) and relapse-free survival.

## Methodology

- 102 radiomics features including shape, first-order, and higher-order features were extracted using Pyradiomics (an open-source Python package).
- 9 different feature selection methods including F-Score, Relief, Mutual Information (MI), Gini Importance, LASSO, Genetic Algorithm (GA), Sequential Backward Search (SBS), Sequential Forward Search (SFS), and Recursive Feature Elimination (RFE) were used.
- Area Under the ROC Curve (AUC), sensitivity and specificity were calculated.
- Different mathematical operations such as erosion, smoothing, dilation, randomization, and ellipse fitting were applied to the original manual VOIs delineated by experts to simulate variations of segmentation masks (Figure 1).

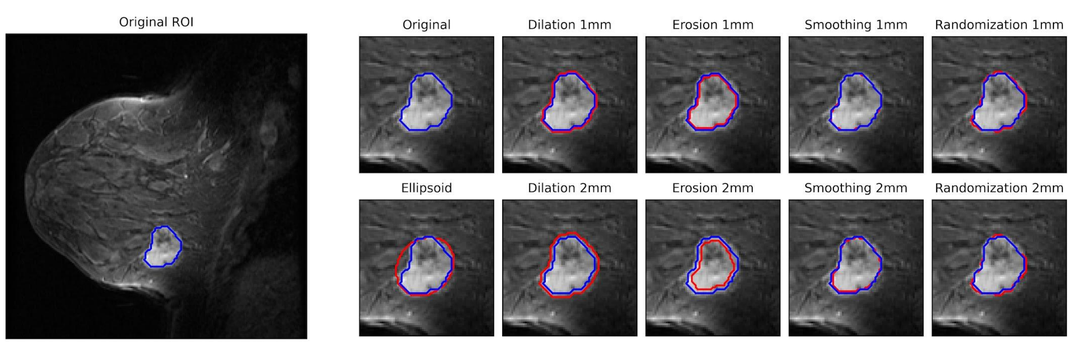


**Figure 1.** Demonstration of the effect of VOI manipulation on a single slice of the breast MRI. Blue contour shows original volume of interest (VOI) outline and red contour shows the corresponding operation.

- The effects of such VOI modifications on various steps of the radiomics workflow, including feature extraction, feature selection, and prediction performance, were evaluated (Figure 2).
- Logistic Regression (LR) and Linear Discriminant Analysis (LDA) were used as classifier.
- Intraclass Correlation Coefficient (ICC) was used to evaluate the similarity between VOI modification and manual segmentations (baseline).

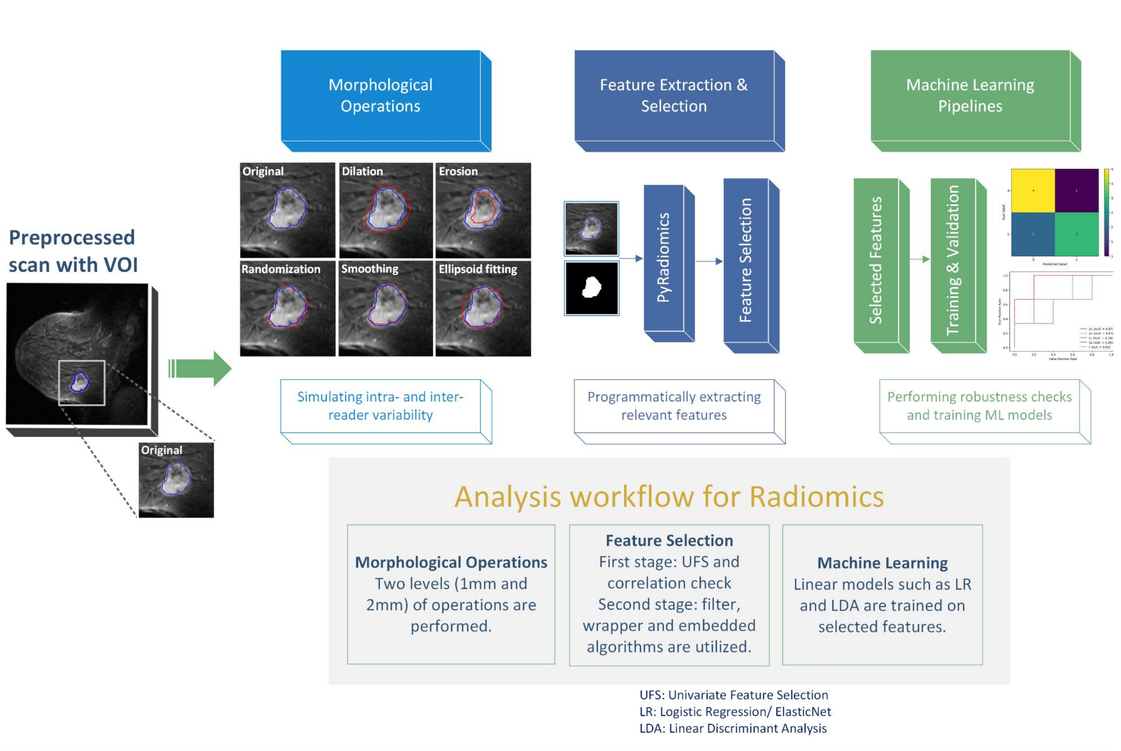


**Figure 2.** A diagram presenting the general proposed radiomics workflow.

## Results

| FS algorithm and classifier | Modification | Train AUC, SE, SP | Test AUC, SE, SP | Avg. ICC of selected features |
|---|---|---|---|---|
| LASSO, LDA | none | 0.82; 0.72; 0.80 | 0.96; 0.68; 1.00 | - |
| LASSO, LDA | Dilation 1 mm | 0.66; 0.55; 0.76 | 0.72; 0.60; 0.96 | 0.94 |
| LASSO, LDA | Dilation 2 mm | 0.19; 0.12; 0.53 | 0.46; 0.36; 0.68 | 0.69 |
| LASSO, LDA | Erosion 1 mm | 0.67; 0.62; 0.66 | 0.88; 0.76; 0.84 | 0.96 |
| LASSO, LDA | Erosion 2 mm | 0.61; 0.48; 0.56 | 0.60; 0.64; 0.60 | 0.79 |
| LASSO, LDA | Ellipsoid fitting | 0.48; 0.20; 0.65 | 0.39; 0.32; 0.52 | 0.22 |
| LASSO, LDA | Randomization 1 mm | 0.78; 0.67; 0.81 | 0.74; 0.60; 1.00 | 0.95 |
| LASSO, LDA | Randomization 2 mm | 0.73; 0.62; 0.70 | 0.69; 0.60; 0.92 | 0.91 |
| LASSO, LDA | Smoothing 1 mm | 0.78; 0.77; 0.75 | 0.93; 0.64; 0.88 | 0.99 |
| LASSO, LDA | Smoothing 2 mm | 0.80; 0.67; 0.70 | 0.80; 0.60; 0.84 | 0.91 |

**Table 1.** The Area under the curve (AUC), sensitivity (SE), specificity (SP), train and test change (%) and average ICC of features related to a selected model using original manual VOIs as well as different VOI modifications for HER2+ breast cancer group.

| FS algorithm and classifier | Modification | Train AUC, SE, SP | Test AUC, SE, SP | Avg. ICC of selected features |
|---|---|---|---|---|
| SFS, LDA | none | 0.94; 0.80; 0.95 | 0.89; 0.60; 0.80 | - |
| SFS, LDA | Dilation 1 mm | 0.80; 0.60; 0.80 | 0.80; 0.40; 0.96 | 0.89 |
| SFS, LDA | Dilation 2 mm | 0.76; 0.60; 0.68 | 0.35; 0.33; 0.56 | 0.68 |
| SFS, LDA | Erosion 1 mm | 0.84; 0.60; 0.90 | 0.92; 0.73; 0.80 | 0.95 |
| SFS, LDA | Erosion 2 mm | 0.74; 0.50; 0.90 | 0.97; 0.80; 0.88 | 0.81 |
| SFS, LDA | Ellipsoid fitting | 0.54; 0.30; 0.85 | 0.35; 0.33; 0.84 | 0.31 |
| SFS, LDA | Randomization 1 mm | 0.86; 0.60; 0.85 | 0.83; 0.40; 1.00 | 0.88 |
| SFS, LDA | Randomization 2 mm | 0.81; 0.50; 0.85 | 0.45; 0.33; 0.92 | 0.82 |
| SFS, LDA | Smoothing 1 mm | 0.92; 0.70; 0.85 | 0.93; 0.67; 0.80 | 0.99 |
| SFS, LDA | Smoothing 2 mm | 0.92; 0.80; 0.85 | 0.85; 0.73; 0.80 | 0.98 |

**Table 2.** The Area under the curve (AUC), sensitivity (SE), specificity (SP), train and test change (%) and average ICC of features related to a selected model using original manual VOIs as well as different VOI modifications for for TNBC breast cancer group.

- Using manual tumour VOIs and radiomics features extracted from baseline MRI scans, an AUC of up to 0.96 and 0.89 was achieved for HER2+ and TNBC, respectively.
- For smoothing and erosion, VOIs yielded the highest number of robust features and the best prediction performance, while ellipse fitting and dilation led to the lowest robustness and prediction performance for both breast cancer subtypes.

## Discussion and conclusion

- This is the first study to comprehensively evaluate the effects of different tumour VOI modifications on radiomics analyses in HER2+ breast cancer and TNBC.
- Our systematic evaluation showed that different VOI modifications can lead to significant differences in radiomics feature values, feature selection and prediction performance.
- Determining a predefined standard for tumour delineation can help develop reliable and robust radiomics models.
- The results of this study can serve as a reference for future radiomics research.

[1] Hatamikia S, George G, Schwarzhans F, Mahbod A, Woitek R. Breast MRI radiomics and machine learning-based predictions of response to neoadjuvant chemotherapy - How are they affected by variations in tumor delineation? Comput Struct Biotechnol J. 2023 Nov 19;23:52-63.

*Sepideh.Hatamikia@dp-uni.ac.at