Validation of a harmonization method to correct for SUV and radiomic features variability in multi-center studies

Fanny Orlhac, Olivier Humbert, Sarah Boughdad, Maud Lasserre, Michael Soussan, Christophe Nioche, Nicholas Ayache, Jacques Darcourt, Frédérique Frouin, Irène Buvat

▶ **To cite this version:**

Fanny Orlhac, Olivier Humbert, Sarah Boughdad, Maud Lasserre, Michael Soussan, et al.. Validation of a harmonization method to correct for SUV and radiomic features variability in multi-center studies. SNMMI Annual Meeting, Jun 2018, Philadelphia, United States. pp.288. hal-01759334

HAL Id: hal-01759334
https://hal.science/hal-01759334

Submitted on 5 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Validation of a harmonization method to correct for SUV and radiomic features variability in multi-center studies

Fanny Orlhac[1,2], Olivier Humbert[3], Sarah Boughdad[2,4], Maud Lasserre[5], Michael Soussan[6,2], Christophe Nioche[2], Nicholas Ayache[1], Jacques Darcourt[3], Frédérique Frouin[2], Irene Buvat[2].

1: Epione, Inria Sophia-Antipolis, Valbonne, France, 2: IMIV, UMR 1023 Inserm/CEA-SHFJ/CNRS/Université Paris-Sud/Université Paris Saclay, Orsay, France, 3: Department of Nuclear Medicine, Centre Antoine Lacassagne/UCA, Nice, France, 4: Department of Nuclear Medicine, Institut Curie-René Huguenin, Saint-Cloud, France, 5: Department of Nuclear Medicine, Centre Georges-François Leclerc, Dijon, France, 6: Department of Nuclear Medicine, Hôpital Avicenne, Bobigny, France.

Objectives: Quantitative characterization of tumor heterogeneity using PET images shows encouraging results to predict patient response or survival, paving the way for precision medicine based on radiomic features. Yet, the vast majority of results are obtained in monocentric studies, without subsequent validation in multicenter settings. One reason is that radiomic features, as well as SUV measurements, are sensitive to the acquisition and reconstruction of PET images, hence a radiomic model established in one center might not be directly applicable to radiomic data from another center. This considerably reduces the dissemination of radiomic models. The goal of this study was to validate the use of the recently described ComBat harmonization method [1] to remove the center effect in a 3-center setting.

Methods: 18F-FDG PET images from 190 patients at initial staging for breast cancer from three centers were included. In center A, 63 patients were scanned using a Gemini TF PET/CT scanner (Philips, BLOB-OS-TF, voxel = 4x4x4 mm). In center B, 74 patients underwent a PET on a Discovery 690 PET/CT scanner (GE Healthcare, OSEM, 2.7x2.7x3.3 mm). In center C, 53 patients were scanned using a Gemini GXL PET/CT scanner (Philips, LORRAMLA, 2x2x1.9 mm). For each patient, the primary lesion was segmented using a threshold equal to 40% of SUVmax. In the resulting volume of interest, we computed SUVmax, SUVmean and 6 textural features (resampling step: 64 gray-levels between 0 and 20 SUV units) using the LIFEx software. The ComBat harmonization method was used to estimate the center effect based on the observed tumor feature values in the three patient groups. To validate its efficiency, for each feature, we investigated the differences between the 3 centers using Kruskal-Wallis test. We also determined the best cut-off value in center A to separate triple-negative (TN) from non-TN lesions by maximizing the Youden index (Sensitivity + Specificity -1). Then we applied this optimal cut-off value on the feature values of the other two centers before and after ComBat. Youden index was used to evaluate the accuracy of TN classification in each case.

Results: In total, 45 lesions were TN (24%) in proportions similar for all patient groups (25% in center A, 20% in center B and 26% in center C). We observed a shift in feature values between the three patient groups, likely due to the scanner effect with a significant difference between the three centers for 7/8 features (p-values < 0.05, Kruskal-Wallis test). After ComBat harmonization, the scanner effect could no longer be detected (p-values > 0.11 for all features). For instance, without harmonization, Homogeneity in center A was lower than Homogeneity in center B, which was itself lower than in center C. The p value of Kruskal-Wallis was less than 0.001 before ComBat and increased to 0.998 after harmonization for Homogeneity. In center A, Youden indices obtained for the 8 features were all between 0.29 and 0.54 for distinguishing TN from non-TN lesions when the cut-off values were determined based on data from center A. When these cut-off values were applied on data from centers B and C without ComBat harmonization, all Youden indices were between 0.04 and 0.29, corresponding to poor performance in TN identification. After ComBat harmonization, all Youden indices increased and were between 0.28 and 0.42, close to values obtained in center A only.

Conclusions: As demonstrated in this study, using ComBat harmonization enables to pool values of radiomic features and SUV measured in PET images acquired in different acquisition and reconstruction conditions. The method is data-driven (no phantom acquisition is needed) and is applied directly to the features values, without requiring any re-calculation of features. ComBat harmonization should facilitate the portability of radiomic models between centers and accelerate their use in clinical practice.