10.1148/radiol.2019190357

Appendix E1

Methods

FDG-PET/CT

Patients fasted for a minimum of 6–8 hours, and blood glucose levels were assessed prior to administration of 12–17 millicuries of FDG. Patients with elevated blood sugar levels were either rescheduled or scanned at the discretion of the attending nuclear medicine physician. Approximately 45–60 minutes after injection, patients were scanned from the skull base to midthigh using bed positions acquired every 2–5 minutes. Continuous axial slices of 3–5 mm thickness was obtained. Noncontrast CT was obtained, and attenuation correction was performed using ordered subset expectation maximization reconstruction. Images were converted to SUV units normalized by patient body weight.

Region of Interest Segmentation

The segmentation threshold for deriving bone segmentations was adjusted on a per-patient basis due to differences in quality of the CT. Manual editing was performed to remove the intervertebral disc space. A single segmentation was performed for each patient in the study by the lead author.

The PET-edge gradient-based semiautomatic segmentation tool was used to contour the MTV which defines the tumor edge based on spatial derivative levels and continuity of the tumor edge. A single distance of 1 cm was chosen and analyzed for the penumbra distance. This was intuitively chosen to avoid oversampling the normal tissue outside the tumor, while sampling enough surrounding voxels (given the large 3–5 mm voxel sizes of the PET image).

Feature Extraction

The total number of features calculated within the MTV, penumbra, MTV plus penumbra, and bone marrow regions are shown in Table 2. In summary, a total of 668 features were extracted including size, sphericity, local volume invariant integral (LVII) shape, histogram intensity, and gray-level co-occurrence matrix (GLCM) texture. These features were chosen to cover a range of feature types and without any predetermination as to their likelihood of their ability to predict lung cancer outcomes.

All texture features were calculated in 3D (13 directions, symmetrical GLCM) with a fixed bin size of 0.2 SUV (28). A total of 12 statistics (trimmed mean (90%), mean, median, kurtosis, skewness, standard deviation, variance, range, maximum, minimum, interquartile range, and mean absolute deviation) were assessed on each of the 12 GLCM features (energy, entropy, correlation, contrast, homogeneity, variance, sum mean, inertia, cluster shade, cluster tendency, maximum probability, and inverse variance) across all directions, resulting in a total of 144 texture features for analysis. The size and shape features were not calculated within the bone marrow, penumbra or within the combined MTV and penumbra. Since the penumbra regions were generated directly from the MTV, the size and shape measures would not be independent measures due to correlation with the size and shape features of the MTV. With regard to the bone

marrow, we were only interested in the metabolic activity as an indicator of malignancy and not the size and shape of the vertebral bodies. Feature values were only extracted once with a prespecified set of parameters by the lead author. Parameter tuning or optimization was not performed in this preliminary study. The feature extraction pipeline used in this study was developed in house and is publicly available on GitHub (https://github.com/riipl/3d_qifp, commit 94c3d12). Prior to clinical implementation, FDA approval and further validation of this software would be needed.

Clinical Features

A total of 11 clinical features were included in this analysis. This included 7 features describing patient demographics, including age, sex, tumor location, histology, smoking status (never and past), cancer stage. We also analyzed 4 blood parameters including white blood cells, hemoglobin, hematocrit, and platelets.

Feature Model Building and Selection

The least absolute shrinkage and selection operator (LASSO) was used for model development since it minimizes possible overfitting of the data by removing highly correlated imaging features, and it allows one to assess the time to event with Cox regression. The optimal tuning parameter, λ, was selected using 100 randomizations of seven-fold cross-validation. Error curves for each λ were averaged across the randomizations and the λ with the minimum error was selected. We set the regularization parameter (alpha) to 1 reduce overfitting by shrinking most of the coefficients to zero to minimize the number of features selected.