4D-PET RECONSTRUCTION OF DYNAMIC NON-SMALL CELL LUNG CANCER [18-F]-FMISO-PET DATA USING ADAPTIVE-KNOT CUBIC B-SPLINES

G. P. Ralli[1], D. R. McGowan[1,2], M. A. Chappell[3], R. A. Sharma[4], G. S. Higgins[1], J. D. Fenwick[5]

[1]Department of Oncology, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, [2]Radiation Physics and Protection, Oxford University Hospitals NHS Foundation Trust, Churchill Hospital, Oxford, OX3 7LE, [3]Institute of Biomedical Engineering, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, [4]UCL Cancer Institute, University College London, 72 Huntley Street, London WC1E 6DD, [5]Clatterbridge Cancer Centre and Institute of Translational Medicine, University of Liverpool, UCD Block, Royal Liverpool University Hospital, Daulby Street, Liverpool L69 3GA.

ABSTRACT

4D-PET reconstruction has the potential to significantly increase the signal-to-noise ratio in dynamic PET by fitting smooth temporal functions during the reconstruction. However, the optimal choice of temporal function remains an open question. A 4D-PET reconstruction algorithm using adaptive-knot cubic B-splines is proposed. Using realistic Monte-Carlo simulated data from a digital patient phantom representing an [18-F]-FMISO-PET scan of a non-small cell lung cancer patient, this method was compared to a spectral model based 4D-PET reconstruction and the conventional MLEM and MAP algorithms. Within the entire patient region the proposed algorithm produced the best bias-noise trade-off, while within the tumor region the spline- and spectral model-based reconstructions gave comparable results.

Index Terms— B-splines, Dynamic PET, Expectation Maximization, NSCLC, Regularization

1. INTRODUCTION

Iterative 4D positron emission tomography (PET) reconstruction improves the signal to noise ratio in dynamic-PET image sequences by fitting temporally smooth functions to the time-activity curve (TAC) of each image voxel during the reconstruction. In most cancer imaging studies, many different tissues are present in the field of view of the scanner, leading to a wide range of TAC shapes. Bias from poorly fitted regions can spatially propagate to well modeled regions during the reconstruction [1], meaning that highly flexible temporal functions are required. The optimal choice of temporal function for 4D-PET reconstruction in situations with a large range of tissues remains an open question in the field.

Linear kinetic models represent TACs of given voxels as linear combinations of pre-defined basis functions,

$$x_{jm}(\boldsymbol{\theta_j}) = \sum_{k=1}^{n_k} B_{km}\theta_{jk}, \quad (1)$$

where x_{jm} is the radioactivity concentration of voxel j in time frame m, n_k is the total number of basis functions, θ_{jk} is the weighting factor of basis function k in frame m, B_{km}, and $\boldsymbol{\theta_j}$ is a vector containing all parameter values in voxel j. Since the basis functions are continuous in time, the value for a given time frame is calculated by averaging these functions across the time frame. Linear kinetic models allow a wide range of TAC shapes to be fitted, given an appropriate choice of basis functions, making them an appropriate choice for 4D-PET reconstruction.

In situations with a wide range of tissues present in the image, the spectral model is commonly used [1, 2]. This is based on the general solution to the compartment model equations, which model each TAC as a tissue response function (a weighted sum of exponential decays) convolved with the arterial input function (AIF). In the spectral model, the basis functions are calculated by defining a range of exponential decay functions (by varying the decay constants) and convolving them with the measured AIF.

Another linear model with the potential to fit a wide range of TACs are B-spline functions. These are defined on a knot vector $\boldsymbol{\tau} = \{\tau_0, \tau_1, \dots, \tau_{N+p}\}$, where the knots are listed in ascending order, and are calculated using the Cox-de Boor recurrence relations [3]. Spline functions of order p (linear combinations of B-splines) and their derivatives up to order p - 1 are continuous at the knots. A notable B-spline based 4D-PET reconstruction framework was proposed by Nichols et. al. [4], with knots placed along equal arc-lengths of the head curve (a temporal histogram of all list-mode data). While B-spline fits do not directly yield any physiologically relevant information, images produced with spline-based 4D-PET reconstruction are likely to contain less noise without a substantial increase in the bias. Fitting compartment models to tumor TACs obtained from images reconstructed using such methods may therefore yield improved kinetic parameter estimates, compared with conventional reconstruction algorithms.

In this work a 4D-PET reconstruction algorithm using adaptive-knot cubic B-splines is proposed. This differs from the methodology developed in [4] by using the nested-EM framework of Wang and Qi [5], which accelerates the convergence rate of 4D-PET reconstruction. Furthermore, the proposed method uses a theoretically-driven adaptive knot placement algorithm to optimize the knot vector for each voxel TAC, rather than using the same knot sequence for all voxels, as was done in [4]. This method is compared to both conventional image reconstruction algorithms and 4D-PET reconstruction with the spectral model.

2. 4D-PET RECONSTRUCTION WITH LINEAR KINETIC MODELS

The expected number of photon counts detected by detector pair i in frame m, $\langle y_{im} \rangle$, is defined as

$$\langle y_{im}(\boldsymbol{\theta}) \rangle = \sum_{j=1}^{n_j} \sum_{k=1}^{n_k} P_{ij} B_{km}\theta_{jk} + \varepsilon_{im}, \quad (2)$$

where n_j is the total number of voxels, ε_{im} is the number of erroneously measured counts (random coincidences and scattered photons) and P_{ij} are the elements of an $n_i \times n_j$ matrix $\boldsymbol{P}$, called the system matrix, where n_i is the number of detector pairs. These components represent the probability of a photon pair originating from voxel j being detected by detector pair i.

The counts y_{im} can be accurately modelled as independent Poisson-distributed variables. The Poisson log-likelihood function

of the scanner data, $L(\boldsymbol{y}|\boldsymbol{\theta})$, shows the log of the probability that a given estimate of the parameters could have generated the measured count data:

$$L(\boldsymbol{\theta}|\boldsymbol{y}) = \sum_{m=1}^{n_m}\sum_{i=1}^{n_i} y_{im}\ln(\langle y_{im}(\boldsymbol{\theta})\rangle) - \langle y_{im}(\boldsymbol{\theta})\rangle, \quad (3)$$

where n_m is the number of time frames. The parameters can then be calculated by maximizing $L(\boldsymbol{\theta}|\boldsymbol{y})$. In this work, this was done using the iterative nested-EM reconstruction algorithm [5].

In this approach intermediate images $\hat{x}_{jm}^{n+1}$ (where n is the iteration number) are first obtained using the image update equation for the maximum-likelihood expectation maximization (MLEM) reconstruction algorithm, which is widely used in conventional 3D-PET reconstruction. The parameters are then fitted using n_l iterations of

$$\theta_{jk}^{n,l+1} = \frac{\theta_{jk}^{n,l}}{\sum_m B_{mk}}\sum_m B_{mk}\frac{\hat{x}_{jm}^{n+1}}{x_{jm}\left(\boldsymbol{\theta}_j^{n,l}\right)}, \quad l=1, ..., n_l, \quad (4)$$

where $\theta_{jk}^{n+1} = \theta_{jk}^{n,n_l+1}$, and $\boldsymbol{x}(\boldsymbol{\theta^{n+1}})$ is used as the image estimate for the next iteration.

In order to incorporate spatial regularization into the reconstruction, a modified objective function, $\Phi(\boldsymbol{\theta}|\boldsymbol{y})$, can also be used,

$$\Phi(\boldsymbol{\theta}|\boldsymbol{y}) = L(\boldsymbol{\theta}|\boldsymbol{y}) - \beta U(\boldsymbol{x}(\boldsymbol{\theta})), \quad (5)$$

where β is a tunable parameter controlling the trade-off between resolution loss and noise suppression and $U(\boldsymbol{x}(\boldsymbol{\theta}))$ is a concave function that is designed to penalize rough images,

$$U(\boldsymbol{x}) = \frac{1}{4}\sum_j\sum_{k\in\mathcal{N}_j} w_{jk}\psi(x_j - x_k), \quad (6)$$

where $\mathcal{N}_j$ is the set of nearest neighbours of voxel j and w_{jk} is a weighting factor equal to the normalized inverse distance between voxels j and k. In this work the Lange function [6],

$$\psi(\xi) = \delta\left(\frac{|\xi|}{\delta} - \ln\left(1+\frac{|\xi|}{\delta}\right)\right), \quad (7)$$

was used, where δ is an additional smoothing parameter. This function allows a good level of noise suppression in fairly uniform regions, while also preserving edges better than the more widely used quadratic function, $\psi(\xi) = \xi^2$ [6].

Reconstruction algorithms designed to maximize (5) are generally referred to as penalized-likelihood or maximum a posteriori (MAP) algorithms. Spatial noise suppression can be incorporated into the nested-EM framework by performing the image update step with the 3D-MAP algorithm. Such an approach is referred to in this work as a nested-MAP algorithm.

3. ADAPTIVE-KNOT PLACEMENT ALGORITHM

Although spline functions with closely spaced knots are highly flexible, their use can lead to over-fitting of the data. We propose an adaptive-knot placement algorithm, based on the shape of the TAC to be fitted, which allows TACs to be accurately fitted with spline functions, using fewer basis functions than if the knots were uniformly spaced.

For a spline of order p, defined on the interval $[a, b]$, the first and last knots are usually repeated with multiplicity p, such that $\tau_0 = ... = \tau_{p-1} = a$ and $\tau_{N+1} = ... = \tau_{N+p} = b$, to allow for discontinuities at the beginning and end of the curve. What remains is to choose the internal knots.

For a given function $f(t)$ and the set of all splines of order p with m internal knots, $\tilde{S}_{p,m}$, the distance operator is defined as

$$\mathrm{dist}(f, \tilde{S}_{p,m}) := \min\{\|f - s\| : s \in \tilde{S}_{p,m}\}, \quad (8)$$

where $\|f - s\| = \max_{a\le t\le b}|f(t) - s(t)|$ is the maximum difference between spline s and f. Theorem XII.5 of [3] shows that for a p times differentiable function $f(t)$,

$$\mathrm{dist}(f, \tilde{S}_{p,m}) \le \Omega_p m^{-p}\left(\int_a^b |f^{(p)}(t)|^{\frac{1}{p}}dt\right)^p, \quad (9)$$

where Ω_p is a constant depending only on p and $f^{(p)}(t)$ is the p^{th} order derivative of f. Choosing the m internal knots such that

$$\int_{\tau_i}^{\tau_{i+1}} |f^{(p)}(t)|^{\frac{1}{p}}dt = \frac{1}{m}\int_a^b |f^{(p)}(t)|^{\frac{1}{p}}dt, \quad (10)$$

therefore gives an approximation error of order m^{-p}, for noise free data. For noisy data adding more than a certain number of knots results in over-fitting, and therefore a worse description of the data.

Calculating the knots for dynamic-PET data using (10) requires function derivatives to be estimated from discrete, noisy data points. We have used a Savitzky-Golay filter [7] (with weighed data points) to suppress noise amplification when calculating the derivatives.

In this methodology a subset of k (an odd number) adjacent data points $\{t_i, x_i\}$, with t_i and x_i being the independent variable and noisy data points respectively, are fitted with a low degree polynomial of degree q, using weighted least squares. In dynamic-PET the weighting factor for time frame m is commonly defined as

$$w_m = \frac{1}{\sigma_m^2} = \frac{\Delta T_m e^{-\lambda t_m}}{x_m}, \quad (11)$$

where σ_m^2 is the variance, ΔT_m is the duration of frame m, with mid-time t_m, λ is the decay constant of the radioisotope.

A change of variable, $z = t - t_c$ (where t_c is the t_i value for the central data point in the subset) is made, prior to the fitting of the polynomial $X(z)$,

$$X(z) = a_1 + a_2 z + \cdots + a_q z^{q-1}. \quad (12)$$

For cubic (order 4) splines, the 4th order derivative is required in order to use (10) to place the knots, meaning $X(z)$ must be quartic or higher ($q \ge 5$). The value of the N^{th} derivative at the central data-point (i.e $z = 0$) of this subset is given by a_{N+1}. This process is used to calculate the derivatives at each of the data points.

This method does not allow the derivative of the first and last $\frac{k-1}{2}$ points to be calculated as they will never fall into the centre of a subset of k adjacent points. This can be accounted for by adding, in reverse order, copies of the first $\frac{k-1}{2}$ points at the beginning and the last $\frac{k-1}{2}$ points at the end, as recommended in [7].

The knot placement algorithm (for a fixed number of knots) used in this work is summarized by the four steps below.

1. Place four repeated knots at the beginning and end of each TAC.
2. Identify the point where the TAC initially begins to rise and place 4 repeated knots there (to handle the discontinuity).
3. Calculate the 4th derivative of all points after the initial rise using the Savitzky-Golay method (k=11, q=5).
4. Place the remaining knots according to (10).

In the cubic B-spline-based 4D-PET reconstruction algorithm proposed in [4] the knots are placed along equal arc lengths of the TAC, which is approximately equivalent to using (10) with p = 2. According to (9) this is suboptimal for cubic B-splines.

4. METHODS

4.1 Digital Phantom and Data Generation

The 4D-XCAT2 digital phantom [8] was used to simulate a non-small cell lung cancer (NSCLC) patient, injected with the [18-F]-FMISO hypoxia tracer, using time activity curves (TACs) obtained from a clinical dynamic FMISO-PET image of a NSCLC patient, with voxel dimensions 3.1×3.1×2.0 mm. Spherical regions of interest (ROIs) with a diameter of 3 cm were drawn in healthy lung

and liver tissue, and the spine. A further ROI was defined in the central part of the descending aorta for five consecutive PET axial slices, and the TAC from this region was used to represent the AIF. Additional ROIs were drawn in three separate regions believed to contain normoxic, hypoxic and necrotic tumor tissue.

The noisy TACs obtained from the 'healthy tissue' ROIs were fitted with adaptive-knot cubic B-splines, the AIF was fitted using the three exponential model of Feng et. al. [9] and the tumor TACs were fitted with a 3-compartment 5-parameter (3C5K) compartment model. These fitted curves were then binned into the following time-frame sequence: [1x30 s, 6x5 s, 6x20 s, 7x60 s, 10x120 s, 3x300 s], and two additional 600 s frames 2 and 4 hours post-injection. These were taken to be the 'ground truth' TACs, and are shown in figure 1, along with images of the final time frame of the phantom.

The PET-SORTEO Monte-Carlo simulator [10] was used to generate PET sinograms representative of those obtained using an mMR PET-MR scanner (Siemens Healthcare) [11] for each time frame of the image sequence, including effects of scattered photons, random co-incidences and attenuation. 50 noise realizations of dynamic-PET sinogram data were simulated from the phantom. Since the aim of the current study is to evaluate the effectiveness of noise suppression using the proposed 4D-PET framework, as a simplifying assumption no patient motion was simulated.

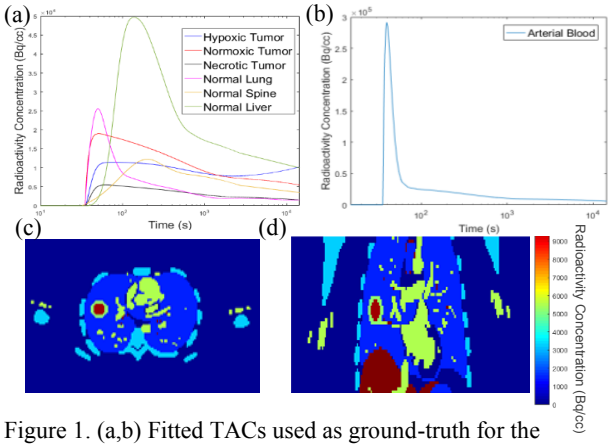

Figure 1. (a,b) Fitted TACs used as ground-truth for the phantom generation, (c) axial and (d) coronal slices of the final phantom time frame.

4.2 Reconstruction Algorithms

An attenuation correction sinogram was calculated by forward-projecting an attenuation map of the patient phantom. A normalization correction sinogram was generated by performing a 20 minute simulated scan of a uniform cylindrical phantom with PET-SORTEO. The attenuation and normalization corrections were modelled as part of the system matrix. Random co-incidences were estimated using a delayed co-incidence window, and scattered photons were estimated with the single scatter simulation algorithm [12], using the attenuation map.

Images were reconstructed using both MLEM and MAP, in order to allow the 4D-PET reconstructions to be compared to more conventional reconstruction algorithms. 4D-PET reconstructions were performed using the nested-MAP algorithm with both adaptive-knot splines and the spectral model.

Voxel-specific knots were calculated on images reconstructed with the MAP algorithm, using the adaptive knot algorithm described in section 3. 11 free knots were used (not including the 4 repeated knots placed at the beginning and end of each TAC). This number was determined by fitting adaptive knot splines to simulated tumor TACs and input functions, derived from fits to real patient TACs. 11 knots minimized the Akaike information criterion (AIC) for the input functions (the fastest varying functions) while not leading to significant increases in the RMS error when fitting tumor TACs.

Another cubic B-spline based nested-MAP reconstruction was performed, using one knot sequence for all voxels, obtained by placing 11 knots on the head-curve using the algorithm described in section 3. Since the current version of PET-SORTEO is unable to produce list-mode data, an approximation to the head curve was calculated by summing the events in each sinogram time frame.

For the spectral model-based reconstruction, 100 basis functions were used. The exponential decay constants were logarithmically spaced between values of the half-life decay and 0.01 s^{-1}. The input function was estimated by drawing a large ROI in the left-ventricle region of the MAP images and fitting the TAC from this region with the three exponential Feng function [9].

A further idealized spectral-model based nested-MAP reconstruction was performed by firstly fitting the spectral basis functions to the ground truth image, and removing any non-contributing basis functions from a given voxel. This represents the 'best possible' optimization method one could apply to the spectral model in order to avoid over-fitting the data by using fewer basis functions. While this method could not be used practically, it allows the proposed spline-based reconstruction algorithm to be compared to both an un-optimized and a 'perfectly' optimized spectral model.

Each algorithm was run for 30 iterations, and the nested-MAP algorithms each used 30 sub-iterations, to ensure convergence.

4.3 Image Quality Metrics

At every iteration of each reconstruction method, images of the bias, Δ, and the weighted standard deviation, σ_w, were calculated:

$$\Delta_i = \frac{1}{T_{scan}} \int_0^t (\bar{x}_i - x_i^{true}) dt \approx \frac{1}{T_{scan}} \sum_m \Delta T_m (\bar{x}_{im} - x_{im}^{true}), \quad (13)$$

$$\sigma_{wi}^m = \left(\frac{\sigma_{m,measured}^2 \Delta T_m}{x_{im}^{true}} \right)^{\frac{1}{2}}, \quad (14)$$

where $\bar{x}_i$ is the mean value of all 50 repeats in voxel i, x_i^{true} is the true value of voxel i, ΔT_m is the time frame duration of frame m, T_{scan} is the scan duration and $\sigma_{measured}^2$ is the measured variance of all 50 repeats. σ_{wi}^m was averaged over all time frames to get σ_{wi}.

5. RESULTS AND DISCUSSION

Plots of bias versus σ_w, averaged over both the whole-patient and the tumor region, are shown in figure 2. TACs for a single noise realization of both the entire hypoxic tumor ROI and a single voxel at the center of the tumor are plotted for the best performing conventional reconstruction algorithm (MAP) and the nested-MAP algorithms with splines (using voxel-specific knots) and optimized spectra in figure 3.

The results in figure 2 show that all the nested-MAP reconstructions are less noisy than the conventional methods, though the un-optimized spectra are also considerably more biased. The MAP algorithm, which includes spatial regularization only, significantly outperforms the MLEM algorithm, which has no regularization. The nested-MAP reconstruction (using splines with voxel-specific knots) gives the best bias-noise trade-off in the

whole patient region, while this and the optimized spectra give similar results in the tumor region.

The results in figure 3 show both the spline- and optimized spectral model nested-MAP reconstructions produce similar TACs at the ROI level, and both significantly outperform the MAP algorithm, particularly in the earlier time frames. At the early time points, the optimized spectral model appears able to capture the shape of the tumor TAC better than the spline functions at the noisier single voxel level. It is important to note that the spectral basis functions were artificially optimized to represent the 'best possible' optimization, and that a more realistic optimization method may provide worse fits in practice.

In our framework, 4D-PET reconstruction is used purely to suppress noise, with the kinetics analysis done post-reconstruction. An alternative 4D-PET framework designed to work in regions with many tissues present, proposed by Kotasidis et. al. [1], adaptively fits a 'primary' kinetic model of interest followed by a more flexible 'secondary' model to fit the residuals in regions where the primary model performs poorly. This limits the effects of bias propagation during the reconstruction while allowing kinetic parameters of interest to be directly obtained from the reconstruction. However, the accuracy of this method is strongly influenced by the choice of secondary model. The results in figure 2 (a) suggest that adaptive-knot splines may be a superior choice of secondary model to the spectral model (used in [1]), due to the ability of the splines to fit all the TACs in the patient region with a lower overall bias without increasing the noise level.

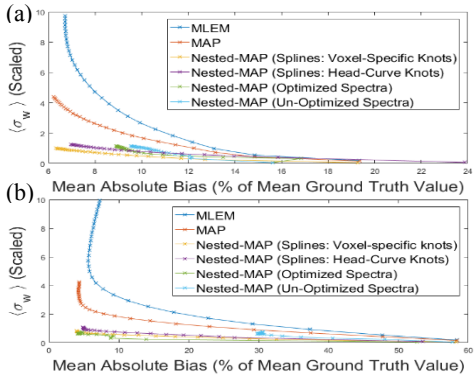

Figure 2. Plots of bias vs σ_w at every iteration, averaged over (a) the entire patient region and (b) the tumor region. σ_w is expressed as a fraction of the best performing algorithm.

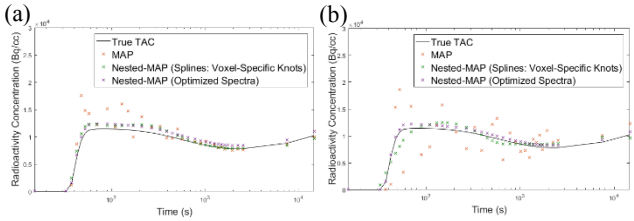

Figure 3. Comparison of hypoxic tumor TACs (single noise realization) from (a) an ROI and (b) a single voxel.

6. CONCLUSION AND FUTURE WORK

Of all the methods tested, 4D-PET reconstruction with adaptive-knot cubic B-splines (using voxel-specific knots) was shown to produce the best bias-noise trade-off in the entire patient region, and to provide comparable results to the optimized spectral model in the tumor region. The optimized spectral model captured the TAC shape best in the single voxel region, though this is possibly due in part to the artificial optimization method used.

The accuracy of compartment model fits to TAC data from the methods used here, as well as the applicability of these methods to clinical datasets, will be investigated in future work.

7. ACKNOWLEDGEMENTS

This work is funded by the CRUK-EPSRC Oxford Cancer Imaging Centre (OCIC), and has been carried out using the University of Oxford Advanced Research Computing facility. We thank Anthonin Reilhac for providing access to the PET-SORTEO simulator, and assisting with its implementation. RAS is supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

The patient data used in this study was collected for the BKM120 trial (EudraCT 2012-003762-40) and we thank the patients who took part and the OUH NHS Foundation Trust. The trial is sponsored by the University of Oxford, supported by Novartis Pharmaceuticals UK Ltd is managed by the Oxford Clinical Trials and Research Unit and was supported by CRUK grant number C34326/A15163 (www.cancerresearchuk.org) with further funding support from the Oxford ECMC, the OCIC, the CRUK Oxford Centre and the NIHR.

8. REFERENCES

[1] F. Kotasidis et. al. "Application of adaptive kinetic modelling for bias propagation reduction in direct 4D image reconstruction", *Phys. Med. Biol.*, vol. 59, no. 20, pp. 6061–6084, 2014.

[2] V. J. Cunningham and T. Jones, "Spectral analysis of dynamic PET studies", *J. Cerebral Blood Flow Metabol.*, vol. 13, pp. 15–23, 1993.

[3] C. De Boor. *A Practical Guide to Splines*, Applied Mathematical Sciences vol. 27, Springer-Verlag, New York, 1978.

[4] T. E. Nichols et. al., "Spatiotemporal Reconstruction of List-Mode PET Data". *IEEE Trans. Med. Imag*, vol. 21, no. 4, pp. 396–404, 2002.

[5] G. Wang and J. Qi. "Acceleration of the direct reconstruction of linear parametric images using nested algorithms", *Phys. Med. Biol.*, vol. 55, no.5, pp. 1505–1517, 2010.

[6] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing, *IEEE Trans. Med. Imag.*, vol. 9, no. 4, pp. 439–446, 1990.

[7] A. Savitzky and M. Golay, "Smoothing and differentiation of data by simplified least squares procedures", *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[8] W. P. Segars et. al., "4D XCAT phantom for multimodality imaging research". *Med. Phys.*, vol. 37, no. 9, pp. 4902–4915, 2010.

[9] D. Feng et. al., "Models for computer simulation studies of input functions for tracer kinetic modelling with positron emission tomography", *Int. J. Biomed. Comput.*, vol. 32, pp. 95–110, 1993.

[10] A. Reilhac et. al., PET-SORTEO: "A Monte-Carlo-based simulator with high count rate capabilities". *IEEE Tras. Nucl. Sci.*, vol. 51, no. 1, pp. 46–52, 2004.

[11] A. Reilhac et. al., "Validation and application of PET-SORTEO for the geometry of the Siemens mMR scanner", PSMR Conference, Cologne, Germany, 2016.

[12] C. C. Watson, "New, faster, image-based scatter correction for 3D-PET", *IEEE Tras. Nucl. Sci.*, vol. 47, no. 4, pp 1587-1594, 2000.