in the validating set, with 0.83 and 0.80 respectively, followed by LASSO+KNN (AUC=0.83, ACC=0.71). Mi+NB or AdaBoost, as well as Wilcoxon+NB or RF, had good performance with an AUC of 0.80. SVM had the best mean performance in the cross-validation and validation cohort (only accuracy). RF had the best mean of AUC in the validation cohort. Conclusion: this study showed that image features obtained from a pretreatment 18F-FDG-PET/CT could predict the metabolic response in recurrent or metastatic breast cancer, by their incorporation in a ML model, which performance depends largely on the feature selection and ML classifier methods selected. LASSO+SVN/RF had the better performance.

References: None

EPS-226
Reproducibility of a semi-automatic gradient-based segmentation approach for lymphoma PET
F. Yousefizivi1, I. Bloise1, P. Martineau2, D. Wilson2, F. Benard3, T. Bradshaw1, A. Rahim1, C. Uribe4, I. Bloise1
1BC Cancer Research Institute, Vancouver, BC, CANADA, 2BC Cancer, Vancouver, BC, CANADA, 3Department of Radiology, University of Wisconsin, Madison, WI, UNITED STATES OF AMERICA, 4Department of Radiology, University of British Columbia, Vancouver, BC, CANADA.

Aim/Introduction: Metabolic tumor volume (MTV) and total lesion glycolysis (TLG) are identified as prognostic metrics for survival analysis and treatment planning in lymphoma. However, these values are not routinely reported due to the time-consuming task of segmentation. Artificial intelligence (AI) methods have shown promise in performing automatic segmentation but their performance is limited to the size of training data that requires reliable labels. However, manual delineations suffer from inter-observer variability, even when using semi-automatic methods (e.g. fixed thresholding). We aim to develop a semi-automatic workflow that is robust and reproducible between observers to simplify the development of an extensive training dataset for automatic MTV reporting on diffuse large B-cell lymphoma (DLBCL) PET scans. Materials and Methods: A semi-automatic workflow was created using the gradient-based segmentation method (MIM Software, USA). Three experienced nuclear medicine physicians independently segmented nine cases from a cohort of DLBCL PET scans. Lesions were labeled with 15 prepopulated anatomical regions (cervical lymph nodes, thoracic, etc.) as the variant site/size of lymphoma lesions can affect the performance of AI techniques. Physicians were informed that delineations would be used for TMTV calculations. Intraclass correlation coefficient (ICC) for the different segmentations was calculated for MTV, total lesion glycolysis (TLG), max and mean standard uptake (SUV) values. Additional features such as Standard deviation to mean ratio (Std_Mean), Kurtosis, Skewness were considered (for heterogeneity assessment). The segmentation results were also evaluated using the STAPLE algorithm that provides an estimation of true segmentation. Results: We observed high repeatability (ICC ≥ 95%) with respect to a number of quantitative measures (i.e. MTV, TLG, and SUVmax) between physicians. The measured ICC values were MTV=98%, TLG=99%, SUVtot=99%, SUVmax=96% (ICC>95%), SUVmean=92% (ICC>90%), Std_Mean=71% (ICC>70%). Based on the ICC values, Skewness=29% and Kurtosis=0.43 (ICC<50%) had significantly low repeatability among the physicians. The Dice scores compared to the ground truth (estimated by STAPLE) were 0.88, 0.79, and 0.88; the Jaccard values were 0.83, 0.72, and 0.82, and the Hausdorff Distances were 3.06, 4.18, and 3.95 respectively. Conclusion: Our initial results suggest that our workflow is reproducible for MTV and a number of quantitative metrics. While some features such as Kurtosis and Skewness depicted poor reproducibility. The segmentation evaluation also showed that the delineated regions were relatively close to one another. This has the potential of facilitating the creation of a multi-institutional dataset to develop reliable AI models and to facilitate routine reporting of TMTV and other features. References: none

EPS-227
Integrating gene mutation mutual exclusion logic and radiomics to improve the efficacy of predicting EGFR mutation in non-small cell lung cancer
J. Wang, F. Kang, J. Wang: Department of Nuclear Medicine, Xiijing Hospital, Fourth Military Medical University, Xi’an, Shaanxi, CHINA.

Aim/Introduction: Gene mutations are mutually exclusive in non-small cell lung cancer(NSCLC). To clarify the significance of mutation mutual exclusion information in the optimization of radiomics algorithms, this study aimed to take EGFR and KRAS as examples, to explore the influence of KRAS mutation information on the accuracy of radiomics algorithms for predicting EGFR mutation. Materials and Methods: We retrospectively analyzed 218 NSCLC patients with 18F-FDG PET/CT scans and results of EGFR and KRAS gene mutations. Patients were randomly divided into training and testing cohorts. The Pyradiomics toolkit was used for radiomics feature extraction. The gradient boosting decision tree (GBDT) algorithm was used to select features and develop a radiomics score(RS). A composite nomogram model combining PET/CT RS and KRAS mutation information were developed using logistic regression. the area under curve (AUC), specificity, sensitivity, accuracy, and precision were calculated for the model performance evaluation on the training and testing cohort. Results: Among the three models, the composite model exhibited the best performance. The composite nomogram model demonstrated highest AUC, accuracy and specificity in both training and testing cohort (AUC: 0.882 and 0.882, accuracy: 0.809 and 0.758, specificity: 0.868 and 0.879), significantly higher than CT RS model (AUC: 0.784 and