

# 日本医療研究開発機構 官民による若手研究者発掘支援事業 事後評価報告書



## I 基本情報

補助事業課題名: (日本語) 機械学習による拡散 MRI 乳がん診断支援システムの開発  
(プログラム名) (英語) Development of a Diffusion MRI Breast Cancer Diagnosis Support System  
Using Machine Learning

実施期間: 令和 4 年 6 月 1 日 ~ 令和 6 年 3 月 31 日

補助事業担当者 氏名: (日本語) 飯間 麻美  
(英語) Ima Mami

補助事業担当者 所属機関・部署・役職:  
(日本語) 京都大学医学部附属病院 先端医療研究開発機構・特定講師  
(英語) Kyoto University Hospital, Program Specific Lecturer

## II 補助事業の概要

乳がんはがん罹患率のトップであり、的確な治療に直結する新たな画像診断法の開発は急務である。本事業では、拡散 MRI による最先端の定量技術と機械学習の手法を融合させることにより乳腺腫瘍の良悪性鑑別やサブタイプ鑑別に有用な新たな診断法を開発し、分子学的予後予測因子や再発・予後等の情報と関連付けることにより新たな診断支援システムを開発する。乳房 MRI 画像を用いて病変の良悪性やサブタイプの識別、治療効果や予後の評価を向上させることを目指す。多数症例の乳房 MRI 画像につき機械学習を用いて詳細に解析することにより乳腺腫瘍の良悪性やサブタイプ鑑別、治療効果や予後評価に真に有用な診断法を開発する。拡散 MRI 定量値においては病理学的裏付けが十分でなく、定量値と組織情報との相関の探索につき、マウス腫瘍モデルを用いて進める。

### 研究開発の成果

約 530 症例の乳房病変を対象とし、組織像が確認された、または経過観察で腫大が認められなかった 334 の乳房病変を解析した。データは 10-fold cross-validation データ(悪性 139 例、良性 67 例)と test データ(悪性 85 例、良性 43 例)に分け、症例の重複は認められなかった。3T 乳房 MRI を用い、造影 MRI に加えて 1500s/mm<sup>2</sup> までの 5 つの(拡散強調の度合いを示す) b 値を用いて拡散強調画像を撮影を撮像した。

良性腫瘍と悪性腫瘍を鑑別可能な機械学習モデルを開発するために、データ補強法を活用し様々な組み合わせにつき比較した： A：ランダム弾性変形、B：ランダムアフィン変換/ランダムノイズ、C：ミックスアップ。Small 2D CNN モデルは、3D CNN や ResNet18 とも比較した。これらの機械学習を用いた拡散強調像と標準的な乳房造影ダイナミック MRI による診断能の比較については、面積下曲線の推定 (AUC) を含む受信者操作特性曲線 (ROC) 分析を用いて行われた。

補強はすべての実験で精度を改善 (AUC:0.86-0.90 VS. cross-validation データでは 0.86、AUC : 0.85-0.88 VS. test データでは 0.87)、データ補強 C (AUC: 0.86 と 0.83) を除くすべての実験で精度が向上した。2D CNN を用いた結果は、3D CNN よりも AUC が良い結果となった (AUC:0.88-0.90 VS 0.72-0.75)。また、テストデータセット (AUC:0.88) に対しても最良であり、標準的な乳房 MRI (AUC:0.89、0.87) と同等であった。特異度は、標準的な乳房 MRI よりも、DW 画像からの 2D CNN による機械学習モデルの方が高い傾向があり (cross-validation データでは 85.1%対 81%、test データでは 81.4%対 74%)、感度は標準的な乳房 MRI の方が高かった (cross-validation データでは 79.9%対 98%、test データでは 85.9%対 99%)。

機械学習を用いて乳房 MRI の診断性能を評価した研究は複数存在するものの、複数の b 値による乳房拡散強調画像 (DWI) のみを機械学習モデルで評価した研究は稀である。機械学習の手法により、病変セグメンテーションを必要とせず、病変を含むスライス情報のみを用いて、乳腺腫瘍の悪性・良性の鑑別において、標準的な乳腺 MRI に匹敵する良好な性能を達成できることが示された。

拡散強調像は標準的な乳房造影ダイナミック MRI よりも特異度が高いが、感度はやや劣る。これらの結果は、乳腺専門医の読影時間を増やすことなく、標準的な乳房 MRI の診断性能を特に特異度改善の点で補完し、造影剤にアレルギーのある患者にも有益であるという大きな利点を持つ。拡散強調 MRI を用いた機械学習モデルは、乳房腫瘍の良性・悪性の鑑別における特異性を向上させ、診断精度を高められる可能性がある。実用化への準備は整っており、本機能を搭載したソフトウェアを日常臨床で使用するためのワークフローの整備中である。

動物実験においては各拡散 MRI 定量値と血管密度などの病理学的情報との相関の有無につき、バイオマーカーとなるか検討した。組織内の微小循環を評価可能な IVIM (intravoxel incoherent motion) パラメータが生体内における抗 PD-1 抗体治療後の組織変化を反映するかどうかを検討した。大腸癌 (CT26) および乳癌 (4T1) 異種移植マウスモデルに抗 PD-1 抗体治療を行い、その後 MR 撮影を実施した。CT26 モデルでは f (flowing blood volume fraction) が治療群で有意に低下したのに対し、4T1 モデルでは変化は見られなかった。また、両モデルとも病変のサイズに有意差は認められなかった。

抗 PD-1 抗体投与異種移植モデル間の f (flowing blood volume fraction) パラメータと病変サイズの差は、治療効果のレベルの違いを反映している可能性がある。

さらに、腫瘍の微小環境を再現した PDX (Patient-derived xenograft) モデルを用いて、トリプルネガティブ乳癌における DW パラメータと Ki-67 発現との関連も調査した。組織学的 Ki-67 染色で検証された ADC と非ガウス拡散パラメータを評価した。全体として、すべての Ki-67 レベルにおいて、拡散時間の増加とともに Sindex は増加し、sADC と ADCo は減少した。S-index は拡散時間 9ms および 27.6ms で Ki-67 レベルと有意な正の相関を示し、sADC は負の相関を示した。ADC は Ki-67 レベルと負の相関を示したが、その相関は短い拡散時間 (5ms と 9ms) でより顕著であった。K と Ki-67 との相関は見られなかった。臨床で撮影可能な拡散時間での DW パラメーターは Ki-67 を予測可能なイメージングバイオマーカーとなる可能性があり、さらにその最適化を他モデルにおいて検証する。

本事業により開発された技術とソフトウェアは、乳がんの正確な診断の補助となり、個別化医療の推進に貢献する。これにより、患者の予後改善と生活の質の向上が期待される。

英文：

This research project aimed to develop a novel diagnostic method for breast cancer by combining advanced quantitative diffusion MRI technology with machine learning techniques. The primary objectives were to distinguish between benign and malignant breast tumors, identify subtypes, and improve treatment efficacy and prognosis evaluation. The study analyzed 334 breast lesions with confirmed histology or no enlargement on follow-up, using 3T breast MRI to obtain diffusion-weighted images with multiple b-values.

To ensure robust evaluation, the data were carefully divided into two sets: a 10-fold cross-validation dataset (139 malignant, 67 benign cases) and a separate test dataset (85 malignant, 43 benign cases). This division allowed for thorough model training and validation, followed by an unbiased assessment of the final model performance. Importantly, there was no overlap of cases between these datasets, maintaining the integrity of the evaluation process.

A significant focus of the study was the development and optimization of machine learning models for tumor classification. Various data augmentation techniques were compared, including random elastic deformation, random affine transformation/random noise, and mix-up. The researchers experimented with both 2D and 3D CNN architectures, as well as ResNet18. The 2D CNN models consistently outperformed 3D CNN models, achieving an AUC of 0.88-0.90 for cross-validation data and 0.88 for test data, comparable to standard breast MRI (AUC: 0.89 and 0.87, respectively). Notably, the machine learning model based on diffusion-weighted images demonstrated higher specificity than standard breast MRI (85.1% vs. 81% for cross-validation data, 81.4% vs. 74% for test data), although sensitivity was lower (79.9% vs. 98% for cross-validation data, 85.9% vs. 99% for test data). These results suggest that the machine learning approach could complement standard MRI by improving specificity without increasing reading time for breast specialists.

In animal experiments, the correlation between each diffusion MRI quantitative value and pathological information such as vascular density was examined to see if it could be used as a biomarker. The study investigated whether IVIM (intravoxel incoherent motion) parameters, which can assess microcirculation in tissues, reflect tissue changes after anti-PD-1 antibody treatment in vivo. Anti-PD-1 antibody treatment was administered to colon cancer (CT26) and breast cancer (4T1) xenograft mouse models, followed by MR imaging.  $f$  (flowing blood volume fraction) in the CT26 model was significantly decreased in the treatment group, whereas no change was observed in the 4T1 model. There was also no significant difference in lesion size in both models. The differences in  $f$  parameters and lesion size between the anti-PD-1 antibody-treated xenograft models may reflect different levels of treatment efficacy.

The research also investigated the association between DW parameters and Ki-67 expression in triple-negative breast cancer using the patient-derived xenograft (PDX) model, which recapitulates the tumor microenvironment. ADC and non-Gaussian diffusion parameters validated by histological Ki-67 staining were evaluated. Overall, at all Ki-67 levels, Sindex increased with increasing diffusion time, while

sADC and ADC<sub>0</sub> decreased. s-index showed a significant positive correlation with Ki-67 levels at diffusion times of 9 ms and 27.6 ms, while sADC showed a negative correlation. ADC showed a negative correlation with Ki-67 levels, more pronounced at shorter diffusion times (5 and 9 ms). No correlation was found between K and Ki67. DW parameters at diffusion times that can be captured in clinical practice may serve as a predictive imaging biomarker for Ki67, and its optimization will be further validated in other models.

The technology and software developed through this project have the potential to assist in the accurate diagnosis of breast cancer, contribute to personalized medicine, and ultimately improve patient prognosis and quality of life. The researchers are preparing for practical implementation, including the development of a workflow for routine clinical use of the software, with plans for PMDA face-to-face advice to facilitate clinical adoption.