SUPPLEMENTAL APPENDIX 1

Let x_i=(*PET1(i), DIFF(i)*) be a 2-component at voxel *i*, where *i*=1,...,*N*, *PET1(i)* is the initial image SUV at voxel *i* and *DIFF(i)=PET2(i)-PET1(i)* is the subtraction image SUV at voxel *i*. The data $(x_1,...,x_N)\in\mathbf{R}^2$ are assumed to be independent and identically distributed with probability density functions :

$$f(x_i)=\sum_{k=1}^{K} p_k f(x_i|\mu_k,V_k)$$

where K represents the number of different evolutions of tissues in the images, each being characterized by a Gaussian probability density function $f(x_i|\mu_k,V_k)$ parameterized with a mean vector μ_k and a variance matrix V_k, and p_k are the mixing proportions ($0<p_k<1$ for all k=1,...,*K* and $\Sigma_k p_k=1$).

Let C=(c_{ik},i=1,...,*N*;k=1,...,*K*) be a partition with c_{ik}=1 if x_i belongs to cluster k and 0 otherwise.

With this convention, the parameter $\Pi=(p_1, ..., p_{k-1}, \mu_1, ... , \mu_k, V_1,..., V_k)$ is determined so as to maximize the log-likelihood:

$$L(\Pi|x_1,...,x_N,c_{ik})=\sum_{i=1}^{N}\ln\left[\sum_{k=1}^{K} c_{ik}p_k f(x_i|\mu_k,V_k)\right]$$

Starting from an initial parameter Π^0 and a partition C^0, an iteration of the algorithm consists in calculating the current conditional probabilities $t_k(x_i)$ ($1\le i\le N;1\le k\le K$) for the *kth* mixture component and the current value of Π (E step):

$$t_k(x_i)=\frac{p_k f(x_i|\mu_k,V_k)}{\sum_{m=1}^{K} p_m f(x_i|\mu_m,V_m)}$$

The M step then estimates (p_k, μ_k, V_k) using the cluster k and the conditional probability $t_k(x_i)$.

SUPPLEMENTAL APPENDIX 2

The biparametric graph (*PET1, DIFF*) is analyzed considering the Euclidian distance between points (corresponding to voxels). The 4 possible clusters (noise, physiological changes, increasing SUV and decreasing SUV) are systematically initialized as follows:

- noise (voxels with no substantial SUV changes between the two PET): the 100 voxels closest to the origin of the biparametric graph (cluster O)

- physiological changes not related to tumor masses: the 10 voxels closest to coordinates (50;0) (cluster F)

- tumor voxels in which SUV increased between the two scans: the 5 voxels closest to the voxel nearest the (0;50) coordinates (cluster P)

- tumor voxels in which SUV decreased between the two scans: the 5 voxels closest to the voxel nearest the (0;-50) coordinates (class N).

Let $(\bar{x}_k, \bar{y}_k)$ the center of mass of each class k (k=O, F, P, N).

The consistency of the resulting classes is checked as follows:

- If $(\bar{x}_N > 0$ or $\bar{x}_P < 0)$, then N (or P) is eliminated, meaning that it is considered that there is no N (or P) cluster in the data.

- If $\bar{y}_F > 0$ and $\sqrt{(\bar{x}_F - \bar{x}_O)^2 + (\bar{y}_F - \bar{y}_O)^2} < \sqrt{(\bar{x}_P - \bar{x}_O)^2 + (\bar{y}_P - \bar{y}_O)^2}$ and $\sqrt{(\bar{x}_F - \bar{x}_P)^2 + (\bar{y}_F - \bar{y}_P)^2} < \sqrt{(\bar{x}_P - \bar{x}_O)^2 + (\bar{y}_P - \bar{y}_O)^2}$, cluster F is eliminated.

- If $\bar{y}_F < 0$ and $\sqrt{(\bar{x}_F - \bar{x}_O)^2 + (\bar{y}_F - \bar{y}_O)^2} < \sqrt{(\bar{x}_N - \bar{x}_O)^2 + (\bar{y}_N - \bar{y}_O)^2}$ and $\sqrt{(\bar{x}_F - \bar{x}_N)^2 + (\bar{y}_F - \bar{y}_N)^2} < \sqrt{(\bar{x}_N - \bar{x}_O)^2 + (\bar{y}_N - \bar{y}_O)^2}$, cluster F is eliminated.

After solving GMM algorithm, the consistency of the final clusters is checked again. If cluster P or N includes more than 20% of the total number of voxels or includes as many voxels as cluster O, then this cluster corresponds to noise or physiological uptake and the classification is restarted assuming there is no cluster P (or N).

Finally, if cluster P (or N) crosses the x-axis and is longer (in the x-axis) than wide (in the y-axis), then this cluster

corresponds to physiological changes. This cluster P (or N) is removed from the initial partition and the classification is restarted without cluster P (or N).