MEDICAL UNIVERSITY OF VIENNA

Identification of aggressive breast cancer lesions with [18F]-FDG-PET/CT in combination with machine learning and data pre-processing

Denis Krajnc[1], Laszlo Papp[1], Thomas S. Nakuz[2], Heinrich F. Magometschnigg[3], Marko Grahovac[2], Clemens P. Spielvogel[2,4], Zsuzsanna Bago-Horvath[5], Alexander Haug[2,4], Georgios Karanikas[2], Thomas Beyer[1], Marcus Hacker[2], Thomas H Helbich[3], Katja Pinker[3,6]

[1] QIMP Team, Center for Medical Physics and Biomedical Engineering, MUW
[2] Department of Biomedical Imaging and Image-guided Therapy, Division of Nuclear Medicine MUW
[3] Department of Biomedical Imaging and Image-guided Therapy, Division of Molecular and Gender Imaging, MUW
[4] Christian Doppler Laboratory for Applied Metabolomics, MUW
[5] Clinical Institute of Pathology, MUW
[6] Department of Radiology, Breast Imaging Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Objective

The aim of this study was to investigate the performance of breast cancer detection and the identification of aggressive breast cancer from [18F]-FDG-PET/CT incorporating data pre-processing algorithms in machine learning predictive models.

Patients and Methods

A cohort of 170 patients with suspicious imaging finding (BI-RADS 4/5) was examined with [18F]-FDG-PET/CT imaging as part of a prospective study. Histopathology was used as the standard of reference. Breast tumours were classified as benign or malignant. Lesions were classified as (a) less aggressive to targeted treatment (luminal A/B, Her2-positive) and (b) more aggressive (triple negative (TN) breast cancer) based on immunohistochemical (IHC) receptor status and proliferation rate (Table 1).
173 lesions from 170 patients were delineated in the Hermes Hybrid 3D software. Radiomic features were extracted from all delineated lesions. Ensemble learning approaches were applied to predict tumour malignancy as well as IHC triple negative subgroup in a 100-fold Monte Carlo cross-validation scheme. This step was performed twice for both reference labels: (a) with training the predictive models on the original training datasets, and (b) with processing the training datasets by data pre-processing approaches. Predictive performance was estimated over the ML fold validation cases with confusion matrix calculations (sensitivity (SENS), specificity (SPEC) and overall accuracy (ACC)) as well as area under the receiver operator characteristics curve (AUC). Conventional PET correlation analyses were performed for TN subgroup and malignant/benign tumor status. Contribution of imaging and patient characteristics features to prediction was estimated as well.

Results

The breast cancer detection ML model yielded 0.81 AUC. Data pre-processing improved the performance of the model by 4% (Acc), 10% (Spec), and 0.04 (AUC). SUV-based model yielded 0.76 (AUC) (Figure 1a).
The triple negative ML model which yielded 0.85 AUC. Here, data pre-processing increased the performance of the ML model by 6% (Acc), 9% (Sens), 2% (Spec), and 0.06 (AUC). SUV-based model yielded 0.70 (AUC) (Figure 1b).
In the breast cancer detection and the triple negative ML models, PET features were highly important compared with CT (60% PET vs 25% CT, and 51% PET vs 32% CT contribution, respectively) (Figure 2).

Conclusion

We demonstrated that ML models based on radiomic tumour features extracted from [18F]-FDG-PET/CT in conjunction with the utilization of data preparation approaches provide breast cancer detection and identification of aggressive breast cancers with high accuracy.

Patient characteristics (n=170)	Value
Age (years), median (IQR)	57.6 (16–86)
Malignancy	n (%)
Malignant	132 (78)
Benign	38 (22)
Estrogen (ER)	n (%)
–	17 (10)
+	88 (52)
NA	65 (38)
Progesterone (PR)	n (%)
–	27 (16)
+	78 (46)
NA	65 (38)
Ki-67	n (%)
–	26 (15)
+	73 (43)
NA	71 (42)
HER2	n (%)
–	84 (49)
+	22 (13)
NA	64 (38)
Triple negative	n (%)
Yes	11 (6)
No	95 (56)
NA	64 (38)
Luminal A/B	n (%)
A	14 (8)
B	81 (48)
NA	75 (44)

Table 1: Patient cohort characteristics for malignancy, estrogen (ER), progesterone (PR), human epidermal growth receptor 2 (HER2), Ki-67 protein expression, triple negative, and luminal A/B status. NA=Not Available.

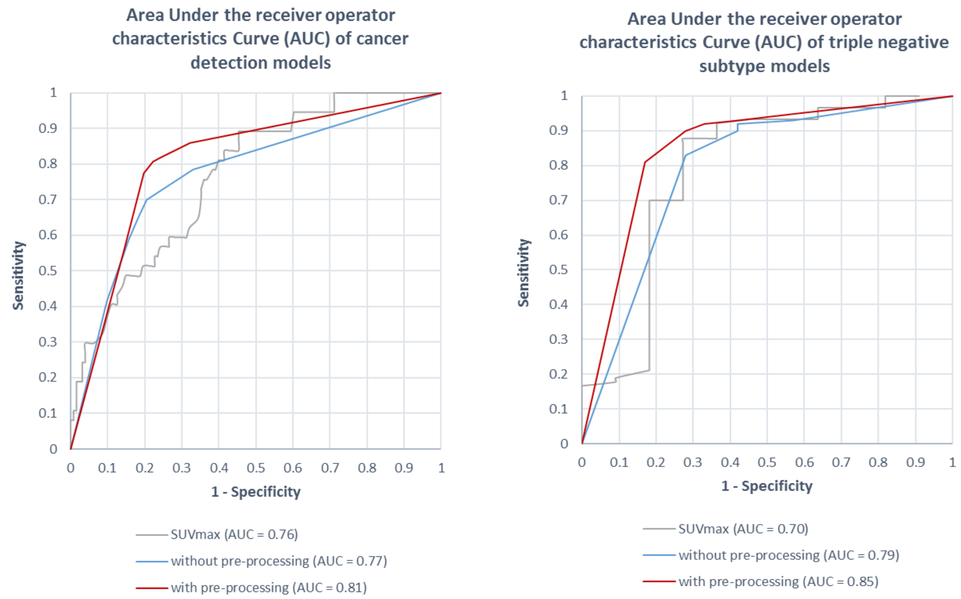

Figure 1: Comparison of area under the receiver operator characteristics curve (AUC) performance of maximum standard uptake value (SUVmax) and machine learning-based ensemble models for (a) cancer detection (b) triple negative subtype

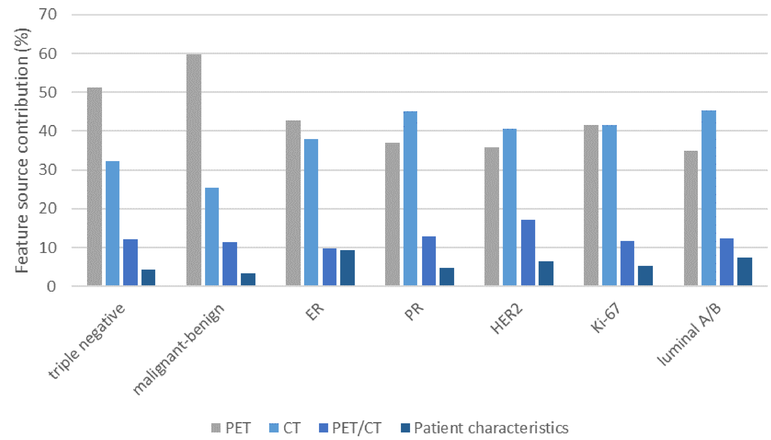

Figure 2: Contributions of imaging and patient characteristics to predict reference labels as determined by the utilized ensemble learning scheme. The weights were collected from the respective best-performing ensemble model variant (original dataset vs pre-processed) of each reference label. Modality importance is expressed in percentages (%)