4D Reconstruction of Oncological Dynamic PET Data



George Ralli

Oriel College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2018

Abstract

4D reconstruction of dynamic positron emission tomography (dPET) data suppresses noise in both the reconstructed image sequences and kinetic parameter estimates by fitting smooth temporal functions to the voxel time-activity-curves (TACs) during the reconstruction. Cancer imaging studies often have many tissues within the scanner field of view, leading to a wide range of TAC shapes to be fitted. The development of flexible temporal functions for use in such situations is an ongoing area of research, and forms the primary focus of this thesis.

Two commonly used models for 4D-PET reconstruction in situations with diverse kinetics are the spectral model, which models TACs as weighted sums of convolutions of the arterial input function (AIF) with exponential decays, and spline functions. In a digital phantom study comparing the performance of 4D-PET reconstruction based on adaptive-knot cubic B-splines and the spectral model, the spline-based reconstruction produced the least biased images overall, though the spectral model-based reconstruction provided a better bias-noise trade-off in the early time-frames.

A spline-residue model describing TACs as weighted sums of convolutions of the AIF with cubic B-spline functions is then proposed. As with the spectral model, convolution with the AIF constrains the spline-residue model at early time-points, potentially enhancing noise suppression in early time-frames, while still allowing a wide range of TAC descriptions over the entire imaged time-course, thus limiting bias. In a digital phantom study comparing 4D-reconstructions based on the spline-residue model, adaptive-knot cubic B-splines, the spectral model and an irreversible two-tissue compartment model, spline-residue based 4D reconstruction produced the highest quality (lowest bias and noise) parametric maps of radiotracer uptake kinetics and was particularly effective at suppressing noise in the early time-frames.

To investigate whether the proposed spline-reside model gives physiologically meaningful kinetic parameters, fits of the spline-residue model to FDG-dPET tumour TACs obtained from a dataset comprising 41 primary breast cancer patients, who also underwent tumour biopsies, are compared to fits of irreversible two- and three-tissue compartment models. These TACs were obtained from analytically reconstructed dPET images. Using goodness-of-fit measures, the spline-residue model was found to give the best description of the patient TACs, and after accounting for multiple hypothesis testing the flux constant from the spline-residue model produced the strongest correlations with gene expression data derived from the tumour biopsies.

Acknowledgements

I would firstly like to thank my supervisors John Fenwick, Michael Chappell, Ricky Sharma and Rebecca Muirhead, for their support throughout the course of this project. I owe a special thanks to John, my primary supervisor, for his advice and guidance, as well as for many interesting and thought-provoking discussions. I'm also grateful to Michael, who agreed to join my supervisory team and provide local support after John moved to Liverpool. I'd also like to thank the Oxford Cancer Imaging Centre for funding this project.

In addition to my supervisors, I'd like to thank everyone who gave me access to their clinical datasets: Geoff Higgins and Daniel McGowan (FMISO-PET data) and Simon Lord (Metformin trial data). Simon Lord, Francesca Buffa and Wei-Chen Cheng also provided me with very useful advice concerning the use of gene expression data, for which I'm grateful. I'd also like to thank Anthonin Reilhac of CERMEP, Lyon, France, who kindly provided me with the PET-SORTEO Monte-Carlo simulation code and assisted with its implementation.

Finally, I'd like to thank my friends and colleagues in Oxford, as well as my family, for their support and encouragement throughout this project.

Contents

List of Figures vii

List of Tables xv

List of Abbreviations xviii

1 Introduction 1
- 1.1 PET Imaging in Oncology . . . 1
- 1.2 Dynamic PET Imaging . . . 3
  - 1.2.1 Static vs. Dynamic PET . . . 3
  - 1.2.2 Limitations of Dynamic PET . . . 4
  - 1.2.3 4D-PET Reconstruction . . . 5
- 1.3 Thesis Outline . . . 7
- 1.4 Journal and Conference Papers . . . 9

2 The PET Image Reconstruction Problem 10
- 2.1 PET Scanner Data . . . 10
  - 2.1.1 Data Acquisition and Storage . . . 10
  - 2.1.2 Data Correction . . . 14
- 2.2 Analytical Image Reconstruction . . . 17
  - 2.2.1 Backprojection . . . 17
  - 2.2.2 Central Section Theorem . . . 18
  - 2.2.3 Filtered Backprojection . . . 19
- 2.3 Iterative Image Reconstruction . . . 21
  - 2.3.1 System Model for Iterative Reconstruction . . . 21
  - 2.3.2 Maximum-Likelihood Expectation-Maximization . . . 24
  - 2.3.3 Ordered Subsets Expectation-Maximization . . . 26
  - 2.3.4 Bayesian Image Reconstruction . . . 28
- 2.4 4D Reconstruction for Dynamic PET . . . 30
  - 2.4.1 Extended Maximum-Likelihood Framework . . . 31
  - 2.4.2 Acceleration with Nested Algorithms . . . 34
  - 2.4.3 Temporal Models for 4D-PET Reconstruction . . . 37
  - 2.4.4 Spatial Bias Propagation . . . 40

3 Validation Framework for 4D-PET Reconstruction Algorithms 42
3.1 Introduction . . . 42
3.2 Data Generation . . . 43
3.2.1 NEMA NU2-Like Image Quality Phantom . . . 43
3.2.2 Dynamic FMISO NSCLC Phantom . . . 45
3.2.3 Generation of Monte-Carlo Simulated Data . . . 48
3.3 Image Reconstruction and Analysis . . . 49
3.3.1 Data Corrections . . . 49
3.3.2 System Matrix Calculation . . . 50
3.3.3 Image Reconstruction . . . 51
3.3.4 Image Analysis . . . 54
3.4 Results . . . 56
3.4.1 NEMA NU2-Like Phantom . . . 56
3.4.2 FMISO NSCLC Phantom . . . 61
3.5 Discussion . . . 63
3.6 Conclusion . . . 66

4 4D-PET Reconstruction with Adaptive-Knot Cubic B-Splines 67
4.1 Introduction . . . 67
4.2 Adaptive-Knot Placement Algorithm . . . 69
4.2.1 B-Spline Basis Functions . . . 69
4.2.2 Good Knot Placement . . . 69
4.2.3 The Weighted Savitzky-Golay Filter . . . 70
4.2.4 Knot Placement Algorithm . . . 72
4.3 Methods . . . 72
4.3.1 Validation of the Adaptive Knot Placement Algorithm with Single TACs . . . 72
4.3.2 4D-PET Reconstruction Study . . . 75
4.4 Results . . . 79
4.4.1 Validation of the Adaptive-Knot Placement Algorithm with Single TACs . . . 79
4.4.2 4D-PET Reconstruction Study . . . 81
4.5 Discussion . . . 86
4.6 Conclusion . . . 89

5 4D-PET Reconstruction using a Spline-Residue Model 90
5.1 Introduction . . . 90
5.2 Nonparametric Spline-Residue Description of dPET TACs . . . 92
5.3 Nested-MAP2: Temporally Regularised 4D-PET Reconstruction Algorithm for Linear Kinetic Models . . . 93
5.4 Methods . . . 97
5.4.1 Digital Phantom Data . . . 97
5.4.2 Image Reconstruction . . . 98
5.4.3 Image Analysis . . . 100
5.5 Results . . . 102
5.5.1 Image Quality Metrics . . . 102
5.5.2 Parametric Maps . . . 106
5.6 Discussion . . . 112
5.7 Conclusion . . . 117

6 Radiogenomic Comparison of the Compartment and Spline-Residue Models 119
6.1 Introduction . . . 119
6.2 Deriving Kinetic Parameters from Fits of the Spline-Residue Model 120
6.3 Methods . . . 122
6.3.1 Patient Data . . . 122
6.3.2 PET-CT Protocol . . . 123
6.3.3 Analysis of Data from Tumour ROIs . . . 125
6.3.4 Genomic Data . . . 128
6.3.5 Correlating Imaging and Genomic Parameters . . . 130
6.4 Results . . . 133
6.4.1 Assessment of Model Fit Quality . . . 133
6.4.2 Correlations Between Imaging and Genomic Measures . . . 135
6.5 Discussion . . . 141
6.6 Conclusion . . . 144

7 Conclusions and Further Work 145
7.1 Summary of Results . . . 145
7.2 Further Work . . . 148

Appendices

A Supplementary Results for Chapter 6 152

References 160

List of Figures

1.1 Illustration of the positron emission and annihilation process that occurs during PET imaging, along with the detection of the resulting coincident 511 keV photons by the PET detector ring that surrounds the patient. 2

1.2 Illustration of the main differences between static and dynamic PET. Static PET produces a single image showing the activity integrated over time, while dynamic PET produces a sequence of images showing how the activity distribution varies over time. 3

2.1 A single PET scanner detector ring with an annihilation event (occurring at the star) being by detected along the LOR depicted by a double-headed arrow. The relationship between the coordinates of the annihilation event in (d_x, d_y) image space and (r, w) space for a given projection angle ϕ is also shown. 11

2.2 Noise-free digital phantom image (a) and corresponding sinogram with 200 angular projections (b). 12

2.3 Illustration of angular mashing in which each group of four adjacent polar angles (and thus sinogram rows) are added together. Note that the mashed sinogram is smaller than the original because it contains a factor of 4 fewer rows, though in this figure it is shown as half the height of the original sinogram to make it easier to visualise. 13

2.4 Illustrations of a true coincidence event (a), an attenuated event (b), a scattered coincidence (c) and a random coincidence (d). Solid arrows depict the photon paths and where appropriate the dashed line shows the erroneous LOR recorded by the scanner. 15

2.5 Analytical reconstructions of noise free sinogram data generated from a digital phantom (shown in (a)). Images of the activity distribution reconstructed with simple and filtered backprojection are shown in (b) and (c), respectively. 20

2.6 Flowchart summarizing the iterative reconstruction procedure. In practice most iterative algorithms are terminated after a fixed number of iterations as opposed to being run until convergence. 22

2.7 Schematic diagram of a two-tissue compartment model. The flow of tracer between compartments is characterised by rate constants (k-values). 37
2.8 Comparison of typical shapes of normalized (a) Patlak, (b) spectral model and (c) cubic B-spline basis functions. 39

3.1 Axial slice of the NEMA-NU2-like digital image quality phantom. This slice shows the centres of each of the spheres that surround the activity-free cylindrical insert. 44
3.2 Schematic diagrams showing irreversible (a) two and (b) three tissue compartment models of FMISO uptake. Flows between compartments are defined by rate-constants (k values) and compartment tracer concentrations. 46
3.3 Illustration of the final time-frame of the FMISO NSCLC dPET phantom. 47
3.4 Spline fits to the clinical hypoxic tumour TAC. The TAC has been binned into the 'finer' time-frame sequence of the original clinical data ([1×30 s, 12×5 s, 6×10 s, 6×30 s, 10×60 s, 6×300 s] + 2×600 s frames 2- and 4-hours post-injection) and the 'coarser' frame sequence used for the dPET phantom ([1×30 s, 6×5 s, 6×20 s, 7×60 s, 10×120 s, 3×300 s] + 2×600 s frames 2- and 4-hours post-injection). Note that the first seven and last five time-frames of the two frame sequences are identical. 48
3.5 Reconstructed images of the NEMA NU2-like digital phantom using OSEM with 4 (a,c) and 10 (b,d) iterations. Images are shown before (a,b) and after (c,d) the application of Gaussian smoothing. 58
3.6 Reconstructed images of the NEMA NU2-like digital phantom using OS-MAP with $\beta = 0.0001$ (a), $\beta = 0.001$ (b), $\beta = 0.01$ (c) and $\beta = 0.1$ (d). 58
3.7 Plots of the MAP objective function defined in equation (3.3) vs iteration number for each of the NEMA-like phantom reconstructions. Note that the reconstruction with $\beta = 0$ corresponds to the OSEM reconstruction. 59
3.8 2C3K, 3C5K and adaptive-knot cubic B-spline fits to the TACs obtained from (a) hypoxic tumour, (b) normoxic tumour, (c) healthy lung and (d) healthy spine ROIs in an FMISO dPET image sequence of a stage IV NSCLC patient. 61

3.9 MAP-reconstructed images of a single noise realization of the final 10 minute time-frame of the FMISO NSCLC patient. This figure illustrates the effect of varying the β smoothing parameter. Note that the MAP reconstruction with $\beta = 0$ is the standard MLEM reconstruction. 63
3.10 Iteration-by-iteration plots of image bias versus noise, averaged over all time-frames and patient voxels, for all MAP reconstructions of the FMISO NSCLC phantom (a), and plots of image bias versus noise in the final reconstruction iteration with varying β (b). MAP reconstructions with $\beta = 0$ correspond to MLEM reconstructions. . 64

4.1 Illustration of the Savitzky-Golay method for calculating the magnitude of the 4^{th} derivative with $k = 11$ and $q = 5$. (a) Shows a simulated noisy TAC and (b) shows the magnitude of the 4^{th} derivative calculated from the noisy TAC data using the Savitzky-Golay method along with the 4^{th} derivative of the true TAC. Note that the simulated noisy TAC was generated using the methodology described in section 4.3.1.1. 71
4.2 Summary of the three cases investigated in the single TAC study to validate the proposed knot placement algorithm. 74
4.3 Single noise realization of spline fits with N_{opt} free knots, 11 free knots and N_{opt} uniform knots to an AIF TAC (a,c,e) and a tumour TAC (b,d,f), for each of the three cases listed in Figure 4.2. Note that the y-scale of the high noise (case 2) tumour TAC in plot (d) is higher than for the other tumour TACs. Spline fits with uniform knots were not performed in case 3. 80
4.4 Reconstructed images of a single noise realization of a 5 minute time-frame of the FMISO NSCLC dPET phantom, starting 40 minutes into the scan. Images produced by every reconstruction algorithm tested in this chapter are shown. 82
4.5 Mean of reconstructed images of a 5 minute time-frame of the FMISO NSCLC dPET phantom, averaged over all 50 noise realizations. Images produced by every reconstruction algorithm tested in this chapter are shown. 83
4.6 Images showing the voxel variance of reconstructed images of a 5 minute time-frame of the FMISO NSCLC dPET phantom, calculated over all 50 noise realizations. The variance values are expressed as a percentage of the mean radioactivity concentration of the ground-truth phantom image of this time-frame. 84

4.7 Iteration-by-iteration plots of image bias versus noise (averaged over all time-frames) for each image reconstruction algorithm, averaged over all patient voxels (a,b) and just the tumour region (c,d). Results are shown for reconstructions with (b,d) and without (a,c) spatial regularization. Note that the scales on the plots are different. 85
4.8 Iteration-by-iteration plots of image bias versus noise (averaged over all time-frames) for each image 4D reconstruction algorithm, averaged over all patient voxels (a) and just the tumour region (b). Results are shown for reconstructions with and without spatial regularization. Note that the scales on the plots are different. 86
4.9 Single noise realization of a measured hypoxic tumour TAC, taken from images reconstructed with the MAP and spline and optimized spectral model-based nested-MAP reconstructions. Tumour TACs taken from the whole hypoxic tumour ROI (a) and a single voxel at the centre of the hypoxic tumour region (b) are shown. 87
4.10 Iteration-by-iteration plots of image bias versus noise (averaged over first 120 s of the scan) for each image nested-MAP reconstruction, averaged over all patient voxels (a) and just the tumour region (b). 87

5.1 Comparison of typical shapes of normalized (a) cubic B-spline, (b) spectral model and (c) spline-residue model basis functions. 93
5.2 Image of the final time-frame of the realistic phantom, with a white box marking the sub-volume of the phantom for which parametric maps were calculated. 102
5.3 Comparison of image reconstruction results obtained for the realistic (first column) and simplified (second column) phantoms using 4D algorithms based on the (a) and (b) cubic splines, (c) and (d) spectral and (e) and (f) spline-residue models with different temporal roughness penalties. The plots show image noise (weighted standard deviation, σ_w) versus absolute bias, averaged over all time-frames and patient voxels, at each reconstruction iteration, and have different scales to allow clear visualisation of performance differences between the different penalty functions. The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown. 103
5.4 Comparison of image quality metrics for 4D reconstructions of the realistic phantom based on linear models (with their optimal temporal roughness penalty) and on the 2C3K model. The noise and bias metrics are averaged over the whole patient (a) and (b) and tumour (c) and (d) phantom regions, and over the entire scan time (a) and (c) and first 120 s (b) and (d). The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown. . 104

5.5 Comparison of image quality metrics for 4D reconstructions of the simplified phantom based on linear models (with their optimal temporal roughness penalty) and on the 2C3K model. The noise and bias metrics are averaged over the whole patient (a) and (b) and tumour (c) and (d) phantom regions, and over the entire scan time (a) and (c) and first 120 s (b) and (d). The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown. 106

5.6 (a) Fractional iteration-to-iteration change in total mean square error (TMSE) versus iteration number, plotted for 2C3K and linear model-based nested-MAP2 reconstructions (with the optimal temporal roughness penalties and the γ'_j hyperparameters free to vary) of the realistic phantom. (b) A magnification of the rectangular region outlined in (a). 107

5.7 Iteration-by-iteration fractional change in total mean square error (TMSE) plotted as a function of nested-MAP2 iteration number, for linear-model based reconstructions of the realistic phantom (using the optimal temporal roughness penalty for each linear model) with γ'_j fixed and free to vary. 107

5.8 Single noise realisation of k_{flux} parametric maps of the subregion of the realistic phantom obtained from image sequences reconstructed using the conventional (non-4D) MAP algorithm, as well as using the 2C3K and linear model-based nested-MAP2 reconstructions (with their optimal temporal roughness penalties). 108

5.9 Voxel-by-voxel plots of (a) the absolute bias and (b) the noise (standard deviation) in the k_{flux} parametric maps of the subregion of the realistic phantom obtained from image sequences reconstructed using the conventional (non-4D) MAP algorithm, as well as using the 2C3K and linear model-based nested-MAP2 reconstructions (with their optimal temporal roughness penalties). Bias and standard deviation values are expressed as percentages of the average ground-truth k_{flux} value across the phantom sub-region. 108

5.10 Fits of the 2C3K model to voxel TACs taken from reconstructed images of the realistic phantom. Voxel TACs taken from the centre of the hypoxic tumour region (a) and the healthy lung region (b) are shown. The 2C3K fits to the true phantom voxel TACs are also shown for comparison. 109

5.11 (a) Absolute bias and (c) noise (standard deviation) of fitted parameter values in parametric maps derived from the 4D-PET reconstructed images of the realistic phantom, averaged over the entire parametric map of the image sub-region shown in Figure 5.2. Values are plotted for each of the kinetic parameters as fractions of those obtained for the same parameter from analysis of conventional (non-4D) MAP-reconstructed images. Results for the k_3 parametric maps obtained from the 2C3K, spectral and spline-residue model based 4D reconstructions are re-plotted on larger scales in (b) and (d) to make them visible. 110

5.12 (a) Absolute bias and (c) noise (standard deviation) of fitted parameter values in parametric maps derived from the 4D-PET reconstructed images of the simplified phantom, averaged over the entire parametric map of the image sub-region shown in Figure 5.2. Values are plotted for each of the kinetic parameters as fractions of those obtained for the same parameter from analysis of conventional (non-4D) MAP-reconstructed images. Results for the k_3 parametric maps obtained from the 2C3K, spectral and spline-residue model based 4D reconstructions are re-plotted on larger scales in (b) and (d) to make them visible. 111

5.13 (a, b) Absolute bias and (c, d) noise (standard deviation) of fitted parameter values in parametric maps derived from the 4D-PET reconstructed images of the realistic phantom, averaged over the hypoxic tumour region alone. For each of the kinetic parameters, the values plotted are fractions of those obtained for the same parameters from analysis of conventional (non-4D) MAP-reconstructed images. 112

5.14 Comparison of image quality metrics for 4D reconstructions of the realistic phantom based on linear models (with their optimal temporal roughness penalty) and on the 2C3K and 3C5K models. The noise (weighted standard deviation, σ_w) and bias metrics are averaged over the whole patient (a, b) and tumour (c, d) phantom regions, and over the entire scan time (a, c) and first 120 s (b, d). The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown. 115

6.1 Schematic diagrams showing irreversible (a) two- and (b) three-tissue compartment models of FDG uptake. Flows between compartments are defined by rate-constants (k values) and compartment tracer concentrations. 122

6.2 Coronal maximum intensity projections of static SUV images of a patient before (a) and after (b) the administration of the drug Metformin. Note that the scale in both images is the same. The white arrow in both images points to the primary breast cancer tumour. The post-Metformin FDG uptake levels of the axillary lymph nodes (circled in (b)) are notably higher. 126
6.3 Examples of fits of the three-exponential Feng AIF model given in equation (6.6) to measured arterial blood TACs from two separate patients. 127
6.4 Schematic diagram summarizing the gene expression process. A more detailed description is given in [114]. 129
6.5 Schematic diagram summarizing the available Genomic and PET imaging parameters, and how they were compared to each other. Note that the analysis shown in this figure was carried out for both the baseline data and the pre- to post-Metformin changes in the data. 132
6.6 Mean weighted RSS error (calculated using leave-one-out cross-validation) for fits of the spline-residue model to a subset of 10 patient TACs as a function of the number of free knots used. 134
6.7 Examples of 2C3K, 3C5K and spline-residue model fits to pre-drug tumour TACs from two different patients. (c) and (d) show the corresponding residue functions from each model fit to the TACs shown in (a) and (b), respectively. 135
6.8 Boxplots of the rate and flux constants from the fitted 2C3K (a), 3C5K (b) and spline-residue model (c), as well as the parameters from the static PET images (d). Any data point that lies further than 1.5x the interquartile range from the first of third quartiles (the top and bottom of each of the boxes, respectively) are shown with red crosses. Note that a logarithmic scale is used to improve the visibility of both the boxes and the outliers. 136
6.9 Boxplots of the K_1 (a) and the k_{flux} (b) parameters produced by each of the different model fits. Any data point that lies further than 1.5x the interquartile range from the first of third quartiles (the top and bottom of each of the boxes, respectively) are shown with red crosses. 136
6.10 Plots of pre- to post-Metformin change in spline-residue k_{flux} versus changes in GLUT1 expression (a), HK2 expression (b), hypoxia signature score (c) and proliferation signature score (d). Linear fits to these plots are shown by the solid black line. All of the correlations shown here passed the Benjamini-Hochberg procedure, and are listed in Table 6.5. Note that the hypoxia and proliferation scores are dimensionless. 138

6.11 Plots of pre- to post-Metformin change in 2C3K k_3 (a) and 3C5K k_2 (b) versus changes in angiogenesis signature score. Linear fits to these plots are shown by the solid black line. One data point on both plots (circled) appears to be an outlier. Both of the correlations shown here passed the Benjamini-Hochberg procedure, and are listed in Table 6.5. Note that the angiogenesis scores are dimensionless. . 138
6.12 Fits of the 2C3K, 3C5K and spline-residue models to a particularly noisy TAC. The 2C3K and 3C5K fits to this TAC give the circled outlier in Figure 6.11 139
6.13 Plots of the k_{flux} parameter derived from spline- residue model fits versus the k_{flux} parameter from 2C3K fits for the baseline data (a) and the pre-post Metformin changes (b). The black line is a plot of y = x. 140

List of Tables

3.1 Settings used for the MAP reconstructions of the NEMA-NU2-like and FMISO NSCLC dPET phantoms. 53

3.2 CRCs for each reconstruction of the NEMA NU2-like phantom. CRCs are expressed as percentages, with errors shown in brackets. The highest CRC for each sphere is highlighted in bold. The acronym GS stands for Gaussian smoothing. 59

3.3 BVs for each reconstruction of the NEMA NU2-like phantom. BVs are expressed as percentages, with errors shown in brackets. The lowest BV for each sphere is highlighted in bold. The acronym GS stands for Gaussian smoothing. 60

3.4 CNRs for each reconstruction of the NEMA NU2-like phantom. BVs are expressed as percentages, with errors shown in brackets. The highest CNR for each sphere is highlighted in bold. The acronym GS stands for Gaussian smoothing. 60

3.5 Weighted RSS errors for model fits to patient TACs calculated using leave-one-out cross-validation. The lowest RSS value for each tissue region is shown in bold. Fits that passed the Wald–Wolfowitz runs test at the 5% significance level are underlined. 62

3.6 Published results from an mMR scan of a real NEMA NU2 phantom [33]. CRCs, BVs and CNRs are shown, with the asociated errors given in brackets. 65

4.1 Settings used for the each of the reconstruction algorithms used in this chapter. 76

4.2 Results for each of the cases shown in Figure 4.2 for the validation of the proposed knot placement algorithm. For each case the range of N_{opt} values and the $\langle RMS \rangle$ values for spline fits using N_{opt} free knots, 11 free knots and N_{opt} uniform knots are shown for both the tumour and AIF TACs. Spline fits with uniform knots were not performed in case 3. 79

6.1 Illustration of how the spline-residue kinetic parameters can be obtained from fits of the 2C3K and 3C5K compartment models. Note that for the 2C3K and 3C5K models the asymptotic (i.e. $T_{scan} \rightarrow \infty$, where T_{scan} is the duration of the scan) expressions for the flux constant k_{flux} are shown. 122
6.2 List of key inclusion and exclusion criteria used when recruiting the patients in this dataset. 124
6.3 Patient and tumour characteristics, as well as a summary of the patient data acquired. The patient estrogen receptor (ER) and human epidermal growth factor receptor-2 (HER2) statuses is also shown. Note that here 'paired' means that data was acquired both before and after the patient was administered with Metformin. . . 125
6.4 Median weighted RSS errors for model fits to patients TACs, calculated using leave-one-out cross validation, and the number of TACs for which each model passes the Wald-Wolfowitz runs test at the 5% significance level. The best result for each metric is highlighted in bold. 134
6.5 Statistically significant results remaining after the Benjamini-Hochberg procedure (with $\alpha = 0.2$) was used to account for multiple hypothesis testing. Note that all of the F-tests for the multiple regression analyses were filtered out the Benjamini-Hochberg procedure. Results are shown in order of ascending p-value. 137
6.6 Results from Z-transform-based test for comparing whether the correlation coefficients between k_{flux} parameters derived from the spline-residue and 2C3K models and the genomic measures were significantly different from each other. Results for both the baseline and pre- to post-Metformin changes are shown. Statistically significant results ($p < 0.05$) are highlighted in bold. 141

A.1 Correlations between GLUT1 expression and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold. 153
A.2 Correlations between HK2 expression and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold. 154
A.3 Correlations between angiogenesis signature score and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold. . 155

A.4 Correlations between hypoxia signature score and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold. 156
A.5 Correlations between proliferation signature score and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold. . 157
A.6 Multiple regression results for the linear regression model containing the expression levels of all available glucose transporter genes. F-test results for baseline and pre-post metformin change data are shown. Statistically significant ($p < 0.05$) results are highlighted in bold. . 158
A.7 Multiple regression results for the linear regression model containing the expression levels of all available hexokinase genes. F-test results for baseline and pre-post metformin change data are shown. Statistically significant ($p < 0.05$) results are highlighted in bold. . 159

List of Abbreviations

2C3K, 3C5K . Irreversible two- and three-tissue compartment models.

1D, 2D, 3D . One-, two- and three-dimensional.

3D-RP . . . . 3D-reprojection.

4D . . . . . Four dimensional, referring in this thesis to dynamic image sequences with two or three spatial dimensions and one temporal dimension.

ACF . . . . . Attenuation Correction Factor.

AICc . . . . . Akaike Information Criterion, corrected for finite sample sizes.

AIF . . . . . Arterial Input Function.

BV . . . . . Background Variability.

CNR . . . . . Contrast to Noise Ratio.

CRC . . . . . Contrast Recovery Coefficient.

CT . . . . . X-ray Computed Tomography.

DCE-MRI . . Dynamic Contrast-Enhanced Magnetic Resonance Imaging.

DNA . . . . . Deoxyribonucleic acid.

DOF . . . . . Degrees of Freedom.

dPET . . . . . Dynamic Positron Emission Tomography.

ECOG . . . . . Eastern Cooperative Oncology Group.

ER . . . . . Estrogen Receptor.

EM . . . . . Expectation-Maximization.

FAZA . . . . . [^{18}F]-fluoroazomycin arabinoside.

FBP . . . . . Filtered Backprojection.

FDG . . . . . [^{18}F]-Fluorodeoxyglucose.

FDR . . . . . False Discovery Rate.

FLT . . . . . 3'-deoxy-3'-[^{18}F]-fluorothymidine.

FMISO . . . . [^{18}F]-fluoromisonidazole.

FORE . . . . . Fourier Re-binning.

FOV . . . . . Field of View.

FPKM . . . . Fragments per Kilobase Million.

GATE . . . . . Geant4 Application for Tomographic Emission.

GCV . . . . . Generalized Cross Validation.

GLUT . . . . . Glucose Transporter.

HER2 . . . . . Human Epidermal growth factor Receptor-2.

HK . . . . . Hexokinase.

IRF . . . . . Impulse Response Function.

LOR . . . . . Line of Response.

LSO . . . . . Cerium-doped Lu_2SiO_5

MAP . . . . . Maximum *a posteriori*.

MB . . . . . Mean Bias.

ML . . . . . Maximum-Likelihood.

MLEM . . . . Maximum-Likelihood Expectation-Maximization.

MR/MRI . . . Magnetic Resonance Imaging.

mRNA . . . . Messenger RNA.

MSE . . . . . Mean Square Error.

NEMA . . . . National Electrical Manufacturers Association.

Nested-MAP2 The 4D nested-MAP algorithm proposed in chapter 5. This algorithm was named nested-MAP2 to distinguish it from the more conventional nested-MAP algorithm used in chapter 4.

NCC . . . . . Normalization Correction Coefficient.

NSCLC . . . . Non-Small Cell Lung Cancer.

OSEM . . . . . Ordered Subsets Expectation-Maximization.

OSL . . . . . One-Step-Late.

OS-MAP . . . Maximum *a Posteriori* reconstruction with Ordered Subsets.

PCG . . . . . Preconditioned Conjugate Gradient.

PENELOPE . Penetration and ENErgy LOss of Positrons and Electrons.

PENELOPET Positron Emission Tomography simulation using PENELOPE.

PET . . . . . Positron Emission Tomography.

PETSIM . . . Simulation package for Positron Emission Tomography.

PET-SORTEO Simulation of Realistic Tridimensional Emitting Objects.

PICD . . . . . Parametric Iterative Coordinate Descent.

PIR . . . . . Parametric Iterative Reconstruction.

PMOLAR-1T Parametric Motion-compensation OSEM List-mode Algorithm for Resolution-recovery for the 1-Tissue compartment model.

POSEM . . . . Parametric Ordered Subsets Expectation-Maximization.

PSF . . . . . Point Spread Function.

PVE . . . . . Partial Volume Effect.

RMS . . . . . Root Mean Square.

RNA . . . . . Ribonucleic acid.

RNASeq . . . RNA Sequencing.

ROI . . . . . Region of Interest.

RSS . . . . . Residual Sum of Squares.

SD . . . . . Standard Deviation.

SIMSET . . . Simulation System for Emission Tomography.

SNR . . . . . Signal-to-Noise Ratio.

SPECT . . . . Single Photon Emission Computed Tomography.

SSRB . . . . . Single Slice Re-binning.

SUV . . . . . Standardised Uptake Value.

SVD . . . . . Singular Value Decomposition.

TAC . . . . . Time-Activity Curve.

TBR . . . . . Tumour to Blood Ratio.

TMSE . . . . . Total Mean Square Error.

TOF . . . . . Time of Flight.

TOR . . . . . Tube of Response.

WLS . . . . . Weighted Least Squares.

Chapter 1

Introduction

1.1 PET Imaging in Oncology

Positron emission tomography (PET) is a functional imaging technique in which a patient is injected with a radioactive tracer and scanned using a dedicated PET scanner [149]. Three-dimensional images of the activity distribution are then reconstructed from the scanner data using either analytical or iterative image reconstruction algorithms. While PET imaging has been applied to neurology [117] and cardiology [61], the most common clinical application of PET is in oncology [39, 86].

The radiotracers used in PET consist of a positron emitting radionuclide attached to a biologically active molecule. The process by which the PET signal is produced and measured is summarized in Figure 1.1. The radionuclide undergoes β^+ decay, resulting in the emission of a positron. This positron rapidly loses kinetic energy via inelastic scattering with atomic electrons within the patient. Once the positron has lost most of its kinetic energy, it annihilates with an electron, resulting in the creation of two 511 keV photons travelling at approximately 180^o to each other. These coincident photons are then detected by a ring of detectors surrounding the patient.

Most clinical PET scans are now performed with hybrid scanners that include an X-ray computed tomography (CT) or magnetic resonance imaging (MRI) scanner as well as the PET scanner. PET-CT and PET-MRI imaging allows the functional information from the PET image to be combined with the anatomical information from the CT or MRI image.

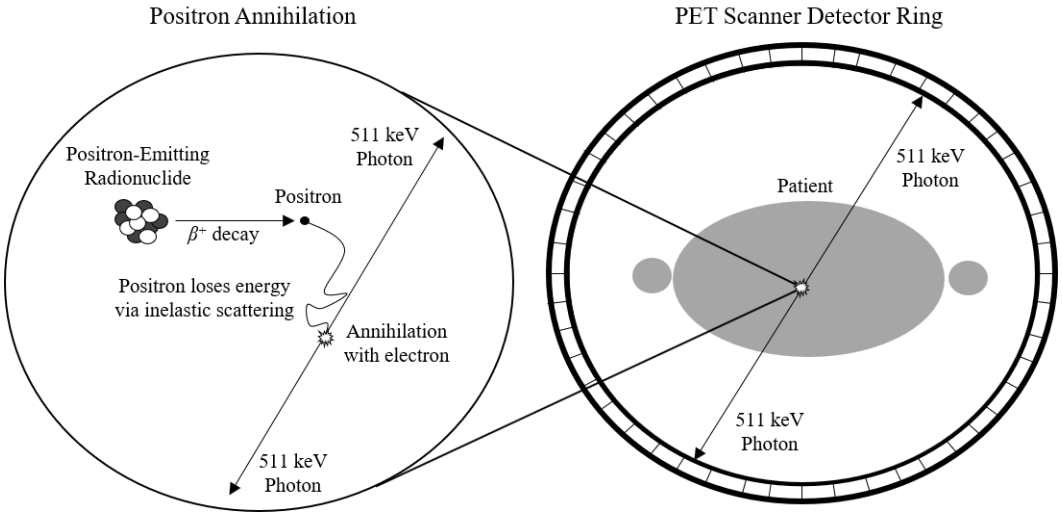


Figure 1.1. Illustration of the positron emission and annihilation process that occurs during PET imaging, along with the detection of the resulting coincident 511 keV photons by the PET detector ring that surrounds the patient.

The most commonly used radiotracer in oncological PET imaging is ^{18}F-labelled fluorodeoxyglucose (FDG), which provides a means for imaging tissue metabolism, specifically glycolysis [26, 115]. Other radiotracers that have been used for cancer imaging with PET include 3'-deoxy-3'-[^{18}F]-fluorothymidine (FLT), a tracer of cell proliferation, and [^{18}F]-fluoromisonidazole (FMISO) and [^{18}F]-fluoroazomycin arabinoside (FAZA), which are tracers of hypoxia [39].

Currently, the main clinical uses of PET imaging in oncology are [12]:

- Detection and differential diagnosis of benign and malignant tumours.
- Tumour staging and prognostic stratification.
- Treatment planning.
- Evaluation of treatment response.
- Restaging and detection of recurrent cancer.
- Development of new anticancer drugs.

Most research into PET imaging for oncology is focused on the development of novel radiotracers [112], the search for new diagnostic applications of PET [39] and the development of image analysis and reconstruction techniques for improving the quantitative information that can be obtained [158, 119]. The work presented in this thesis is focused on the last of these research areas.

1.2 Dynamic PET Imaging

1.2.1 Static vs. Dynamic PET

In the majority of PET scans, the patient is injected with the radiotracer and then sent to a waiting room to allow time for the tracer to be taken up by the patient tissues prior to the scan. Following the scan, the scanner data is reconstructed into a single image, showing the activity concentration integrated over the time-course of the scan. This is known as static PET imaging.

Dynamic PET (dPET) imaging is an alternative methodology in which the patient is scanned immediately after injection. The scanner data is then divided into time-frames, and each one is reconstructed as a separate image. The duration of a dPET scan is typically longer than that of a static PET scan. The differences between static and dynamic PET imaging are illustrated in Figure 1.2.

The main advantage of dPET over static PET is that many individual biological processes are more closely related to the rate of change of tracer uptake rather than the uptake at any one time-point. In dPET imaging, quantitative information concerning tracer uptake kinetics can be obtained by extracting time-activity curves

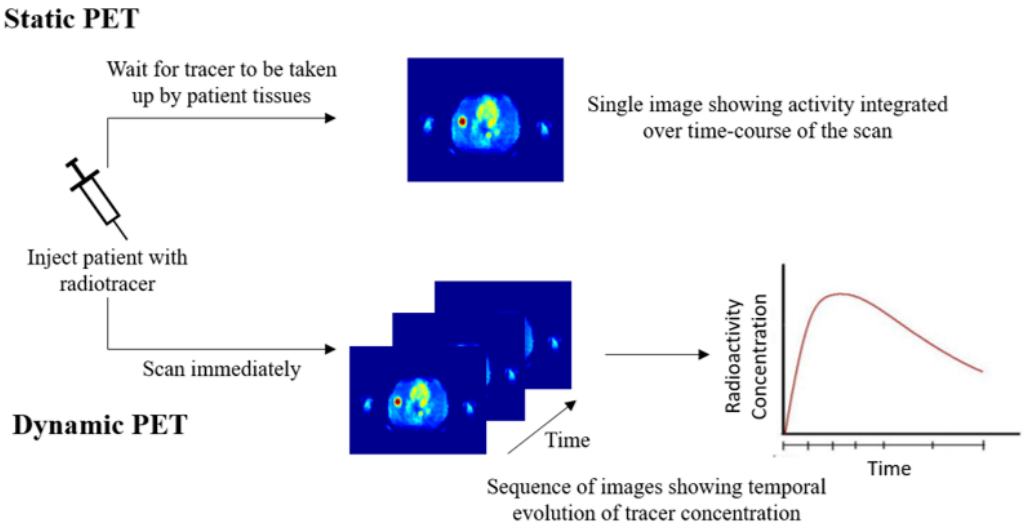


Figure 1.2. Illustration of the main differences between static and dynamic PET. Static PET produces a single image showing the activity integrated over time, while dynamic PET produces a sequence of images showing how the activity distribution varies over time.

(TACs) from regions of interest (ROIs) drawn on the dPET image sequences and fitting them with tracer kinetic models.

1.2.2 Limitations of Dynamic PET

Despite its potential for providing improved quantification of biological processes relative to static PET, dPET suffers from a number of limitations:

- Patient movement during the scan can lead to blurring artefacts and misalignment of anatomical features between time-frames. In the context of oncological imaging this can make it harder to detect small lesions and to accurately quantify the tracer uptake characteristics of the tumours being studied.
- The arterial input function (AIF), defined as the time-course of the radioactivity concentration in the arterial blood, is required by most kinetic models, though in some situations the AIF can be hard to measure.
- Selection of the time-frame durations is somewhat arbitrary.
- Dividing the scanner data into time-frames results in a limited number of photon counts in each frame. This in turn leads to a poor signal-to-noise ratio (SNR) in the reconstructed images.
- Noise is amplified by the image reconstruction process. Furthermore, the noise in the reconstructed activity images is spatially correlated, object dependant and unlike the raw scanner data it is not well described with a Poisson distribution. Accurate noise modelling of the TAC data is rarely performed, as it is both complex and time-consuming [7, 118].

Correction for patient motion is an ongoing area of research, and references [123, 48] provide detailed reviews of the methods proposed so far.

Measurement of the AIF can be achieved by serial blood sampling of the patient during the scan or by drawing an ROI in a blood pool within the dPET image sequences. The latter approach has the advantage of being non-invasive, though for small blood volumes the measurement can be degraded by partial volume effects

(PVEs). Methods to correct for PVEs in image-derived AIFs from small blood regions have been proposed [38, 56, 143].

There exists a trade-off between noise and temporal resolution when selecting the dPET time frames, with longer time-frames having a better SNR but a worse temporal resolution. Short time-frames are normally used at the start of the scan to capture the faster TAC variation at early time-points, while longer frames are used towards the end. The time-frame durations are generally defined using known information on the dynamics of the tracer being used. However, there is no standardised method for determining the time-frame durations, meaning that the exact time-frame durations used for a given tracer often vary from centre to centre. Furthermore, the tracer kinetics can vary considerably from patient to patient. While some published research has investigated ways in which a time-frame sequence can be calculated for an individual TAC [90, 91, 92], there are currently no methods that extend these ideas to entire dPET images in order to rigorously calculate patient-specific time-frame sequences.

The high noise levels in the reconstructed image sequences result in noisy parameter estimates when fitting tracer kinetic models to the reconstructed TACs. Furthermore, because the exact noise distribution of the TAC data is not usually known, errors in modelling the noise when fitting the models can lead to biased kinetic parameter estimates. This thesis is focused on addressing these issues through the use of 4D image reconstruction algorithms, which are introduced in the next subsection.

1.2.3 4D-PET Reconstruction

In conventional dPET imaging, each time-frame is independently reconstructed as an individual image. 4D-PET reconstruction is an alternative methodology in which temporal models are fitted to each voxel TAC during the image reconstruction process. This takes advantage of the fact that, unlike the reconstructed image sequences, the raw scanner data can be accurately modelled as following a Poisson noise distribution. Furthermore, the models used by

most 4D-PET reconstruction algorithms are temporally smooth, and by fitting smooth functions during the reconstruction the noise amplification that typically occurs during image reconstruction is suppressed at its source [124, 129, 157, 178]. However, some 4D-PET reconstruction algorithms (e.g. [128]) do not use temporally-smooth basis functions.

Some researchers have also proposed 4D-PET reconstruction algorithms that simultaneously fit temporal models and estimate motion correction parameters (e.g. [68, 73, 72]). Such approaches have been shown to give more accurate motion correction compared to applying motion correction to 3D-reconstructed images, though the computational complexity of 4D-PET reconstruction algorithms that incorporate motion correction is substantially higher than 4D-PET reconstruction algorithms that don't.

While 4D-PET reconstruction has been shown to provide improvements in the SNR of both the reconstructed image sequences and the kinetic parameters, there is no consensus on what temporal function works best [124, 129]. In some instances the kinetic model of interest is used as the temporal function, allowing the kinetic parameters to be obtained directly from the reconstruction. This approach is sometimes referred to as direct or parametric PET image reconstruction [124, 178].

In cancer imaging many tissues can be present in the scanner field of view (FOV). The tracer can behave very differently in each tissue, leading to a wide range of TAC shapes to be fitted. It has been shown that bias from poorly modelled regions can spatially propagate to well modelled regions during 4D-reconstruction [81], meaning that a kinetic model designed to accurately describe the uptake kinetics within a tumour may perform poorly in other imaged regions, introducing bias into kinetic parameter estimates even in the well described regions. In such situations it is therefore advantageous to use highly flexible functions which can adequately fit TACs in all regions, though the optimal choice of function for doing so remains an open question [124, 129].

1.3 Thesis Outline

This thesis presents an investigation into the use of 4D-PET reconstruction algorithms in situations with diverse kinetics, with a particular emphasis on cancer imaging.

Chapter 2 presents the theoretical foundations of both static and 4D-PET image reconstruction, along with a review of the main static and 4D reconstruction algorithms that have been proposed to date. Most of the methodologies discussed are also applicable to single photon emission computed tomography (SPECT) image reconstruction.

Chapter 3 introduces the Monte-Carlo simulation methodology used in the rest of the thesis to generate PET scanner data from digital phantoms. The methods used to calculate the system matrix and evaluate the quality of reconstructed images in the rest of the thesis are also presented, along with the results from two 'test' simulation studies.

Of the temporal models that are frequently used in 4D-reconstruction, the spectral model [25], which describes TACs as weighted sums of exponential decays convolved with the AIF, and spline functions are among the most flexible. In chapter 4, a 4D-PET reconstruction algorithm using cubic B-splines with a theoretically-driven knot placement algorithm is proposed. This is compared to 'conventional' (non-4D) reconstructions and spectral model-based 4D reconstruction. The algorithms are tested using Monte-Carlo simulated data, generated for a digital thoracic phantom with uptake kinetics based on an FMISO dPET scan of a non-small cell lung cancer (NSCLC) patient. The effects of incorporating spatial regularization into the reconstruction are also investigated.

While the adaptive-knot spline-based reconstruction is found to perform well overall, producing the least biased images of all the algorithms being compared, at early time-frames the spectral model-based reconstruction is found to produce a better bias-noise trade-off. TACs vary most rapidly early on during imaging, and so short time-frames are used for the early time-points, making them particularly noisy. Because the basis functions of the spectral model are generated by convolving

exponential decays with the AIF, their shapes are much more constrained at these early time-points than those of B-splines; but for the same reason they can still describe the early parts of TACs well. Consequently the spectral model may have an advantage at early time-points.

In chapter 5, a new kinetic model describing TACs as weighted sums of convolutions of the AIF with cubic B-spline basis functions is proposed. This new model is termed the spline-residue model. Convolution with the input function constrains the spline-residue model at early time-points, potentially enhancing noise suppression in early time-frames, while still allowing a wide range of TAC descriptions over the entire imaged time-course, thus limiting bias. A similar model has been proposed for kinetics analysis of post-reconstruction dPET data [111], but to our knowledge the use of the model for 4D reconstruction has not been proposed or evaluated. Spline-residue model-based reconstruction is compared to 4D reconstructions using the spectral model, adaptive-knot cubic B-splines and an irreversible two-tissue compartment model. The 4D reconstructions in this chapter are carried out using a novel algorithm that incorporates both spatial and temporal roughness penalties into the objective function. Of the methods compared in this chapter, spline-residue model-based reconstruction is found to perform the best. Incorporation of a temporal roughness penalty into the objective function is found to improve the bias-noise trade-off obtained with 4D-reconstruction based on the spline, spectral and spline-residue models.

The proposed spline-residue model falls under a wider class of 'nonparametric' kinetic models. A number of studies have shown that, due to their inherent flexibility, such models provide improved descriptions of TACs compared to more commonly used compartment models [41, 59, 71, 110, 111]. However, to our knowledge no studies have investigated whether the kinetic parameter estimates obtained from these models provide an improved description of the underlying patient biology.

In chapter 6, the spline-residue model is used to analyse FDG-dPET data from a clinical dataset comprising 41 primary breast cancer patients. These patients underwent dPET scans and tumour biopsies before and after the administration

of the drug Metformin. Fits of the spline-residue model are compared to those of irreversible two- and three-tissue compartment models using goodness-of-fit measures. To investigate which model best reflects the underlying biology, the kinetic parameters obtained from the model fits are correlated with gene expression data measured from the patient biopsies.

Finally, chapter 7 presents a summary of the main conclusions and potential avenues of future work arising from this thesis.

1.4 Journal and Conference Papers

The research presented in this thesis has led to the publication of the following journal and conference papers:

- G. P. Ralli, D. R. McGowan, M. A. Chappell, R. A. Sharma, G. S. Higgins and J. D. Fenwick. 4D-PET reconstruction of dynamic non-small cell lung cancer [18-F]-FMISO-PET data using adaptive-knot cubic B-splines. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 1189-1192, 2017.
- G. P. Ralli, M. A. Chappell, D. R. McGowan, R. A. Sharma, G. S. Higgins and J. D. Fenwick. 4D-PET reconstruction using a spline-residue model with spatial and temporal roughness penalties. *Physics in Medicine & Biology*, 63(9):095013, 2018.

The work presented in chapter 6 is also intended for future publication.

Chapter 2

The PET Image Reconstruction Problem

2.1 PET Scanner Data

2.1.1 Data Acquisition and Storage

When two photons are detected by a pair of detectors within a given time window, they are considered to have originated from a single annihilation event. When this occurs, a prompt coincident event is recorded along the line connecting these two detectors. This line is called a line of response (LOR).

Scintillation detectors, consisting of a dense crystalline scintillator material that emits visible light in response to energy deposition from an incoming gamma ray, are the most widely used 511 keV photon detectors currently employed by PET scanners. This visible light is subsequently detected by a visible light photon detector, such as a photomultiplier tube, and converted into an electrical current.

While many PET scanners are unable to record the position of an annihilation event along a given LOR, modern time-of-flight (TOF) PET systems can achieve timing accuracy at the level of a few hundred picoseconds using fast scintillation crystals such as $LaBr_3$ and cerium-doped Lu_2SiO_5 (LSO) [24]. By recording the difference in arrival times of the coincident photons, TOF PET scanners can achieve event localization to within a few centimetres along the LOR.

PET scanners have a cylindrical geometry, and consist of multiple consecutive rings of detectors that surround the patient. For a given detector ring, the scanner

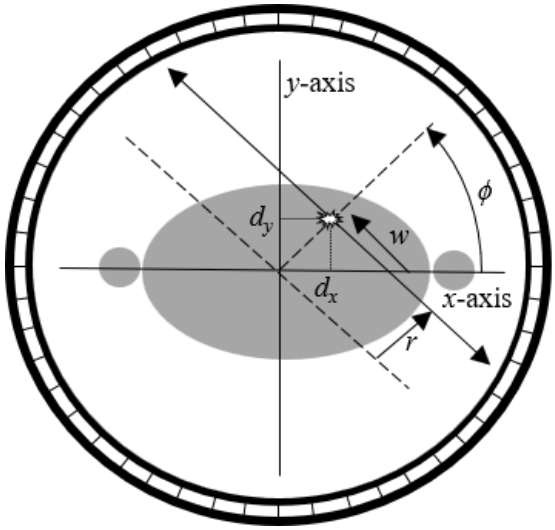


Figure 2.1. A single PET scanner detector ring with an annihilation event (occurring at the star) being by detected along the LOR depicted by a double-headed arrow. The relationship between the coordinates of the annihilation event in (d_x, d_y) image space and (r, w) space for a given projection angle ϕ is also shown.

data is usually histogrammed into a 2D matrix called a sinogram, with each matrix element corresponding to the number of events recorded along a particular LOR. The sinogram is arranged such that each row represents parallel line integrals through the activity distribution at a given angle (a projection), and each column represents the radial offset from the center of the scanner, r.

Consider a set of Cartesian coordinates with the origin at the centre of the scanner, and a photon pair travelling at projection angle ϕ and originating at a positron annihilation occurring at distances d_x and d_y along the x- and y-axes, respectively, as illustrated in Figure 2.1. This figure also shows that another set of (r, w) coordinates, defined as

$$r = d_x \cos(\phi) + d_y \sin(\phi), \tag{2.1}$$

and

$$w = d_y \cos(\phi) - d_x \sin(\phi), \tag{2.2}$$

can be used to describe the position of the annihilation. If r and w are known then d_x and d_y are given by:

$$d_x = r \cos(\phi) - w \sin(\phi), \tag{2.3}$$

and

$$d_y = r\sin(\phi) + w\cos(\phi). \tag{2.4}$$

The value of a given sinogram bin at radial offset r and projection angle ϕ, $s(r, \phi)$, is given by integrating the corresponding LOR along w through the activity distribution $a(d_x, d_y)$:

$$s(r, \phi) = \int_{-\infty}^{+\infty} a(r\cos(\phi) - w\sin(\phi), r\sin(\phi) + w\cos(\phi))dw. \tag{2.5}$$

This relationship is called the Radon transform [116]. The calculation of the sinogram from $a(d_x, d_y)$ via the Radon transform is called forward projection. Figure 2.2 shows an example of a 2D digital phantom and its corresponding sinogram, generated via the Radon transform with 200 angular projections. The Radon transform provides noise-free sinograms from a given activity distribution $a(d_x, d_y)$, though sinograms from real PET scans have Poisson-distributed noise. [116].

PET scans can be run in 2D and 3D acquisition mode. In 2D-mode acquisitions, tungsten septa are placed between each ring of detectors, thus only allowing events along LORs parallel to the plane of each detector ring to be recorded. In some 2D-mode aquisitions, cross-plane events between adjacent detector rings are also recorded. In 3D-mode acquisition the septa are removed, allowing oblique events along LORs between different detector rings to be recorded. While 3D-mode

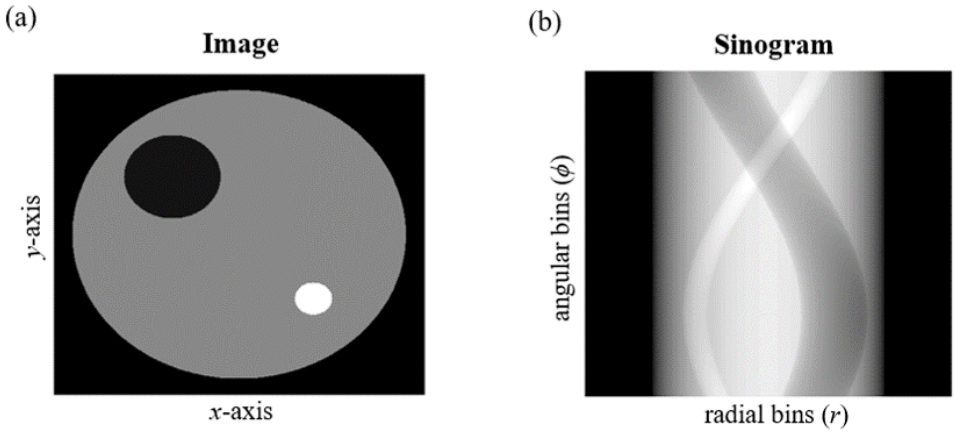


Figure 2.2. Noise-free digital phantom image (a) and corresponding sinogram with 200 angular projections (b).

acquisition substantially increases the size of the raw data files, it has been shown to improve the sensitivity by a factor of 5-7 compared to 2D-mode [23, 24, 155].

For 2D- and 3D-mode acquisitions, adjacent projection angles ϕ within the sinograms can be added together to compress the sinogram files. This compression technique is called angular mashing. An illustration of angular mashing in which each group of four adjacent projection angles are added together is shown in Figure 2.3. It can be seen from this figure that angular mashing corresponds to adding together adjacent sinogram rows. Note that for 3D-mode acquisitions, LORs with adjacent oblique angles can also be added together. The main drawback of angular mashing is that it reduces both the in-plane and - if mashing of the oblique angles is performed on a 3D-mode acquisition - the axial resolution at off-centre positions [32].

Most PET scanners also give the option of acquiring the data in listmode, where each coincidence event is recorded along with a time stamp that indicates when the event was detected. This results in a single file containing a list of coincidence events in the order that they arrived. This is particularly useful for dPET imaging, as it allows the scanner data to be re-binned into multiple time-frame sequences after the scan is completed.

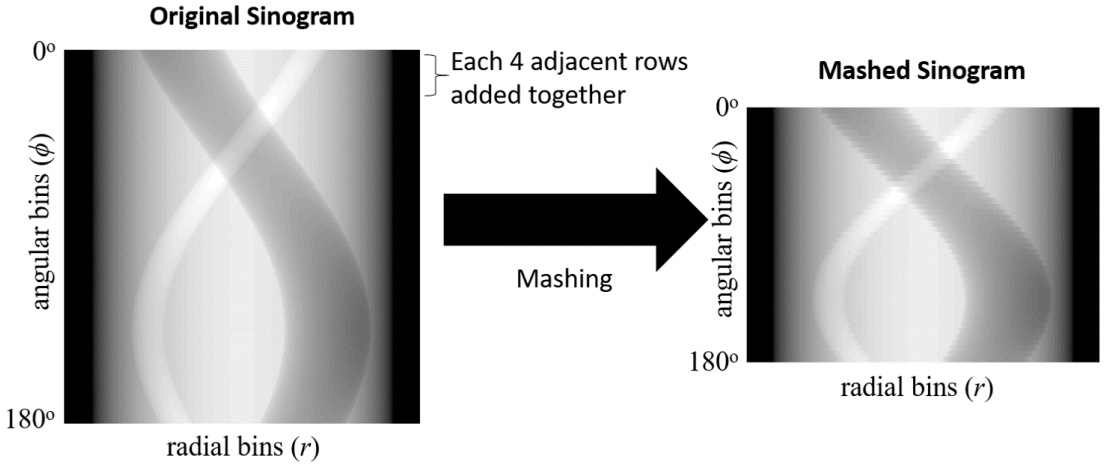


Figure 2.3. Illustration of angular mashing in which each group of four adjacent polar angles (and thus sinogram rows) are added together. Note that the mashed sinogram is smaller than the original because it contains a factor of 4 fewer rows, though in this figure it is shown as half the height of the original sinogram to make it easier to visualise.

2.1.2 Data Correction

To reconstruct an image volume that accurately reflects the true imaged activity distribution, a number of corrections need to be applied to the raw scanner data.

2.1.2.1 Normalization

Nonuniformities in individual detector efficiencies, scattering of photons within the detector crystals, the angle that the LOR makes with the detector faces and detector electronics all contribute to variations in coincidence detection efficiency between different LORs.

To correct for this, a uniform phantom can be scanned such that every LOR is illuminated by the same activity source. This can be achieved through scanning a uniform plane source of activity, positioned at 6 to 8 uniformly spaced projection angles [62]. The normalization correction coefficient (NCC) for a given LOR is then proportional to the inverse of the number of counts detected along it.

Alternatively, the NCCs can be calculated using a component based method, where the NCC for a given pair of detectors is given by:

$$NCC_{ab} = \frac{1}{\epsilon_a \epsilon_b g_{ab}}, \tag{2.6}$$

where ϵ_a and ϵ_b are the efficiencies of detectors a and b, respectively, and g_{ab} is a geometric factor [62, 116]. The geometric factor accounts for the angle of incidence of the LOR, systematic variations of crystal efficiency (which depend on the location of the crystal within the detector block) and relative plane efficiency. The detector efficiencies can be calculated by scanning a cylindrical phantom containing uniform activity [62, 17, 116]. Methods for calculating the g_{ab} factors are reviewed in [4, 152].

2.1.2.2 Detector Dead Time Correction

It takes a finite amount of time for a detector to process a photon signal and convert it to an electrical pulse. Any photons that hit the detector while another event is being processed will not be detected. This phenomenon is known as detector dead time. The most common way to correct for detector dead time is

to perform scans of uniform phantoms at a range of activity levels and then to fit empirical models of detector counting rate as a function of the true radioactivity concentration to the resulting data [27, 36, 45, 102].

2.1.2.3 Attenuation Correction

At 511 keV there is a relatively high probability that one or both of the coincident photon pairs will be attenuated by the patient tissues, as illustrated in Figure 2.4 (b), resulting in the non-detection of the coincidence event. The most accurate way to correct for this is with an attenuation map of the object being scanned. This map is a 3D image in which each voxel value is the linear attenuation coefficient, for a 511 keV photon, of the tissue contained within it. The attenuation correction factors (ACFs) can be calculated by taking line integrals through the attenuation map along each LOR.

The attenuation map can be reconstructed using data from a transmission scan. Here, the patient is placed in the scanner and a positron emitting rod source, usually containing ^{68}Ge, is rotated around them. The rod acts as an external source of 511 keV photon pairs produced from within the rod itself. Alternatively, for hybrid PET-CT systems the voxel intensities of the CT image, conventionally expressed in Houndsfield units, can be converted to linear attenuation coefficients for 511 keV

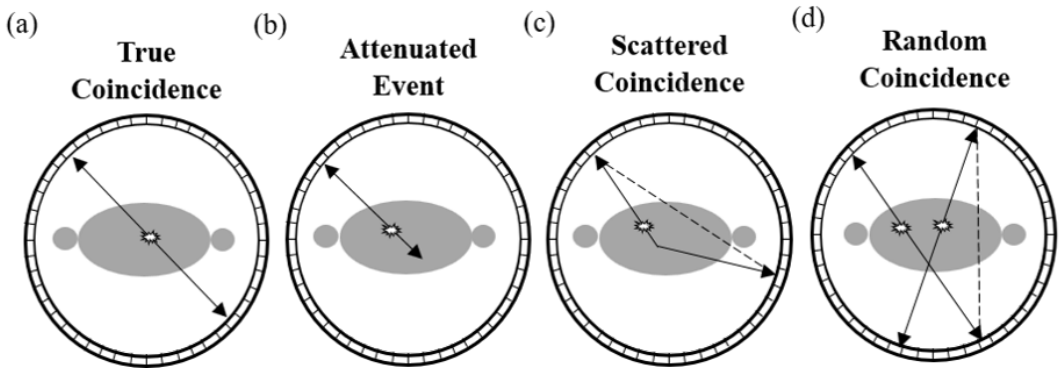


Figure 2.4. Illustrations of a true coincidence event (a), an attenuated event (b), a scattered coincidence (c) and a random coincidence (d). Solid arrows depict the photon paths and where appropriate the dashed line shows the erroneous LOR recorded by the scanner.

photons. The main advantages of using the CT image rather than a transmission scan are that the CT scan is faster and produces images of a higher statistical quality [24].

For hybrid PET-MRI scanners, the MR signal is not directly correlated to the attenuation coefficients and thus cannot be easily converted to a set of ACFs. Furthermore, the small bore inside MRI systems often won't accomodate a rotating PET transmission source. Various approaches for attenuation correction in PET-MRI imaging have been developed based on templates, atlas information, direct segmentation of T1-weighted MR images or segmentation of images from special MR sequences, and are reviewed in [166].

2.1.2.4 Correction of Scattered Coincidences

At 511 keV, the dominant mechanism of interaction between coincident photons and atomic electons within biological tissues is Compton scattering [24, 116]. In addition to causing photons to be attenuated, Compton scattering can also cause a photon to be deflected from its original path, leading to the event being recorded along the wrong LOR. This is called a scattered coincidence, and is illustrated in Figure 2.4 (c).

Because Compton scattering is the predominant cause of photon attenuation, the attenuation maps are approximately equivalent to maps of the coefficients for Compton scatter interactions [24]. Simulation-based scatter correction methods use the attenuation map together with an estimate of the activity distribution (usually an image reconstructed without scatter correction) to estimate the number of scatter coincidences. Widely used methods include the single scatter simulation algorithm [182, 183] and Monte-Carlo-based methods [63, 88]. Scatter correction can also be performed using convolution methods [5] and dual energy methods [53], though these are less accurate than simulation-based methods, especially for 3D-mode data acquisitions [116].

2.1.2.5 Correction of Random Coincidences

Two photons from different positron annihilations can reach their respective detectors within the coincidence time window, and thus be incorrectly recorded as a photon

pair originating along the LOR connecting these two detectors. Such an event is called a random coincidence, and is illustrated in Figure 2.4 (d).

It can be shown that the expected rate of randoms detection, $\langle R_{random} \rangle$, on a given LOR can be estimated by

$$\langle R_{random} \rangle = 2\tau R_{single,a} R_{single,b} \tag{2.7}$$

where τ is the length of the coincidence timing window and $R_{single,1}$ and $R_{single,2}$ are the rate of individual photon detection (singles) on the individual detectors a and b [116]. If the singles rate of each detector is recorded, then R_{random} can be calculated for each LOR.

Another method for estimating the number of random coincidences is the delayed coincidence method. Here, the timing signals on one of the detectors of the LOR are delayed by a time substantially higher than τ. This means that the two photons in each event recorded on this delayed coincidence channel must have originated from different positron annihilations, and the number of detected coincidences is a good estimate of the number of randoms in the prompt data. The advantage of this method is that the delayed channel has identical dead time properties to the prompt channel, though the statistical quality is worse than the randoms estimate provided by (2.7), since the singles detection rate is typically higher than the coincidence detection rate [116].

2.2 Analytical Image Reconstruction

2.2.1 Backprojection

An important component of analytical image reconstruction algorithms is backprojection, which provides a mapping from sinogram space to image space. Given a sinogram $s(r, \phi)$ with N_ϕ projection angles, the backprojected activity is defined as

$$\tilde{a}(d_x, d_y) = \frac{1}{N_\phi} \sum_{n=1}^{N_\phi} s(r, \phi_n) = \frac{1}{N_\phi} \sum_{n=1}^{N_\phi} s(d_x \cos(\phi_n) + d_y \sin(\phi_n), \phi_n), \tag{2.8}$$

where $\tilde{a}(d_x, d_y)$ is the value of the backprojected activity at a point (d_x, d_y) within the image, using the Cartesian coordinate system defined in Figure 2.1. Each voxel of a backprojected image is therefore given by the average of the counts along each LOR that passes through that voxel.

Figure 2.5 (b) shows a backprojected image from the noise-free sinogram shown in Figure 2.2 (b). Although backprojection does produce an image that resembles the true activity distribution, the result is very blurry due to counts being placed outside the object boundaries. It can be shown that the mathematical relationship between the backprojected image $\tilde{a}(d_x, d_y)$ and the true activity distribution $a(d_x, d_y)$ is given by

$$\tilde{a}(d_x, d_y) = a(d_x, d_y) \otimes \frac{1}{r}, \tag{2.9}$$

where $\otimes$ denotes the convolution operation [116].

2.2.2 Central Section Theorem

To resolve this blurring issue, a method based on the central section theorem (sometimes called the Fourier or projection slice theorem) can be used. This theorem states that the 1D-Fourier transform of a projection at angle ϕ_n is equal to the 2D-Fourier transform of the image, evaluated along a radial profile at angle ϕ_n from the x-axis:

$$S(\nu_r, \phi_n) = A(\nu_x, \nu_y)|_{\nu_x = \nu_r \cos(\phi_n), \nu_y = \nu_r \sin(\phi_n)}, \tag{2.10}$$

where ν_r is the spatial frequency in the radial direction; ν_x and ν_y are respectively the spatial frequencies in the x- and y-directions; $S(\nu_r, \phi_n)$ is the 1D-Fourier transform of $s(r, \phi_n)$; and $A(\nu_x, \nu_y)$ is the 2D-Fourier transform of the activity distribution $a(d_x, d_y)$ [184].

The image can therefore be reconstructed by taking the 1D-Fourier transform of each sinogram row and interpolating the result onto a 2D rectangular grid $A(\nu_x, \nu_y)$, according to the above equation. The image $a(d_x, d_y)$ is obtained via the inverse Fourier transform of $A(\nu_x, \nu_y)$.

Interpolating $S(\nu_r, \phi_n)$, which is sampled along radial lines, onto the rectangular grid $A(\nu_x, \nu_y)$ is computationally intensive, and the final image is very sensitive to interpolation errors [116]. Consequently, this method is rarely used in practice.

2.2.3 Filtered Backprojection

Filtered backprojection (FBP) is a commonly used analytical PET image reconstruction algorithm. This algorithm can be derived from the central section theorem.

The 2D inverse Fourier transform of $A(\nu_x, \nu_y)$ is given by:

$$a(d_x, d_y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\nu_x, \nu_y) \exp(2\pi i(\nu_x d_x + \nu_y d_y)) d\nu_x d\nu_y. \tag{2.11}$$

Transforming to polar coordinates gives $\nu_x = \nu_r \cos(\phi)$ and $\nu_y = \nu_r \sin(\phi)$. Restricting ν_r and ϕ to $-\infty \leq \nu_r \leq \infty$ and $0 \leq \phi \leq \pi$, respectively, implies that $d\nu_x d\nu_y = |\nu_r| d\nu_r d\phi$ [184]. Applying this to the above integral, and substituting $S(\nu_r, \phi)$ for $A(\nu_x, \nu_y)$, as permitted by the central section theorem in (2.10), gives:

$$a(d_x, d_y) = \frac{1}{(2\pi)^2} \int_{0}^{\pi} \int_{-\infty}^{\infty} S(\nu_r, \phi) \exp(2\pi \nu_r i(d_x \cos(\phi) + d_y \sin(\phi))) |\nu_r| d\nu_r d\phi. \tag{2.12}$$

From equation (2.1) $r = d_x \cos(\phi) + d_y \sin(\phi)$, meaning:

$$\begin{aligned} a(d_x, d_y) &= \frac{1}{(2\pi)^2} \int_{0}^{\pi} \int_{-\infty}^{\infty} S(\nu_r, \phi) \exp(2\pi i r \nu_r) |\nu_r| d\nu_r d\phi \\ &= \frac{1}{2\pi} \int_{0}^{\pi} \tilde{s}(r, \phi) d\phi, \end{aligned} \tag{2.13}$$

where $\tilde{s}(r, \phi)$ is a filtered sinogram:

$$\tilde{s}(r, \phi) = F^{-1} \left[S(\nu_r, \phi) |\nu_r| \right], \tag{2.14}$$

and $F^{-1}[\cdot]$ denotes the 1D inverse Fourier transform with respect to ν_r.

Equation (2.13) represents backprojection with continuous ϕ [184], though real PET sinogram data contains a discrete number of N_ϕ projection angles. Re-writing equation (2.13) using the discrete backprojection defined in equation (2.8) gives:

$$a(d_x, d_y) = \frac{1}{N_\phi} \sum_{n=1}^{N_\phi} \tilde{s}(r, \phi_n). \tag{2.15}$$

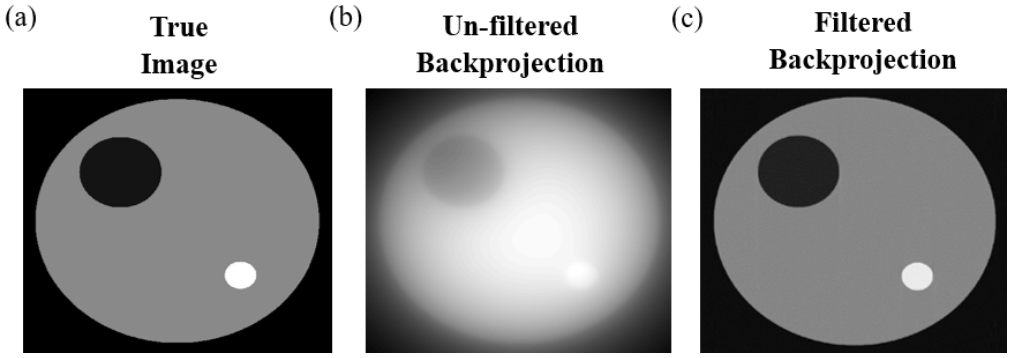


Figure 2.5. Analytical reconstructions of noise free sinogram data generated from a digital phantom (shown in (a)). Images of the activity distribution reconstructed with simple and filtered backprojection are shown in (b) and (c), respectively.

FBP reconstruction proceeds by first calculating the filtered sinogram via (2.14) and then backprojecting this sinogram to get the activity distribution.

The ramp filter, $|\nu_r|$, acts as a high-pass filter, thus suppressing the low spatial frequency components that lead to blurring when using un-filtered backprojection. In practice, the ramp filter is set to zero for $|\nu_r| > \nu_{cut-off}$. The cut-off frequency is often given by the Nyquist sampling frequency, which represents the highest frequency that can be recorded in discretely sampled data:

$$\nu_{cut-off} = \frac{1}{2\Delta r}, \tag{2.16}$$

where Δr is the distance between radial samples in the sinogram [116].

Figure 2.5 (c) shows an image reconstructed from the sinogram in Figure 2.2 (b) using FBP with a ramp filter. The FBP-reconstructed image is clearly more representative of the true image than the one reconstructed with simple backprojection.

FBP reconstruction can only be applied to sinogram data obtained from 2D-mode data acquisitions. A widely used extension of the FBP method to 3D-mode data is the 3D-reprojection (3D-RP) algorithm [79]. Alternatively, the data from the 3D-mode acquisitions can be re-sampled into a set of 2D sinograms using methods such as single slice re-binning (SSRB) [28] or Fourier re-binning (FORE) [32], and reconstructed with standard FBP.

2.3 Iterative Image Reconstruction

While analytical algorithms like FBP and 3D-RP are computationally efficient, they have a number of drawbacks. The underlying mathematics accurately describes the relationship between a source distribution and noise-free projection data, but real scanner data is noisy. Analytical reconstruction algorithms also make a number of simplifying assumptions about the data acquisition process, such as modelling the LORs as line integrals between point detectors, and are unable to model any of the other degrading factors that occur in a PET scanner, such as intercrystal scatter, positron range and non-colinearity [116]. For these reasons, iterative image reconstruction algorithms which are designed to address these issues are becoming more widely used [24, 119].

The basic procedure of iterative image reconstruction is summarized in Figure 2.6. One begins with an initial estimate of the image, and forward projects it using a system model describing the scanner physics. The projection data calculated from the image estimate is then compared to the data obtained from the PET scan. An optimization algorithm is used to update the image estimate based on the difference between the measured and calculated scanner data. Not all iterative algorithms are globally convergent, and furthermore noise is often amplified at later iterations [119]. For these reasons iterative algorithms are usually terminated after a fixed number of iterations rather than run to convergence.

In addition to being able to incorporate a more sophisticated model of the scanner physics than analytical reconstruction algorithms, iterative algorithms are derived from statistical modelling of the underlying scanner data, and are thus better able to account for noise when reconstructing images [119]. In this section the main approaches to iterative PET image reconstruction are summarized.

2.3.1 System Model for Iterative Reconstruction

In order to perform the forward projection step, iterative reconstruction algorithms require a system model that describes the data acquisition process. The system models used are typically more sophisticated than the simple Radon transform

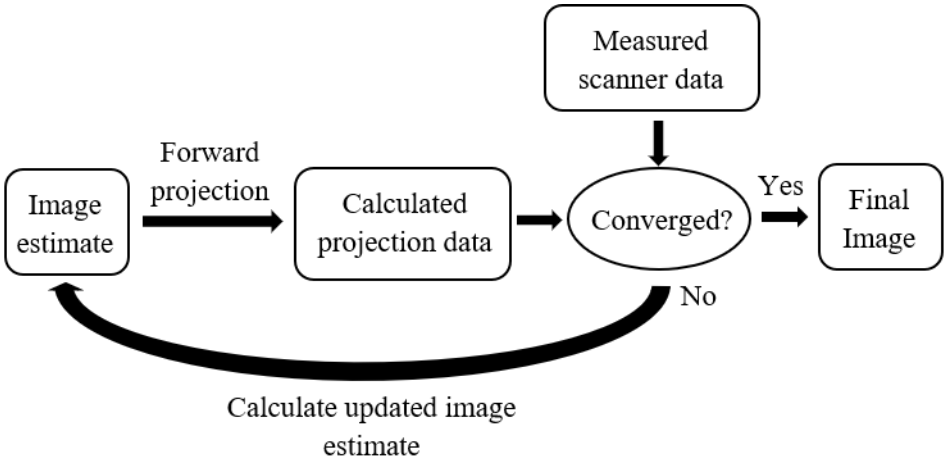


Figure 2.6. Flowchart summarizing the iterative reconstruction procedure. In practice most iterative algorithms are terminated after a fixed number of iterations as opposed to being run until convergence.

shown in equation (2.5), and model a wider range of physical effects that affect the scanner data.

A PET dataset with N_{LOR} LORs can be expressed as an N_{LOR}-dimensional vector $\boldsymbol{y}$, with each element y_i representing the counts measured along the i^{th} LOR, and the activity image can be expressed as a vector $\boldsymbol{x}$, with each element x_j representing the number of counts originating from voxel j. The expected scanner data can be calculated from the image estimate using

$$\langle y_i \rangle = \sum_{j=1}^{N_V} P_{ij} x_j + r_i + s_i, \tag{2.17}$$

where N_V is the number of voxels in the image, r_i and s_i are respectively the number of random and scattered coincidences along LOR i and P_{ij} are the elements of an $N_{LOR} \times N_V$ matrix $\boldsymbol{P}$ called the system matrix. The element P_{ij} represents the probability that a pair of photons originating in voxel j are detected along LOR i. The relationship between the expected scanner data and the image can also be expressed in matrix form:

$$\langle \boldsymbol{y} \rangle = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{r} + \boldsymbol{s}, \tag{2.18}$$

where $\boldsymbol{r}$ and $\boldsymbol{s}$ are N_{LOR}-dimensional vectors containing the elements r_i and s_i, respectively.

Equation (2.17) describes an alternative forward projection model to the simple Radon transform. Because the full system matrix is very large, Qi *et al.* [120] proposed an approach that utilizes the sparsity and symmetry properties of the system matrix elements to express it in a factorized form:

$$\boldsymbol{P} = \boldsymbol{P}_{det.sens}\boldsymbol{P}_{det.blur}\boldsymbol{P}_{attn}\boldsymbol{P}_{geom}\boldsymbol{P}_{positron}, \tag{2.19}$$

where $\boldsymbol{P}_{positron}$ models the blurring effects due to the positron travelling a finite distance before annihilating with an electron (for ^{18}F this distance is sub-millimetre, and so for tracers such as FDG and FMISO this is often ignored); $\boldsymbol{P}_{geom}$ describes the mapping from the image to the scanner data; $\boldsymbol{P}_{attn}$ contains the attenuation correction factors for each LOR; $\boldsymbol{P}_{det.blur}$ represents the blurring that occurs at the detector level due to photon noncolinearity, inter-crystal scatter and photon penetration; and $\boldsymbol{P}_{det.sens}$ contains the dead time and normalization correction factors.

The elements of $\boldsymbol{P}_{geom}$ are often calculated by taking line integrals along each LOR through the image voxel grid. This can be efficiently achieved via the ray tracing algorithm of Siddon [141], or an accelerated version of it [67, 191]. Such approaches assume that the PET scanner rings consist of point detectors, though real PET detectors have a finite surface area. Not accounting for the finite detector area results in a reduced resolution in the reconstructed images [135].

Resolution modelling methods that account for the point spread function (PSF) of the scanner in the system matrix calculation have been proposed. One method for PSF modelling is to incorporate an image space blurring kernel into the system matrix [122], though such approaches have been shown to introduce greater artefacts into the reconstructed images than methods that explicitly account for the finite detector areas when calculating $\boldsymbol{P}_{geom}$ [135]. To account for the finite detector areas, methods based on solid angle calculations [19, 120, 30], volumes of intersection between voxels and tubes of response (TORs) connecting the faces of each detector pair [96, 150, 137, 136, 135] and Monte-Carlo calculations [140, 161, 121, 144] have been proposed. Methods for system matrix calculation are reviewed in [1, 66].

Estimates of the numbers of random and scattered coincidences are usually calculated prior to the reconstruction. The methods described in section 2.1.2.5 provide good pre-reconstruction estimates of the number of random events. However, performing simulation-based scatter correction pre-reconstruction still requires an activity estimate. This is usually taken to be an image reconstructed without scatter correction, which is suboptimal because the activity estimate used already contains scattered events, though in practice it does provide reasonable scatter estimates [183]. A method that incorporates the single scatter simulation algorithm into the forward projection step of iterative reconstruction, thus incorporating scatter correction into the reconstruction, has been shown to provide improved scatter estimates [185].

2.3.2 Maximum-Likelihood Expectation-Maximization

The maximum-likelihood expectation-maximization (MLEM) algorithm of Shepp and Vardi [139, 159] and Lange and Carson [84] forms the basis of most iterative PET reconstruction algorithms.

The measured counts y_i can be accurately modelled as independent Poisson-distributed variables, $y_i \sim \text{Poisson}(\langle y_i \rangle)$. The probability of obtaining the measured scanner data from a given image estimate $\boldsymbol{x}$ is therefore given by the Poisson likelihood function:

$$P(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{N_{LOR}} \frac{\langle y_i \rangle^{y_i} \exp(-\langle y_i \rangle)}{y_i!}. \tag{2.20}$$

The goal of MLEM reconstruction is to find the image $\boldsymbol{x}$ that maximizes $P(\boldsymbol{y}|\boldsymbol{x})$.

For reasons of computational convenience, one can equivalently find the image $\boldsymbol{x}$ that maximizes the natural logarithm of the likelihood function (i.e. the log-likelihood function):

$$\ln(P(\boldsymbol{y}|\boldsymbol{x})) = \sum_{i=1}^{N_{LOR}} \left[y_i \ln(\langle y_i \rangle) - \langle y_i \rangle - \ln(y_i!) \right]. \tag{2.21}$$

Substituting (2.17) into to the above expression and removing the $\ln(y_i!)$ term, which doesn't depend on the image estimate, gives the objective function for

MLEM reconstruction:

$$L(\boldsymbol{y}|\boldsymbol{x}) = \sum_{i=1}^{N_{LOR}} \left[y_i \ln \left(\sum_{j=1}^{N_V} P_{ij} x_j + r_i + s_i \right) - \left[\sum_{j=1}^{N_V} P_{ij} x_j + r_i + s_i \right] \right]. \tag{2.22}$$

Despite the absence of the $\ln(y_i!)$ term, $L(\boldsymbol{y}|\boldsymbol{x})$ is usually referred to as the log-likelihood function in the PET image reconstruction literature.

MLEM uses the iterative expectation-maximization (EM) algorithm of Dempster *et al.* [34] to maximize $L(\boldsymbol{y}|\boldsymbol{x})$. The EM algorithm is designed to solve maximum-likelihood (ML) problems through the introduction of a set of 'complete data' which, if observed, would make the ML problem more easily solvable. For MLEM PET image reconstruction, the complete data is defined as a matrix of unobservable random variables $\boldsymbol{u}$ with element u_{ij} representing the total number of photon pairs from voxel j that travel along LOR i. The n^{th} EM iteration consists of calculating the conditional expectation of the log-likelihood of the complete data given the measured data and current image estimate $\boldsymbol{x}^n$ (the E-step) and the maximization of this conditional expectation with respect to the image (the M-step).

At iteration n, the E-step gives:

$$E\left[L(\boldsymbol{u}|\boldsymbol{x})|\boldsymbol{y}, \boldsymbol{x}^n\right] = \sum_{j=1}^{N_V} \left[x_j^n \sum_{i=1}^{N_{LOR}} \frac{P_{ij} y_i \ln(P_{ij} x_j)}{\sum_{k=1}^{N_V} \left[P_{ik} x_k^n + r_i + s_i\right]} - x_j \sum_{i=1}^{N_{LOR}} P_{ij} \right], \tag{2.23}$$

as shown in [139]. In the M-step the updated image $\boldsymbol{x}^{n+1}$ is calculated by maximizing $E\left[L(\boldsymbol{u}|\boldsymbol{x})|\boldsymbol{y}, \boldsymbol{x}^n\right]$ with respect to $\boldsymbol{x}$. Taking partial derivatives with respect to the voxel activities x_j gives

$$\frac{\partial}{\partial x_j} E\left[L(\boldsymbol{u}|\boldsymbol{x})|\boldsymbol{y}, \boldsymbol{x}^n\right] = \frac{x_j^n}{x_j} \sum_{i=1}^{N_{LOR}} \frac{P_{ij} y_i}{\sum_{k=1}^{N_V} \left[P_{ik} x_k^n + r_i + s_i\right]} - \sum_{i=1}^{N_{LOR}} P_{ij}. \tag{2.24}$$

The value of x_j^{n+1} is then found be setting the partial derivatives to zero and rearranging for x_j:

$$x_j^{n+1} = \frac{x_j^n}{\sum_{i=1}^{N_{LOR}} P_{ij}} \sum_{i=1}^{N_{LOR}} P_{ij} \frac{y_i}{\sum_{k=1}^{N_V} \left[P_{ik} x_k^n + r_i + s_i\right]}. \tag{2.25}$$

The MLEM algorithm proceeds by taking an initial image estimate (often a blank image with the radioactivity concentration in each voxel set to the same non-zero

value) and using equation (2.25) to update each image voxel. This procedure is repeated either for a fixed number of iterations or until the image has converged.

Because the MLEM update equation is multiplicative, if all the voxels in the initial image estimate are positive then it is not possible for the MLEM algorithm to assign negative values to any of the voxels. This non-negativity constraint has been shown to introduce positive bias into MLEM-reconstructed images in situations with low numbers of counts [168, 52]. The MLEM algorithm can be shown to be a scaled gradient ascent algorithm by rewriting equation (2.25) as:

$$x_j^{n+1} = x_j^n + \frac{x_j^n}{\sum_{i=1}^{N_{LOR}} P_{ij}} \sum_{i=1}^{N_{LOR}} P_{ij} \left(\frac{y_i}{\sum_{k=1}^{N_V} [P_{ik} x_k^n + r_i + s_i]} - 1 \right) = x_j^n + \Delta_j \frac{\partial}{\partial x_j} L(\boldsymbol{y}|\boldsymbol{x}), \tag{2.26}$$

with step size $\Delta_j = \frac{x_j^n}{\sum_{i=1}^{N_{LOR}} P_{ij}}$. Because the step size is proportional to the current activity concentration estimate, the updates in regions with low activity are small, thus slowing down the convergence in these regions. Large numbers of MLEM iterations can cause high variance in the reconstructed images, and so the reconstruction is often stopped before convergence. Since regions of lower activity converge more slowly, they are thus more likely to have higher bias levels when the reconstruction is terminated early. In dPET studies, this has been shown to translate into biased kinetic parameters as well [57, 170].

2.3.3 Ordered Subsets Expectation-Maximization

While the MLEM algorithm has been shown to produce higher quality images than analytical algorithms, this method takes many iterations to converge [24, 119]. To accelerate the convergence, Hudson and Larkin [65] proposed the ordered subsets EM (OSEM) algorithm, which is currently the most widely used iterative PET reconstruction algorithm [24].

In this methodology, the data is divided into N_s subsets. Each data subset, $\zeta_q \subseteq \{y_i\}_{i=1}^{N_{LOR}}$, is defined such than $\{y_i\}_{i=1}^{N_{LOR}} = \bigcup_{q=1}^{N_s} \zeta_q$ and $\zeta_a \cap \zeta_b = \emptyset$ for $a \neq b$. The MLEM update is applied using each subset in turn (meaning that each complete OSEM iteration involves N_s sub-iterations). At OSEM iteration n, the initial image

estimate used for the first sub-iteration is defined as the final image from iteration $n-1$, $x_j^{n,0} = x_j^{n-1}$. For a given sub-iteration, using subset ζ_q, the update equation is:

$$x_j^{n,q} = \frac{x_j^{n,q-1}}{\sum_{i\in\zeta_q} P_{ij}} \sum_{i\in\zeta_q} P_{ij} \frac{y_i}{\sum_{k=1}^{N_V} \left[P_{ik} x_k^{n,q-1} + r_i + s_i\right]}. \tag{2.27}$$

Once all of the sub-iterations are complete, the updated image is defined as $x_j^{n+1} = x_j^{n,N_s}$. The time taken to perform all N_s sub-iterations is approximately equal to the time required for one MLEM iteration [65].

To avoid artefacts in the reconstructed images, the projections are usually chosen to have maximum separation angle in each subset. Hudson and Larkin [65] also recommended that the subsets be chosen such that $\sum_{i\in\zeta_q} P_{ij}$ is independent of q, meaning that a coincidence event from a given voxel has an equal probability of being detected in each of the subsets. This is known as subset balance, and is difficult to achieve in practice due to spatially varying detector efficiencies and photon attenuation.

OSEM has been shown to provide substantial improvements in convergence rates when a reasonable number of subsets are used [65, 119]. The main disadvantage of OSEM is that if $N_s > 1$ then it does not generally converge to the ML solution. Furthermore, OSEM has been observed in some cases to enter limit cycles after a certain number of iterations [15]. Often, however, OSEM is terminated after a low number of iterations to limit the noise inflation that typically occurs at late iterations, and is thus less affected by this issue. Modified versions of OSEM with improved convergence properties have also been proposed [11, 2, 64, 78].

A further issue with OSEM is that in situations with low numbers of counts it is possible that dividing the data into subsets may result in one or more subsets not containing any counts. Applying equation (2.27) with a subset containing zero counts can lead to the voxel being erroneously assigned an activity concentration of zero. The multiplicative nature of the update equations means that once a voxel is assigned a value of zero it remains trapped there. This zero-trapping issue in situations with low number of counts can lead to negatively-biased images [168, 70]. OSEM-reconstructed dPET image sequences are particularly susceptible

to low-count bias, since dividing the scanner data into time-frames can result in individual time-frames containing low numbers of coincidence events.

2.3.4 Bayesian Image Reconstruction

A problem with all iterative reconstruction algorithms designed to maximize $L(\boldsymbol{y}|\boldsymbol{x})$ is that the underlying ML problem is ill-conditioned. This means that small changes in the data can lead to large changes in the reconstructed images, resulting in image estimates with high variance [119]. Bayesian PET image reconstruction algorithms, also known as maximum *a posteriori* (MAP) algorithms, are designed to address this issue through the use of regularization.

The posterior probability distribution, $P(\boldsymbol{x}|\boldsymbol{y})$, is defined in terms of the likelihood and the prior probability, $P(\boldsymbol{x})$, using Bayes' rule:

$$P(\boldsymbol{x}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{x})P(\boldsymbol{x})}{P(\boldsymbol{y})}. \tag{2.28}$$

Taking the log of the above expression, and removing any constant terms that don't depend on the image $\boldsymbol{x}$, gives:

$$\Phi(\boldsymbol{x}) = L(\boldsymbol{y}|\boldsymbol{x}) + \ln(P(\boldsymbol{x})). \tag{2.29}$$

The most commonly used prior in MAP reconstruction algorithms is the Gibbs prior

$$P(\boldsymbol{x}) = \frac{\exp(-\beta U(\boldsymbol{x}))}{Z}, \tag{2.30}$$

where β is a constant number, $U(\boldsymbol{x})$ is an 'energy function' and Z is a normalization constant. Substituting this prior into (2.29), and removing the constant $\ln(Z)$ term, gives the regularized objective function:

$$\Phi(\boldsymbol{x}) = L(\boldsymbol{y}|\boldsymbol{x}) - \beta U(\boldsymbol{x}). \tag{2.31}$$

Instead of deriving the regularized objective function using a Bayesian framwork, some authors view it as a penalized likelihood problem [119]. In practice, however, both approaches result in the same objective function.

The energy function $U(\boldsymbol{x})$ is usually treated as a penalty function for spatial roughness:

$$U(\boldsymbol{x}) = \sum_{j=1}^{N_V} \sum_{k \in \mathcal{N}_j} \omega_{jk} \psi(x_j - x_k), \tag{2.32}$$

where $\mathcal{N}_j$ is the set of nearest neighbours of voxel j, $\psi(\kappa)$ is a convex function of the difference between neighbouring voxels and ω_{jk} is a weighting factor equal to the normalized inverse distance between voxels j and k. When $U(\boldsymbol{x})$ is taken as a spatial roughness penalty, the tunable parameter β controls the trade-off between resolution loss due to over-smoothing and noise reduction.

A commonly used penalty function is the quadratic penalty $\psi(\kappa) = \kappa^2$. While this function is effective at smoothing out noise in regions of fairly uniform activity, it can also blur the edges between different regions. Conversely, the absolute function, $\psi(\kappa) = |\kappa|$, has better edge preservation properties than the quadratic function, but is less effective at penalizing noise in uniform regions. Another function, proposed by Lange [83], is defined as

$$\psi(\kappa) = \left(\frac{|\kappa|}{\delta} - \ln\left(1 + \frac{|\kappa|}{\delta}\right) \right) \delta, \tag{2.33}$$

where δ is an additional smoothing parameter. This function is approximately quadratic when $|\kappa| \ll \delta$ (i.e. in fairly uniform regions) and approximates the absolute function when $|\kappa| \gg \delta$ (i.e. when close to a boundary between two objects). Provided δ is well chosen, this function is therefore able to combine the advantages of the quadratic penalty in uniform regions with the superior edge preservation of the absolute penalty. Penalty functions of the form $\psi(\kappa) = |\kappa|^p$, with $1 < p < 2$, have also been proposed [9, 10]. Furthermore, non-convex penalty functions (e.g. [3, 44]), have been proposed, though these increase the chances of the reconstruction getting trapped in a local maximum, and are therefore not commonly used.

The EM algorithm can be used to maximize the objective function in (2.31). The M-step produces an image that maximizes the expectation of the MAP

objective function:

$$\boldsymbol{x}^{n+1} = \max_{\boldsymbol{x}} \left\{ \sum_{j=1}^{N_V} \left[x_j^n \sum_{i=1}^{N_{LOR}} \frac{P_{ij} y_i \ln(P_{ij} x_j)}{\sum_{k=1}^{N_V} [P_{ik} x_k^n + r_i + s_i]} - x_j \sum_{i=1}^{N_{LOR}} P_{ij} \right] - \beta U(\boldsymbol{x}) \right\}. \tag{2.34}$$

A commonly used method for maximizing the above expression is the one-step-late (OSL) algorithm of Green [51], in which the partial derivatives of $U(\boldsymbol{x})$ are evaluated at the current image estimate $\boldsymbol{x}^n$, giving the following update equation:

$$x_j^{n+1} = \frac{x_j^n}{\sum_{i=1}^{N_{LOR}} P_{ij} + \beta \frac{\partial}{\partial x_j} U(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}^n}} \sum_{i=1}^{N_{LOR}} P_{ij} \frac{y_i}{\sum_{k=1}^{N_V} [P_{ik} x_k^n + r_i + s_i]}. \tag{2.35}$$

The main disadvantage of OSL is that it doesn't generally converge to the MAP solution [83]. Many MAP algorithms with superior computational efficiency and convergence properties have been proposed, with [130, 119, 87, 108] providing good reviews.

2.4 4D Reconstruction for Dynamic PET

In dPET imaging the scanner data is conventionally divided into time-frames and each frame is reconstructed as a separate image, using either iterative or analytical reconstruction algorithms. Dividing the scanner data into time-frames results in a low number of counts in each frame, which in turn leads to a poor SNR in the reconstructed images. Furthermore, the ill-conditioned nature of the PET image reconstruction problem means that the noise in the reconstructed images is both amplified and can no longer be modelled as following a Poisson distribution [7, 118]. These issues degrade the quality of parameters obtained from fitting tracer kinetic models from TACs obtained from conventionally reconstructed dPET image sequences.

4D-PET reconstruction extends the iterative framework described in the previous section to incorporate the fitting of smooth temporal models to each voxel TAC during the reconstruction. This methodology acts to temporally regularize the dPET image sequences and suppress the noise amplification that occurs during the image reconstruction process. Furthermore, the raw scanner data can be accurately

modelled as being Poisson-distributed, thus allowing for improved model fits compared to post-reconstruction fitting. 4D-PET reconstruction was independently proposed by Snyder [142] and Carson and Lange [16], though neither of them tested their methods on clinical data or with realistic simulations. Subsequent work in the field, however, has demonstrated that such approaches have the potential to substantially improve the bias and noise properties of both the reconstructed images and the kinetic parameters [124, 129, 157, 178].

It should be noted that two other spatio-temporal reconstruction methods have been proposed to reduce noise in dPET image sequences. One method uses conventional iterative algorithms to reconstruct each time-frame, and applies spatio-temporal smoothing to the image sequence estimates after each iteration [104, 169]. The other method modifies the MAP framework described in section 2.3.4 to include temporal as well as spatial roughness penalties [50, 148]. These methods are sometimes referred to as 4D reconstruction algorithms as well. Both methods rely on the *ad hoc* assumption that voxels in neighbouring time-frames have similar intensities, which is not appropriate in situations with fast kinetics. While the the latter method is sometimes used in cardiac SPECT imaging [74], these methods are not often used for 4D-PET reconstruction, and are not considered further.

2.4.1 Extended Maximum-Likelihood Framework

A kinetic model $f(t, \boldsymbol{\theta})$ describes the radioactivity concentration as a function of time, t, and a set of model parameters $\boldsymbol{\theta}$. If $f(t, \boldsymbol{\theta})$ is used to model the decay-corrected activity in voxel j, then the number of counts originating from this voxel at time-frame m is given by:

$$x_{jm}(\boldsymbol{\theta}_j) = \int_{T_{ms}}^{T_{mf}} f(t, \boldsymbol{\theta}_j) \exp(-\lambda t) dt, \tag{2.36}$$

where T_{ms} and T_{mf} are respectively the start and end times of time-frame m, $\boldsymbol{\theta}_j$ are the model parameters for voxel j and λ is the radiotracer decay constant.

This can be substituted into the system model described in equation (2.17), thus giving the expected scanner data in time-frame m as a function of the

model parameters:

$$\langle y_{im}(\boldsymbol{\Theta})\rangle = \sum_{j=1}^{N_V} P_{ij} x_{jm}(\boldsymbol{\theta}_j) + r_{im} + s_{im}, \tag{2.37}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \ldots \boldsymbol{\theta}_{N_V}]^T$, and the T superscript indicates the matrix transpose. Here the system matrix is considered to be independent of time, however if it incorporates time-dependent effects such as detector dead time then one obtains matrix elements which also depend on time. Defining the dynamic scanner data with N_T time-frames as an $N_{LOR} \times N_T$ matrix $\boldsymbol{Y}$(with each row corresponding to a given LOR and each column a given time frame), the log-likelihood of the scanner data given a set of model parameters becomes

$$L(\boldsymbol{Y}|\boldsymbol{\Theta}) = \sum_{m=1}^{N_T} \sum_{i=1}^{N_{LOR}} \left[y_{im} \ln \left(\langle y_{im}(\boldsymbol{\Theta})\rangle \right) - \langle y_{im}(\boldsymbol{\Theta})\rangle \right]. \tag{2.38}$$

The parameters are generally found by iteratively maximizing $L(\boldsymbol{Y}|\boldsymbol{\Theta})$ with respect to $\boldsymbol{\Theta}$. MAP frameworks that use a modified version of the objective function (2.31):

$$\Phi(\boldsymbol{\Theta}) = L(\boldsymbol{Y}|\boldsymbol{\Theta}) - \beta U(\boldsymbol{\Theta}). \tag{2.39}$$

have been proposed to incorporate spatial regularization into the reconstruction (e.g. [75, 76, 172, 174]). Additionally, an objective function incorporating both temporal and spatial regularization, as well as non-negativity constraints, has been proposed in [89, 109]

$$\Psi(\boldsymbol{\Theta}) = L(\boldsymbol{Y}|\boldsymbol{\Theta}) - \beta U(\boldsymbol{\Theta}) - \gamma \Gamma(\boldsymbol{\Theta}) - \mu \rho(\boldsymbol{\Theta}), \tag{2.40}$$

where $\Gamma(\boldsymbol{\Theta})$ is a temporal roughness penalty function, $\rho(\boldsymbol{\Theta})$ penalizes negativity and γ and μ are additional tunable hyperparameters.

Many 4D reconstructions use linear kinetic models. These models represent TACs as linear combinations of pre-defined basis functions

$$f(t, \boldsymbol{\theta}) = \sum_{k=1}^{N_k} \theta_k B_k(t), \tag{2.41}$$

where N_k is the number of basis functions and $B_k(t)$ is the k^{th} basis function with coefficient θ_k. Because the basis functions are pre-defined, only the coefficients need to be found.

For linear models, Matthews *et al.* [100] proposed a 4D reconstruction method called parametric iterative reconstruction (PIR), which extends the MLEM framework to obtain a similar set of update equations for the θ_k parameters. Substituting the definition of the linear model into equation (2.36) shows that for linear models the image voxel value at a given time-frame is given by

$$x_{jm}(\boldsymbol{\theta}_j) = \sum_{k=1}^{N_k} \theta_{jk} B_{km}, \tag{2.42}$$

where the value of the k^{th} basis function in time frame m, B_{km}, is given by

$$B_{km} = \int_{T_{ms}}^{T_{mf}} B_k(t) \exp(-\lambda t) dt. \tag{2.43}$$

For linear models the expected scanner data is therefore given by

$$\langle y_{im}(\boldsymbol{\Theta}) \rangle = \sum_{j=1}^{N_V} \sum_{k=1}^{N_k} P_{ij} B_{km} \theta_{jk} + r_{im} + s_{im}, \tag{2.44}$$

or in matrix form

$$\langle \boldsymbol{Y}(\boldsymbol{\Theta}) \rangle = \boldsymbol{F}\boldsymbol{\Theta} + \boldsymbol{R} + \boldsymbol{S}, \tag{2.45}$$

where $\boldsymbol{R}$ and $\boldsymbol{S}$ are $N_{LOR} \times N_T$ matrices containing the elements r_{im} and s_{im}, respectively, and $\boldsymbol{F} = \boldsymbol{P} \odot \boldsymbol{B}$, with the matrix $\boldsymbol{B}$ containing the elements B_{km} and $\odot$ denoting the Kronecker product.

The above equation is very similar to the system model defined in equation (2.18) for iterative reconstruction of activity images. Following the procedure described in section 2.3.2 for deriving the MLEM update equation, but now for the θ_{jk} parameters, gives:

$$\theta_{jk}^{n+1} = \frac{\theta_{jk}^{n}}{\sum_{m=1}^{N_T} \sum_{i=1}^{N_{LOR}} P_{ij} B_{mk}} \sum_{m=1}^{N_T} \sum_{i=1}^{N_{LOR}} P_{ij} B_{mk} \frac{y_{im}}{\sum_{a=1}^{N_V} \sum_{b=1}^{N_k} P_{ia} B_{bm} \theta_{ab} + r_{im} + s_{im}}. \tag{2.46}$$

Like the MLEM algorithm, PIR suffers from slow convergence rates. Tsoumpas *et al.* [156] proposed a modified version of PIR called parametric OSEM (POSEM),

which uses the ordered subsets methodology described in section 2.3.3 to accelerate the convergence. Preconditioned conjugate gradient (PCG) algorithms have also been proposed for faster convergence of linear model-based reconstructions with objective functions in the forms shown in equations (2.38-2.40) [109, 89, 172].

To directly maximize the objective functions (2.38-2.40) for non-linear kinetic models, EM-based algorithms such as Parametric Motion-compensation OSEM List-mode Algorithm for Resolution-recovery for the 1-tissue compartment model (PMOLAR-1T) [187, 188, 46] and iterative coordinate descent algorithms such as the parametric iterative coordinate descent (PICD) algorithm [75, 76] have been proposed.

2.4.2 Acceleration with Nested Algorithms

An alternative methodology for improving the convergence rate of 4D reconstructions is via nested algorithms. These methods split each reconstruction iteration into two separate steps: an image update step with a conventional reconstruction algorithm such as MLEM followed by a step where the temporal model is fitted to each voxel TAC. This methodology was first proposed by Reader *et al.* [125], though they did not formally prove that this method maximizes $L(\boldsymbol{Y}|\boldsymbol{\Theta})$.

Using the method of optimization transfer [85], Wang and Qi [175] proposed a statistically derived nested algorithm called nested-EM. Here, at each iteration n, a minorizing surrogate function $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n)$ satisfying

$$Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n) \leq L(\boldsymbol{Y}|\boldsymbol{\Theta}), \tag{2.47}$$

with equality if and only if $\boldsymbol{\Theta} = \boldsymbol{\Theta}^n$, is maximized with respect to $\boldsymbol{\Theta}$. Defining $\boldsymbol{\Theta}^{n+1}$ as the value of $\boldsymbol{\Theta}$ that maximizes $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n)$ implies that:

$$\begin{aligned} L(\boldsymbol{Y}|\boldsymbol{\Theta}^{n+1}) &= Q(\boldsymbol{\Theta}^{n+1}|\boldsymbol{\Theta}^n) + L(\boldsymbol{Y}|\boldsymbol{\Theta}^{n+1}) - Q(\boldsymbol{\Theta}^{n+1}|\boldsymbol{\Theta}^n) \\ &\geq Q(\boldsymbol{\Theta}^n|\boldsymbol{\Theta}^n) + L(\boldsymbol{Y}|\boldsymbol{\Theta}^n) - Q(\boldsymbol{\Theta}^n|\boldsymbol{\Theta}^n) \\ &= L(\boldsymbol{Y}|\boldsymbol{\Theta}^n). \end{aligned} \tag{2.48}$$

Calculating and then maximizing $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n)$ at iteration n therefore guarantees that $L(\boldsymbol{Y}|\boldsymbol{\Theta})$ is monotonically increased.

Building on the work of Carson and Lange [16], Wang and Qi [175] defined the following surrogate function:

$$Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n) = \sum_{j=1}^{N_V} \left(\sum_{i=1}^{N_{LOR}} P_{ij} \right) \sum_{m=1}^{N_T} \left[\hat{x}_{jm}^{n+1} \ln(x_{jm}(\boldsymbol{\theta}_j)) - x_{jm}(\boldsymbol{\theta}_j) \right], \tag{2.49}$$

where $\hat{x}_{jm}^{n+1}$ is the MLEM update applied to the activity image in frame m given by the current parameter estimate, $x_{jm}(\boldsymbol{\theta}_j^n)$:

$$\hat{x}_{jm}^{n+1} = \frac{x_{jm}(\boldsymbol{\theta}_j^n)}{\sum_{i=1}^{N_{LOR}} P_{ij}} \sum_{i=1}^{N_{LOR}} P_{ij} \frac{y_i}{\sum_{k=1}^{N_V} \left[P_{ik} x_{km}(\boldsymbol{\theta}_k^n) + r_i + s_i \right]}. \tag{2.50}$$

In each iteration of the nested-EM algorithm, the activity images are first updated using equation (2.50). Next, because equation (2.49) is separable in voxels, the kinetic parameters in each voxel are calculated using

$$\boldsymbol{\theta}_j^{n+1} = \max_{\boldsymbol{\theta}_j} \left\{ \sum_{m=1}^{N_T} \left[\hat{x}_{jm}^{n+1} \ln(x_{jm}(\boldsymbol{\theta}_j)) - x_{jm}(\boldsymbol{\theta}_j) \right] \right\}. \tag{2.51}$$

As an alternative to maximizing equation (2.51), Matthews *et al.* [99] proposed fitting the model using weighted least squares (WLS)

$$\boldsymbol{\theta}_j^{n+1} = \min_{\boldsymbol{\theta}_j} \left\{ \sum_{m=1}^{N_T} w_{jm} \left[\hat{x}_{jm}^{n+1} - x_{jm}(\boldsymbol{\theta}_j) \right]^2 \right\}, \tag{2.52}$$

with weighting factors

$$w_{jm} = \frac{1}{x_{jm}(\boldsymbol{\theta}_j^n)}. \tag{2.53}$$

The advantage of the WLS approach is that standard WLS algorithms can be employed to perform the model fitting step. Wang and Qi have noted in [176] that monotonic convergence to the ML solution is not guaranteed when the model fitting step is converted to a WLS problem, however convergence has been observed in practice [99].

For linear kinetic models, an iterative equation for solving equation (2.51) can be derived [175]. Substituting the definition of a linear model, given in equation (2.42), into equation (2.51) gives

$$\boldsymbol{\theta}_j^{n+1} = \max_{\boldsymbol{\theta}_j} \left\{ \sum_{m=1}^{N_T} \left[\hat{x}_{jm}^{n+1} \ln \left(\sum_{k=1}^{N_k} \theta_{jk} B_{km} \right) - \sum_{k=1}^{N_k} \theta_{jk} B_{km} \right] \right\}. \tag{2.54}$$

This expression is directly analogous to the log-likelihood function $L(\boldsymbol{y}|\boldsymbol{x})$ defined in equation (2.22), and is maximized using N_l iterations of

$$\theta_{jk}^{n,l+1} = \frac{\theta_{jk}^{n,l}}{\sum_{m=1}^{N_T} B_{km}} \sum_{m=1}^{N_T} B_{km} \frac{\hat{x}_{jm}^{n+1}}{\sum_{a=1}^{N_k} B_{am} \theta_{ja}^{n,l}}. \tag{2.55}$$

Each iteration of the linear model-based nested-EM algorithm [175] consists of an image update using equation (2.50) followed by N_l sub-iterations of equation (2.55) to fit the model. The updated parameters are defined as $\theta_{jk}^{n+1} = \theta_{jk}^{n,N_l}$. Spatial regularization can be incorporated into the nested-EM reconstruction by replacing the image update step in equation (2.50) with a 3D-MAP image update step [175, 189].

Substituting the definition of $\hat{x}_{jm}^{n+1}$ into equation (2.55) shows that the PIR algorithm introduced in section 2.4.1 can be considered a special case of linear model-based nested-EM with $N_l = 1$. Running the nested-EM algorithm with multiple sub-iterations of (2.55) has been shown to result in faster convergence rates than the PIR algorithm [175]. Because the matrix $\boldsymbol{B}$ is considerably smaller than $\boldsymbol{P}$, the computational cost of equation (2.55) is substantially lower than the image update step, meaning that using extra sub-iterations doesn't greatly increase the time required for each nested-EM iteration. In the same paper, Wang and Qi [175] also showed that further improvements in the convergence rate can be achieved with a conjugate gradient algorithm, using the nested-EM update as an implicit preconditioner.

Equations (2.51) and (2.52) can both be used when fitting non-linear models in a nested 4D reconstruction framework, with $L(\boldsymbol{Y}|\boldsymbol{\Theta})$ as the objective function. For spatially regularized objective functions in the form of $\Phi(\boldsymbol{\Theta})$, given in equation (2.39), Wang and Qi [173] proposed an optimization transfer algorithm for non-linear models using paraboloidal surrogate functions. They later proposed an improved algorithm for non-linear models, which combines the nested-EM surrogate function in equation (2.49) with a separable surrogate function for the spatial roughness penalty, and maximizes this surrogate function using a modified Levenberg-Marquardt algorithm [176].

2.4.3 Temporal Models for 4D-PET Reconstruction

2.4.3.1 Compartment Models

Non-linear compartment models are often used in dPET studies. These models contain rate constants (sometimes referred to as 'micro-parameters') which are conceptually linked to specific biological processes, making them an attractive choice [55]. Because of this, many compartment model-based 4D reconstruction algorithms have been proposed, with [75, 188, 99, 173, 176, 22] being notable examples.

Compartment models describe the individual processes affecting the tracer uptake as separate compartments. The flow of tracer between compartments is characterised by individual rate constants. Figure 2.7 shows an example of a two-tissue compartment model. When PET image voxel values are given as radioactivity concentrations, which is often the case, the rate constant describing the uptake of the tracer from the blood to the first compartment is conventionally capitalized as K_1. This is because the units of K_1 are millilitre plasma (or blood) per minute per millilitre tissue (mL min^{-1} mL^{-1}), while the other rate constants have units min^{-1} [55].

For the two-tissue compartment model in Figure 2.7, the rate of change of tracer concentration within each compartment is given by

$$\frac{d}{dt}\begin{pmatrix} C_1(t) \\ C_2(t) \end{pmatrix} = \begin{pmatrix} -(k_2+k_3) & k_4 \\ k_3 & -k_4 \end{pmatrix}\begin{pmatrix} C_1(t) \\ C_2(t) \end{pmatrix} + \begin{pmatrix} K_1 \\ 0 \end{pmatrix} C_p(t), \tag{2.56}$$

where $C_n(t)$ is the tracer concentration in compartment n as a function of time, $C_p(t)$ is the AIF and the k-values are the rate constants shown in Figure 2.7. The full TAC, $f(t)$, is given by

$$f(t) = V_b C_p(t) + (1 - V_b)(C_1(t) + C_2(t)), \tag{2.57}$$



Figure 2.7. Schematic diagram of a two-tissue compartment model. The flow of tracer between compartments is characterised by rate constants (k-values).

where V_b is the fractional blood volume in the region the TAC was extracted from.

The rate equations for an arbitrary number of compartments is given by:

$$\frac{d}{dt}\boldsymbol{C}(t) = \boldsymbol{Q}\boldsymbol{C}(t) + \boldsymbol{b}C_p(t), \tag{2.58}$$

where $\boldsymbol{C}(t)$ is a column vector whose n^{th} element is $C_n(t)$, $\boldsymbol{Q}$ is a square matrix whose elements are functions of the rate constants and $\boldsymbol{b}$ is a column vector whose only non-zero element is its first, which is equal to K_1. Gunn *et al.* [55] showed that the general solution to the above equations is

$$f(t) = V_b C_p(t) + (1 - V_b)C_p(t) \otimes \sum_{q=1}^{N_c} \zeta_q \exp(-\eta_q t), \tag{2.59}$$

where N_c is the number of compartments, $\otimes$ denotes a convolution and ζ_q and η_q are free parameters which are defined as functions of the rate constants.

2.4.3.2 Linear Kinetic Models

Linear kinetic models, defined in (2.41), can be more flexible than compartment models, provided that the basis functions are well chosen. Commonly used linear models in 4D-PET reconstruction include graphical analysis methods such as the Patlak plot (e.g. [172, 156, 21, 77]), the spectal model (e.g. [100, 107, 126]) and spline functions (e.g. [89, 109, 162, 163]).

The Patlak model [113] is defined as

$$f(t, \boldsymbol{\theta}) = \theta_1 \int_0^t C_p(\tau)d\tau + \theta_2 C_p(t). \tag{2.60}$$

This model can be considered as an irreversible one-tissue compartment model (i.e. a compartment model where only K_1 is non-zero) where a fraction of the tracer enters the tissue during the scan and remains there throughout the remaining scan time [113].

The spectral model [25] is based on the general solution to the compartment model equations shown in equation (2.59). Each TAC is described as the AIF convolved with a weighted sum of N_k exponential decays, along with an extra term to account for the finite blood volume in the region:

$$f(t, \boldsymbol{\theta}) = \theta_0 C_p(t) + C_p(t) \otimes \sum_{k=1}^{N_k} \theta_k \exp(-\alpha_k t) \tag{2.61}$$

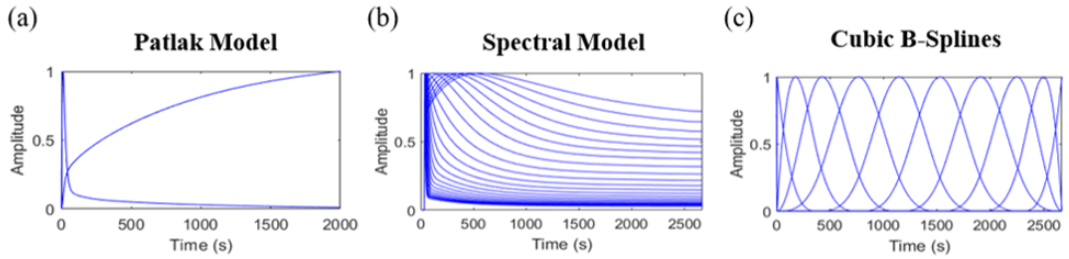


Figure 2.8. Comparison of typical shapes of normalized (a) Patlak, (b) spectral model and (c) cubic B-spline basis functions.

where $\otimes$ denotes a convolution and α_k are decay constants. The number and values of the exponential decay constants are defined in advance and held constant, meaning that only the θ_k coefficients are optimized during the model fitting.

Spline functions [29] are piecewise polynomial functions, joined continuously at sites called knots, and are often expressed as linear combinations of B-spline basis functions. Figure 2.8 compares the typical shapes of Patlak, spectral model and cubic B-spline basis functions.

Although graphical analysis methods and the spectral model allow macro-parameters, which are defined as functions of micro-parameters, such as the volume of distribution to be derived from the model fits, the micro-parameter values obtained from non-linear compartment model fits are conceptually more directly linked to specific biological processes.

B-spline basis functions do not contain any physiological information, and so spline-based 4D reconstruction can only be used to filter noise from the dPET image sequences, which can then be analysed with a physiologically motivated kinetic model post-reconstruction.

Another potential approach to linear-model based 4D-PET reconstruction is to use data-derived basis functions. Such methods make no *a priori* assumptions about the underlying TAC shapes, and because the resulting basis functions are tailored to each dataset they are likely to describe the underlying TACs well. Matthews *et al.* [100] proposed a method in which singular value decomposition (SVD) is applied to conventionally reconstructed dPET image sequences, providing a set of data-derived temporal basis functions for use in 4D-PET reconstruction. An issue

with this technique is that the resulting basis functions may contain negative values [100, 124, 129]. Reader *et al.* [127, 128] proposed an alternative methodology in which the temporal basis functions as well as their coefficients are estimated during the reconstruction process, allowing adaptive redefinitions of the basis functions as the reconstruction proceeds. Although this approach appears promising, Rahmim and Tang [124] note that it is not globally convergent. Because of the current limitations of using data-derived basis functions, they are less commonly used in 4D-PET reconstruction than the spline, spectral and Patlak models.

2.4.4 Spatial Bias Propagation

In situations with many tissues in the scanner FOV, which is often the case in cancer imaging studies, there can be a wide range of TAC shapes to be fitted. Kotasidis *et al.* [81] have shown that bias from poorly modelled regions can spatially propagate to well modelled regions during 4D reconstruction.

Although non-linear compartment models contain micro-parameters that are closely linked to specific biological processes, they also make specific assumptions about the behaviour of the radiotracer in the tissue being modelled. This limits the range of TAC shapes they can describe. In situations with diverse kinetics, it may therefore be better to perform a 4D reconstruction using flexible linear kinetic models such as splines or the spectral model to produce dPET image sequences with reduced noise, followed by post-reconstruction analysis with a compartment model.

An alternative technique for addressing this issue is called 'adaptive kinetic modelling'. Adaptive kinetic modelling fits a 'primary' model containing kinetic parameters of interest, and then a more flexible 'secondary' model is fitted to the residuals. The combined fitting of the primary and secondary model allows good descriptions of all TACs, even where the primary model fit is poor, while still allowing the kinetic parameters of interest to be obtained from the primary model fit. Adaptive kinetic modelling was first proposed by Matthews *et al.* [101], though they only tested the method on post-reconstruction dPET image

sequences. Kotasidis *et al.* [80] have since incorporated this method into a nested 4D reconstruction algorithm.

Both methods require a flexible temporal model that can describe a wide range of TAC shapes without fitting the noise: in the former to produce smooth images with low bias levels and in the latter to act as the secondary model. Although many temporal models have been proposed for 4D reconstruction, few studies have inter-compared their performance. The optimal choice of temporal model therefore remains an open question.

Chapter 3

Validation Framework for 4D-PET Reconstruction Algorithms

3.1 Introduction

The goal of 4D-PET reconstruction is to produce dPET image sequences and kinetic parameters with lower bias and noise than those obtained from conventional reconstruction methods. The most common method for evaluating the performance of novel 4D-PET reconstruction algorithms is by reconstructing simulated PET scanner data that has been generated from a digital phantom. Digital phantoms are computer-generated images that provide a known ground-truth to validate the reconstructed images against.

The 4D-PET reconstruction algorithms proposed in the next two chapters of this thesis have been validated using a digital thoracic phantom simulating an FMISO dPET scan of a stage IV non-small cell lung cancer (NSCLC) patient. This chapter describes the construction of this phantom, as well as the methods used to generate simulated PET data from it. Furthermore, in order to validate the simulation methodology, a digital phantom based on the National Electrical Manufacturers Association (NEMA) NU2 image quality phantom [151] is developed. The results from a simulated static PET scan of this NEMA-like phantom are compared to published data of a real static PET scan of a NEMA phantom using the same scanner model and phantom activities.

One method for producing simulated PET scanner data from a digital phantom is to use the system matrix to forward project the digital phantom via (2.17). Poisson-distributed noise can then be generated from the noise free sinograms. A more realistic approach is to use Monte-Carlo simulation software, which uses random sampling to simulate the radiation transport processes and the photon detection systems.

Commonly used Monte-Carlo software packages for producing simulated PET detector data from digital phantoms include PETSIM [153], SIMSET (Simulation System for Emission Tomography) [58], GATE (Geant4 Application for Tomographic Emission) [69], PENELOPET [37] (which is based on the PENELOPE (Penetration and ENErgy LOss of Positrons and Electrons) code [6]) and PET-SORTEO (Simulation of Realistic Tridimensional Emitting Objects) [132, 131, 105]. Reviews of some (though not all) of these Monte-Carlo simulation packages are available in [14, 190].

In this thesis the PET-SORTEO software was used to generate simulated PET data from the digital phantoms. PET-SORTEO allows the user to choose from a number of commercially available PET systems, for which the simulation software has been both written and validated, and produces PET detector data that is representative of what would have been produced by the scanner being simulated. Here the software was used to simulate a Siemens Biograph mMR PET-MR scanner, which is the most modern PET scanner that PET-SORTEO has been validated for to date [133].

3.2 Data Generation

3.2.1 NEMA NU2-Like Image Quality Phantom

A digital phantom based on the NEMA NU2 image quality phantom [151] was created to test the simulation and data correction methods used in this thesis. This phantom has a cylindrical volume of length 17.5 cm and diameter 22.4 cm, and contains six spheres of diameter 10, 13, 17, 22, 28, and 37 mm, arranged such that their centres are in the same axial plane and at the same distance from the centre

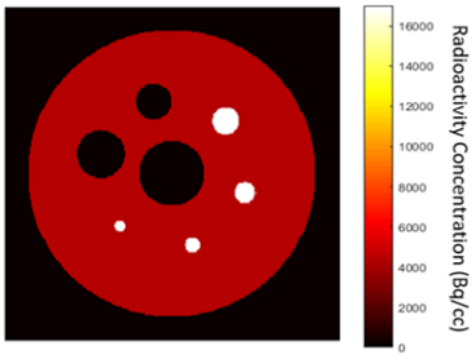


Figure 3.1. Axial slice of the NEMA-NU2-like digital image quality phantom. This slice shows the centres of each of the spheres that surround the activity-free cylindrical insert.

of the cylinder (5.72 cm). A cylindrical insert (length 17.5 cm, diameter 5 cm) was placed along the axis of the cylinder. For 511 keV photons, this insert has a linear attenuation coefficient of 0.034 cm^{-1} (similar to that of a lung). Other than the central cylindrical insert, the linear attenuation coefficient of water for 511 keV photons (0.096 cm^{-1}) was used for the rest of the phantom. It is important to note that the outer casing of a real NEMA NU2 image quality phantom has an irregular shape, to allow the phantom to be placed on the scanner bed without rolling.

The phantom was sampled onto a voxel grid with dimensions 2.0×2.0×2.0 mm^3. The smallest four spheres were given a radioactivity concentration of 17000 Bq/cc, the two largest spheres and the central cylindrical insert were given no radioactivity concentration and the radioactivity concentration in the rest of the cylinder was set to 4270 Bq/cc. These activity levels were chosen to match those used in a published study in which a real NEMA-NU2 phantom was scanned with an mMR scanner [33]. The spheres with non-zero activity concentration are henceforth referred to as the 'hot' spheres, while those containing no activity are referred to as 'cold' spheres. The activity concentrations used in this phantom correspond to a total radioactivity of 30 MBq.

Figure 3.1 shows the axial slice of the NEMA NU2-like digital phantom that contains the centres of each of the six spheres.

3.2.2 Dynamic FMISO NSCLC Phantom

The 4D-XCAT2 digital phantom package [138] was used to simulate a single slice of a dPET image of an NSCLC patient who had been injected with the FMISO hypoxia tracer. Tracer activity concentrations in different regions and time-frames were chosen to match temporally smoothed TACs taken from a clinical FMISO-dPET image of a patient with stage IV NSCLC.

3.2.2.1 Generation of TACs from clinical data

The clinical NSCLC dPET image was obtained from a dPET scan carried out on a GE Discovery 690 PET-CT scanner. The patient was injected with 370 MBq FMISO 30 seconds into PET imaging, which continued for 45 minutes and resumed for 10 minute intervals at two and four hours post-injection. The first 45 minutes of scanner data was binned into [1×30 s, 12×5 s, 6×10 s, 6×30 s, 10×60 s, 6×300 s] time-frames, and the two later 10 minute intervals were processed as single time-frames. Each frame was reconstructed onto a grid of 5.5×5.5×3.3 mm^3 voxels using a TOF-OSEM reconstruction algorithm (2 iterations, 24 subsets) provided by the scanner manufacturer, and 6.4 mm spatial Gaussian smoothing was applied post-reconstruction.

Lung and bone TACs were obtained from spherical ROIs of 3 cm diameter placed in healthy lung and spine regions in the real patient image, and an AIF TAC was taken from a cylindrical ROI of diameter 10 mm located in the centre of the descending aorta on five consecutive PET axial slices. Tumour TACs were obtained from irregularly-shaped ROIs considered to contain hypoxic and normoxic tumour tissue. The ROIs were drawn by an experienced radiologist and checked by a second radiologist.

For temporal smoothing, each TAC except the AIF was fitted with cubic splines, adaptively placing the knots according to the algorithm proposed in chapter 4, and with irreversible two- and three-tissue compartment models having 3 and 5 rate-constants respectively. These compartment models are schematically drawn in Figure 3.2 and subsequently referred to as 2C3K and 3C5K. The AIF TAC was

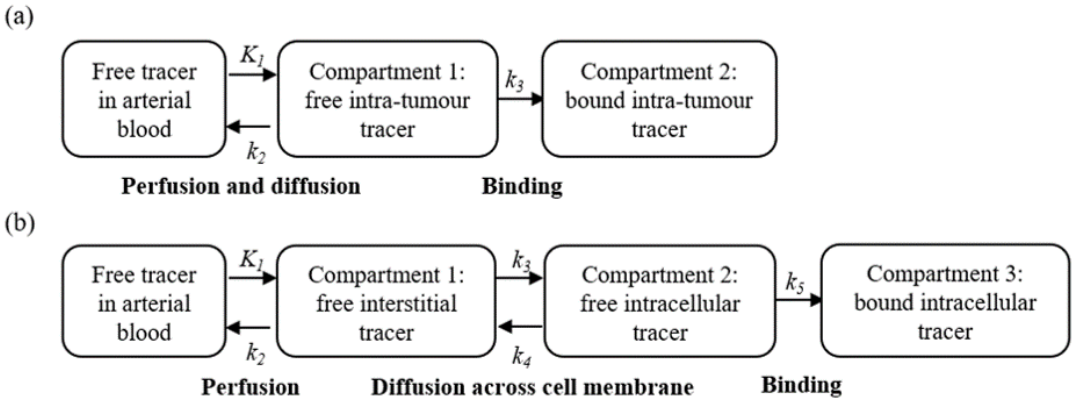


Figure 3.2. Schematic diagrams showing irreversible (a) two and (b) three tissue compartment models of FMISO uptake. Flows between compartments are defined by rate-constants (k values) and compartment tracer concentrations.

fitted with the phenomenological three-exponential model of Feng *et al.* [40] alone. Weighted least squares (WLS) was used for all fitting, with the weighting factors

$$w_m = \frac{\Delta T_m e^{-\lambda t_m}}{a_m}, \tag{3.1}$$

where a_m is the decay-corrected average activity concentration within the ROI during frame m, ΔT_m is the frame duration, t_m is the mid-point of the frame and λ is the radiotracer decay constant [20]. The WLS fits of the non-linear compartment models were carried out with the Levenberg-Marquardt algorithm available in the MATLAB optimization toolbox.

Model fit quality was assessed for each TAC using leave-one-out cross-validation, calculating the weighted residual sum of squares (RSS) error of the fit with the weighting factors defined in (3.1). The Wald-Wolfowitz runs test was used to check whether any significant structure remained in the residuals of each model fit at the 5% significance level. Of the models that passed the runs test for a given patient TAC, the best model fit was taken to be the one with the lowest weighted RSS value.

3.2.2.2 Digital dPET phantom

The patient phantom has voxel dimensions of 3.1×3.1×2.0 mm^3. The default settings of the XCAT phantom software package gives phantom voxel dimensions of 3.1×3.1×3.1 mm^3. Although the first two dimensions are already very similar

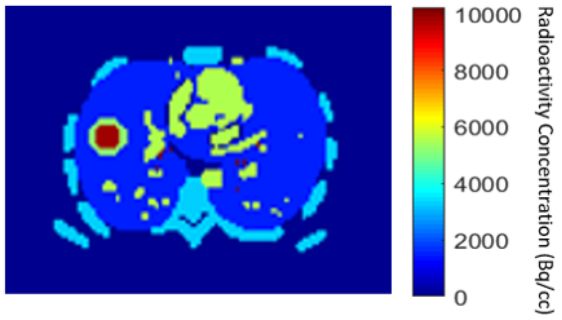


Figure 3.3. Illustration of the final time-frame of the FMISO NSCLC dPET phantom.

to those used in clinical PET scans, and were thus left the same here, the third dimension was re-set to 2.0 mm to allow the phantom slice to align with the mMR scanner detector rings.

Lung, bone, blood, normoxic and hypoxic tumour regions were filled with noise-free ground-truth activity concentrations that varied with time according to best model fits to the TACs obtained from the corresponding regions in the patient image. The fitted curves were binned into a [1×30 s, 6×5 s, 6×20 s, 7×60 s, 10×120 s, 3×300 s] time-frame sequence followed by two 10 minute frames at 2 and 4 hours post-injection. The FMISO NSCLC phantom geometry is illustrated in Figure 3.3, which shows the final time-frame of the phantom.

The reason for using fewer time-frames for the phantom than the clinical FMISO dPET image is that this coarser frame sequence was found to produce kinetic model fits that were very similar those fitted to TACs binned into the finer sequence used for the clinical dPET image. To illustrate this, Figure 3.4 shows adaptive-knot spline fits to the clinical hypoxic tumour TAC with the scan data binned into both frame sequences. The spline fits are compared here as they are more flexible than the compartment models, and thus more sensitive to changes in the underlying data. This figure shows that the spline fits to data binned into the two frame sequences are not substantially different to each other, though there are some differences in the early time-frames. Using fewer time-frames allows for faster reconstruction times, and so the coarser frame sequence was used for the phantom.

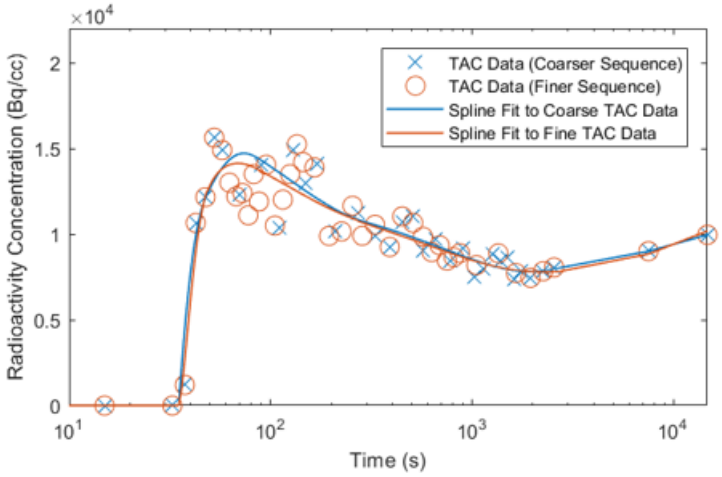


Figure 3.4. Spline fits to the clinical hypoxic tumour TAC. The TAC has been binned into the 'finer' time-frame sequence of the original clinical data ([1×30 s, 12×5 s, 6×10 s, 6×30 s, 10×60 s, 6×300 s] + 2×600 s frames 2- and 4-hours post-injection) and the 'coarser' frame sequence used for the dPET phantom ([1×30 s, 6×5 s, 6×20 s, 7×60 s, 10×120 s, 3×300 s] + 2×600 s frames 2- and 4-hours post-injection). Note that the first seven and last five time-frames of the two frame sequences are identical.

3.2.3 Generation of Monte-Carlo Simulated Data

Simulated PET detector data was generated from the two digital phantoms described in the previous subsection using the PET-SORTEO simulation software[1]. The software was used to produce sinograms representative of those produced by a Siemens Biograph mMR scanner. This scanner is made up of 8 rings of 56 detector blocks each. Each block is composed of 8×8 LSO crystal elements of dimension 4×4×20 mm^3, making up a total of 64 rings of 448 crystal elements [33]. The axial and transaxial dimensions of the scanner FOV are 25.8 cm and 59.4 cm, respectively. This scanner does not have TOF capabilities.

A 10 minute 3D-mode scan was simulated for the NEMA NU2-like image quality phantom. Following the standard mMR settings for 3D-mode scans, the maximum ring difference was set to 60. Axial mashing, where sinograms with adjacent oblique angles are added together, was performed to compress the data. Oblique LORs are those that connect detectors from different rings. The size of the groupings

[1]Although the PET-SORTEO software has been used to generate a library of digital phantom data for researchers to download, the software itself is not commonly available to download. The PET-SORTEO software was made available to me by A. Reilhac at CERMEP, Lyon, France, who originally developed the software.

of sinograms with adjacent oblique angles being added together (the 'span') was set to 11. For a given span and maximum ring difference, the total number of oblique samples stored is given by:

$$\text{Number of oblique samples} = \frac{2 \times \text{Max. Ring Difference} + 1}{\text{span}}. \tag{3.2}$$

For the settings used here this results in 11 oblique samples. The total data from the 3D-mode scan gives 837 sinograms arranged into the 11 oblique samples. The total number of detected counts in this simulated scan was 145 million.

For the FMISO NSCLC dPET phantom the simulated scans were run in 2D-mode. 50 noise realizations of dPET sinogram data were generated. On average, the total number of counts in each noise realization was 3.5 million. The scans of this phantom were performed in 2D-mode because generating and reconstructing multiple noise realizations of 3D-mode scan data for a dPET phantom with many time-frames is much more time consuming, and requires large amounts of disk space to store. Because the purpose of the FMISO NSCLC dPET phantom simulations is to study the noise-suppressing properties of 4D-PET reconstruction, as a simplifying assumption no patient motion was simulated.

3.3 Image Reconstruction and Analysis

3.3.1 Data Corrections

The PET-SORTEO software was used to perform the following data corrections:

- **Normalization.** Normalization correction is performed using the component-based method described in [17].
- **Detector dead time.** Dead time corrections were preformed using the method described in [131].
- **Attenuation** The attenuation correction factors for each phantom were calculated by taking line integrals along each LOR through the attenuation maps of the phantoms.

- **Scattered coincidences.** The number of scattered events were estimated using the single scatter simulation algorithm [183]. This algorithm uses the phantom attenuation map and an estimate of the phantom activity. The phantom activity estimate used here is an FBP-reconstructed image from the data with no scatter correction applied.
- **Random coincidences.** Each simulated scan included the simulation of a delayed coincidence window, which was used to estimate the number of random coincidences.

Further details of the way in which the PET-SORTEO software calculates the data correction factors are available in [131].

The ground-truth phantom attenuation maps were used for the attenuation and scatter corrections. This is somewhat unrealistic, since in reality a measured attenuation map would contain both bias and noise. However, the purpose of this project is to evaluate the performance of 4D-PET reconstruction algorithms. Because the same attenuation correction factors are used in each of the reconstruction algorithms being compared in this work, the fact that they are calculated in an idealized way does not give any one algorithm an advantage over the others.

Rather than pre-correct the data prior to reconstruction, the dead time, normalization and attenuation corrections were modelled as part of the system matrix, using the factorized form described in (2.19). The estimated random and scattered coincidences were added to the forward-projected image estimate at each reconstruction iteration, as shown in (2.18).

3.3.2 System Matrix Calculation

The calculations of the data correction components of the system matrix are described in the previous subsection. To calculate the geometric component, $\boldsymbol{P}_{geom}$, which describes the mapping from image to sinogram space, a method that calculates element $P_{geom,ij}$ as the volume of intersection between voxel j and tube of response (TOR) i was used. TORs are tubes of constant cross section that connect the faces of each detector pair, thus accounting for the finite surface area of the detector

faces, unlike LORs. This methodology has been shown to provide substantial improvements to image quality compared to more commonly used line integral based approaches, which estimate the sinogram by taking line integrals along each LOR through the activity distribution [135].

Because the full matrix $\boldsymbol{P}_{geom}$ is large, it is advantageous to calculate the elements on the fly, rather than pre-calculating and storing the whole matrix. However, calculating volumes of intersection between cuboidal TORs and voxels is a time-consuming process.

In this project a method proposed by Lougovski *et al.* [96] was used. This method approximates the TORs as cylinders and the voxels with spheres. The cross sectional area of the cylindrical TORs is equal to that of the detector faces, and the spherical voxels are defined such that their volume is equal to that of the real cuboidal voxels. The volume of intersection of a cylinder and a sphere depends only on the distance from the centre of the sphere to the cylinder axis [82], allowing efficient on the fly calculations of the matrix elements. Despite the approximations being made, Lougovski *et al.* [96] tested their method on data from two different PET scanners and found that it substantially outperformed the line integral-based reconstruction software provided by the two scanner manufacturers.

This method is similar to an earlier approach that approximates voxels as spherically symmetric volume elements ('blobs') made up of Kaiser-Bessel basis functions [98]. In a follow-up study Lougovski *et al.* [95] showed that their method provides slightly better results than the blob-based approach while also being easier to implement.

3.3.3 Image Reconstruction

3.3.3.1 MAP Reconstruction Algorithm

Spatially regularized images of both phantoms were reconstructed using the MAP methodology described in section 2.3.4. Recall that the MAP objective function is given by

$$\Phi(\boldsymbol{x}) = L(\boldsymbol{y}|\boldsymbol{x}) - \beta U(\boldsymbol{x}), \tag{3.3}$$

where β is a hyperparameter controlling the trade-off between image resolution and noise and $U(\boldsymbol{x})$ is a spatial roughness penalty

$$U(\boldsymbol{x}) = \sum_{j=1}^{N_V} \sum_{k \in \mathcal{N}_j} \omega_{jk} \psi(x_j - x_k), \tag{3.4}$$

where $\mathcal{N}_j$ is the set of nearest neighbours of voxel j, $\psi(\kappa)$ is a convex function of the difference between neighbouring voxels and ω_{jk} is a weighting factor equal to the normalized inverse distance between voxels j and k.

In this work the Lange function was chosen as the spatial roughness penalty function, as it provides a good trade-off between image resolution and noise, as discussed in section 2.3.4. The Lange penalty function is defined as:

$$\psi(\kappa) = \left(\frac{|\kappa|}{\delta} - \ln\left(1 + \frac{|\kappa|}{\delta}\right) \right) \delta, \tag{3.5}$$

where δ is an additional tunable parameter.

The MAP updates were performed using an algorithm proposed by Wang and Qi [177], which is itself based on an earlier algorithm proposed by De Pierro [31]. This method is derived using the method of optimization transfer, and provides a computationally efficient way of maximizing the MAP objective function $\Phi(\boldsymbol{x})$.

Each iteration of this algorithm proceeds in three steps. Firstly, the current image estimate at iteration n, x_j^n, is updated using the MLEM update equation defined in equation (2.25):

$$\tilde{x}_{j,EM}^{n+1} = \frac{x_j^n}{\sum_{i=1}^{N_{LOR}} P_{ij}} \sum_{i=1}^{N_{LOR}} P_{ij} \frac{y_i}{\sum_{k=1}^{N_V} [P_{ik} x_k^n + r_i + s_i]}. \tag{3.6}$$

Next, an intermediate regularized image is calculated from the current image estimate using:

$$\tilde{x}_{j,Reg}^{n+1} = \frac{1}{2H_j^n} \sum_{k \in \mathcal{N}_j} \omega_{jk} h_{jk}^{\psi}(z_{jk}^n)(x_j^n + x_k^n), \tag{3.7}$$

where $z_{jk}^n = x_j^n - x_k^n$, $h_{jk}^{\psi}(z_{jk}^n)$ is given by

$$h_{jk}^{\psi}(z_{jk}^n) = \frac{1}{|z_{jk}^n| + \delta}, \tag{3.8}$$

when using the Lange function (see [177] for details) and $H_j^n = \sum_{k \in \mathcal{N}_j} \omega_{jk} h_{jk}^{\psi}(z_{jk}^n)$.

Once $\tilde{x}_{j,EM}^{n+1}$ and $\tilde{x}_{j,Reg}^{n+1}$ are calculated, the updated image estimate is given by

$$x_j^{n+1} = \frac{2\tilde{x}_{j,EM}^{n+1}}{(1-\beta_j^n\tilde{x}_{j,Reg}^{n+1}) + \sqrt{(1-\beta_j^n\tilde{x}_{j,Reg}^{n+1})^2 + 4\beta_j^n\tilde{x}_{j,EM}^{n+1}}}, \tag{3.9}$$

where

$$\beta_j^n = \frac{\beta H_j^n}{\sum_{i=1}^{N_{LOR}} P_{ij}}. \tag{3.10}$$

The full derivation of this methodology is given in [177].

This methodology can also be extended to include the ordered subsets methodology described in section 2.3.3. To do so, in each iteration the three steps described in equations (3.6), (3.7) and (3.9) are applied to each subset in turn.

3.3.3.2 Image Reconstruction Settings

Images of both phantoms were reconstructed from the simulated data with the MAP algorithm described in subsection 3.3.3.1, using a range of β parameters. The settings used are described in Table 3.1. Reconstructions with $\beta = 0$ correspond to the standard OSEM reconstruction (or the MLEM reconstruction when only 1 subset is used). Because the 2D-mode simulation of the FMISO NSCLC phantom is noisier than that of the 3D-mode simulation of the NEMA NU2-like phantom, one more β value was used for this phantom. Here MAP reconstructions using more than one subset are referred to as OS-MAP reconstructions, while reconstructions with only one subset are simply referred to as MAP reconstructions.

The values of δ were kept at 0.1 in this study because preliminary experiments revealed that reconstructions with $\delta = 0.01, 1.00$ were not substantially different

Table 3.1. Settings used for the MAP reconstructions of the NEMA-NU2-like and FMISO NSCLC dPET phantoms.

Phantom	Number of Subsets	Number of Iterations	δ	β Values
NEMA-NU2-like Phantom	8	10	0.1	0, 0.0001, 0.001, 0.01, 0.1
FMISO NSCLC dPET Phantom	1	30	0.1	0, 0.0001, 0.001, 0.01, 0.1, 1

to those when $\delta = 0.1$. Furthermore, other studies have also found that $\delta = 0.1$ provides a good trade-off between edge preservation and noise suppression [177].

Ordered subsets were not used for the reconstructions of the FMISO NSCLC dPET phantom as the 4D-PET reconstruction algorithms that these non-4D reconstructions will be compared to in later chapters do not use ordered subsets. The subsets for the NEMA NU2-like phantom were chosen such that the angle between the projections contained in each subset was maximized.

Because axial mashing was used to compress the simulated 3D-mode NEMA phantom data, the same span-11 axial mashing procedure was applied to the expected scanner data produced by the forward projection step of each reconstruction iteration to allow the expected and measured data to be compared.

Each time-frame of the FMISO NSCLC dPET phantom was reconstructed onto a 2D 172×172 voxel grid with voxel dimensions 3.1×3.1×2.0 mm^3. Images of the NEMA NU2-like phantom were reconstructed onto a 3D 128×128×127 voxel grid with voxel dimensions 2.0×2.0×2.0 mm^3.

3.3.4 Image Analysis

3.3.4.1 NEMA NU2-Like Phantom

The image quality measures defined in the NEMA-NU2 (2012) literature [151] were used to compare the quality of the reconstructed NEMA NU2-like phantom images produced by different algorithms.

Spherical ROIs were drawn in each of the spheres, and for each of these six ROIs another ROI of equal size was drawn in the hot cylindrical region of the phantom that contains the spheres.

Contrast recovery coefficients (CRCs) were calculated for each of the spheres. For the hot spheres, the contrast recovery coefficient is defined as

$$CRC_{H,s} = \frac{\frac{C_s}{B_s} - 1}{\frac{C_{Ts}}{B_{Ts}} - 1} \times 100, \tag{3.11}$$

where C_s is the mean activity concentration in the ROI of sphere s, B_s is the mean activity concentration in the background ROI of the same size and C_{Ts} and

B_{Ts} are the true activity concentrations in the hot sphere and background ROIs, respectively. For the cold spheres, the contrast recovery coefficient is defined as

$$CRC_{C,s} = \left(1 - \frac{C_s}{B_s}\right) \times 100. \tag{3.12}$$

The background variability (BV) for all six of the background ROIs was calculated using

$$BV_s = \frac{\sigma_s}{B_s} \times 100, \tag{3.13}$$

where σ_s is the standard deviation in the measured activity concentrations in the voxels contained by the background ROI s.

The factor of 100 the CRC and BV calculations is used so that they can be expressed as percentages.

To assess the trade-off between contrast and noise, the contrast to noise ratio (CNR) was calculated for each of the spheres using

$$CNR_s = \frac{CRC_s}{BV_s}, \tag{3.14}$$

where CRC_s is either $CRC_{H,s}$ or $CRC_{C,s}$, depending on whether sphere s is a hot or cold sphere.

The above measures were calculated for each of the reconstructed images of the NEMA NU2-like phantom. For the OSEM-reconstructed images, this analysis was performed after both 4 and 10 iterations, the reason being that clinical OSEM reconstructions are often terminated before convergence to limit the noise amplification that occurs in later iterations [119]. Because OSEM does not provide any form of spatial regularization, the OSEM-reconstructed images were analyzed both with and without 4 mm Gaussian smoothing.

3.3.4.2 Dynamic FMISO NSCLC Phantom

To characterise the accuracy of the reconstructed images of the FMISO NSCLC phantom, the average absolute bias of imaged activity concentrations over the time-course of the scan was calculated for every voxel j:

$$[\text{Image Bias}]_j = \frac{1}{T_{scan}} \sum_{m=1}^{N_T} \Delta T_m |\bar{a}_{jm} - a_{jm}^{true}|, \tag{3.15}$$

where T_{scan} is the scan duration, ΔT_m is the duration of time-frame m, N_T is the total number of time-frames, $\bar{a}_{jm}$ is the mean activity concentration in voxel j at time-frame m in all 50 noise realizations and a_{jm}^{true} is the true activity concentration.

The noise in each voxel j at every time-frame m was calculated using the weighted standard deviation $\sigma_{w,jm}$

$$\sigma_{w,jm} = \left(\frac{\sigma_{jm,measured}^2 \Delta T_m}{a_{jm}^{true}} \right)^{\frac{1}{2}}, \tag{3.16}$$

where $\sigma_{jm,measured}^2$ is the measured variance of the 50 noise realizations of a_{jm} values, and the weighting factors $a_{jm}^{true}/\Delta T_m$ are based on the dPET noise model of Chen *et al.* [20] and nominally account for intrinsic variations in noise between frames. The average noise for voxel j was then calculated across all N_T time-frames:

$$[\text{Image Noise}]_j = \frac{1}{N_T} \sum_{m=1}^{N_T} \sigma_{w,jm}. \tag{3.17}$$

Overall bias and noise were characterised as the mean absolute bias and mean image noise ($\langle \sigma_w \rangle$) averaged over all voxels within the patient. To facilitate algorithm inter-comparison, normalised mean image noise and absolute bias values were computed at every iteration of all reconstructions.

3.4 Results

3.4.1 NEMA NU2-Like Phantom

OSEM-reconstructed images of the NEMA NU2-like phantom using 4 and 10 iterations are shown in Figure 3.5. This figure shows images both with and without Gaussian smoothing. OS-MAP reconstructed images with all the β values listed in Table 3.1 are shown in Figure 3.6. The axial slices shown in Figures 3.5 and 3.6 are the slice containing the centres of each of the spheres, like that of Figure 3.1.

In Figure 3.6 there appears to be a chequerboard-like pattern in the images reconstructed with high β. It is likely that this artefact is also present in the OSEM and low β OS-MAP reconstructions as well, though in those images it is obscured by the noise. The exact reason for this artefact is not obvious, although

since this artefact was not observed in the images of the clinical dynamic phantom, which was simulated in 2D-mode, it could be due to the handling of the oblique angles in the 3D-mode NEMA phantom data.

In order to check the convergence of the NEMA-like phantom reconstructions, for each of the reconstructions the MAP objective function given in equation (3.3) was calculated at each reconstruction iteration. The MAP objective function is plotted as a function of iteration number for each of the NEMA-like phantom reconstructions in Figure 3.7. Note that in this figure $\beta = 0$ corresponds to the OSEM reconstruction, and the MAP objective function with $\beta = 0$ reduces to the log-likelihood function. Figure 3.7 shows that the rate of convergence was slower for MAP reconstructions with higher β values, though by 10 iterations the objective function values of all the NEMA-like reconstructions had levelled off.

The CRCs and BVs of each sphere are shown in Tables 3.2 and 3.3 respectively. The standard deviations of the activities in the sphere and background ROIs were also measured, and with these the errors in the CRCs and BVs were calculated using standard error propagation formulae [151]. These errors are also shown in Tables 3.2 and 3.3. The CNRs of each sphere are shown in Table 3.4.

The optimal CRC values are obtained when using either OSEM with 10 iterations and no smoothing or OS-MAP with $\beta = 0.0001$ (which is itself very similar to OSEM with no smoothing). However, both of these methods result in high background variances, ultimately leading to poor CNR values. OSEM with four iterations and Gaussian smoothing yields the best contrast to noise ratios of the OSEM reconstructions attempted.

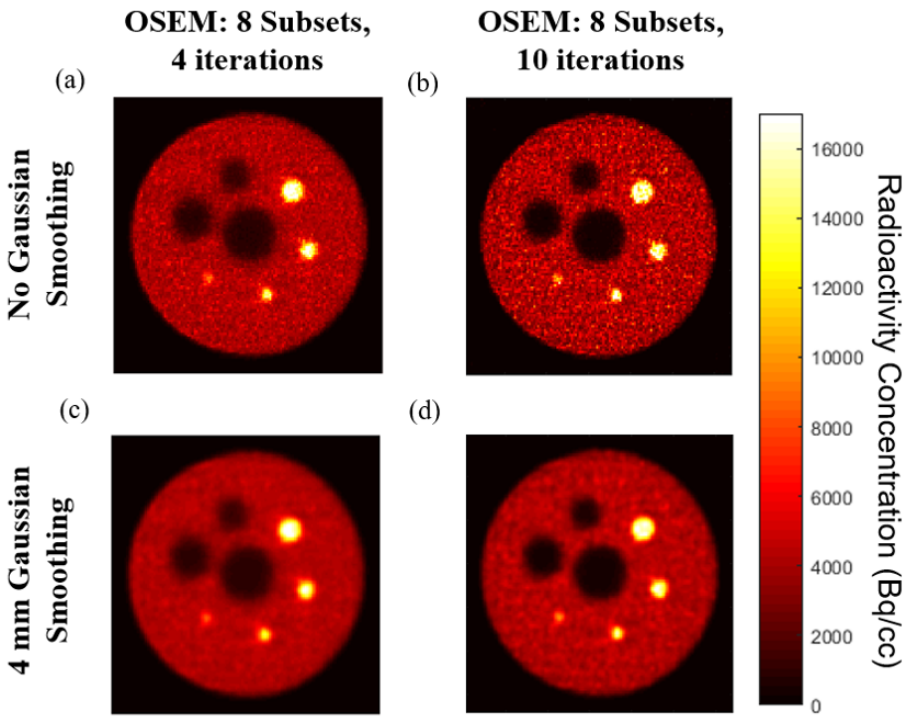


Figure 3.5. Reconstructed images of the NEMA NU2-like digital phantom using OSEM with 4 (a,c) and 10 (b,d) iterations. Images are shown before (a,b) and after (c,d) the application of Gaussian smoothing.

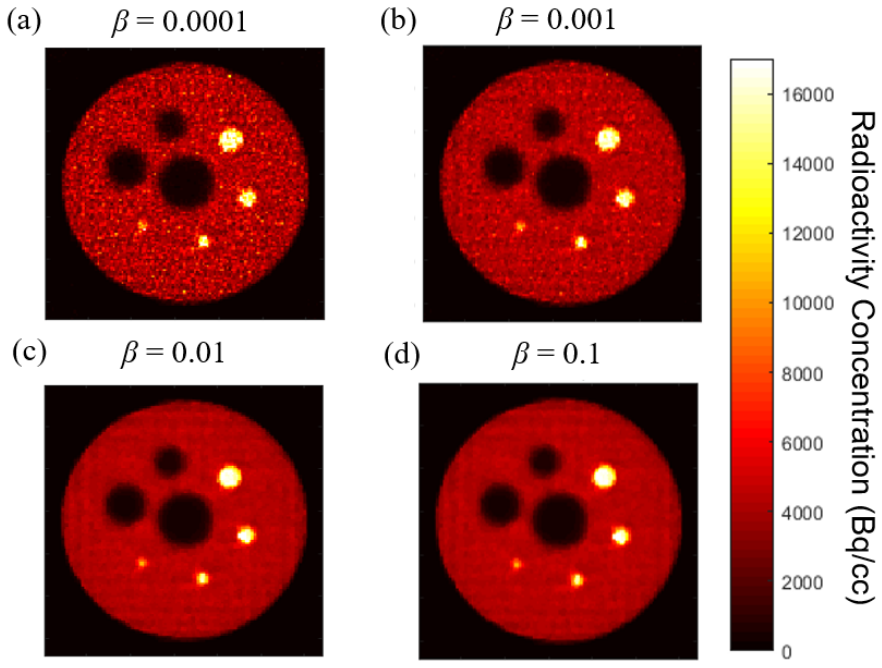


Figure 3.6. Reconstructed images of the NEMA NU2-like digital phantom using OS-MAP with $\beta = 0.0001$ (a), $\beta = 0.001$ (b), $\beta = 0.01$ (c) and $\beta = 0.1$ (d).

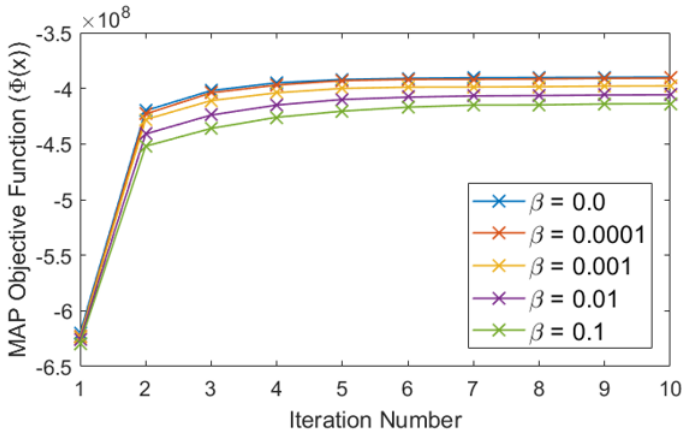


Figure 3.7. Plots of the MAP objective function defined in equation (3.3) vs iteration number for each of the NEMA-like phantom reconstructions. Note that the reconstruction with $\beta = 0$ corresponds to the OSEM reconstruction.

OS-MAP with $\beta = 0.01$ gives the best CNR values of the OS-MAP reconstructions (apart from the largest two spheres, where $\beta = 0.1$ performs better). It also yields higher CRC values than OSEM with four iterations and Gaussian smoothing. OS-MAP with $\beta = 0.01$ therefore gives superior resolution compared to OSEM with Gaussian smoothing, but at the expense of slightly higher noise.

Table 3.2. CRCs for each reconstruction of the NEMA NU2-like phantom. CRCs are expressed as percentages, with errors shown in brackets. The highest CRC for each sphere is highlighted in bold. The acronym GS stands for Gaussian smoothing.

Reconstruction Algorithm	Sphere Diameter (mm)					
	10	13	17	22	28	37
OSEM	42.6	58.9	62.8	74.0	66.8	73.7
(4 iterations)	(3.0)	(2.6)	(2.5)	(2.1)	(14.2)	(16.0)
OSEM+GS	35.0	51.7	60.8	68.0	66.8	73.7
(4 iterations)	(1.3)	(1.2)	(1.1)	(1.0)	(3.7)	(4.9)
OSEM	50.6	**66.0**	64.7	76.1	**83.8**	**87.8**
(10 iterations)	(7.3)	**(6.9)**	(6.4)	(6.4)	**(31.2)**	**(35.8)**
OSEM+GS	47.0	61.4	59.3	72.8	79.7	83.1
(10 iterations)	(3.3)	(3.3)	(3.0)	(12.8)	(12.8)	(15.7)
OS-MAP	**51.0**	58.5	**68.6**	**80.8**	76.3	79.6
($\beta = 0.0001$)	**(1.9)**	(1.7)	**(1.4)**	**(1.3)**	(2.9)	(3.2)
OS-MAP	49.4	57.1	66.4	80.2	75.9	79.3
($\beta = 0.001$)	(1.4)	(1.3)	(1.1)	(0.9)	(2.5)	(2.8)
OS-MAP	44.2	53.6	63.7	79.7	75.1	78.8
($\beta = 0.01$)	(1.0)	(1.1)	(0.7)	(0.5)	(1.3)	(1.8)
OS-MAP	40.8	51.5	60.2	79.2	73.4	78.6
($\beta = 0.1$)	(0.7)	(1.1)	(0.5)	(0.5)	(1.2)	(1.7)

Table 3.3. BVs for each reconstruction of the NEMA NU2-like phantom. BVs are expressed as percentages, with errors shown in brackets. The lowest BV for each sphere is highlighted in bold. The acronym GS stands for Gaussian smoothing.

Reconstruction	Sphere Diameter (mm)					
Algorithm	10	13	17	22	28	37
OSEM	18.3	17.8	17.4	16.3	14.3	13.0
(4 iterations)	(0.8)	(0.9)	(0.8)	(0.7)	(0.6)	(0.5)
OSEM+GS	**6.3**	**5.6**	**4.0**	**3.9**	3.6	3.8
(4 iterations)	**(0.2)**	**(0.2)**	**(0.1)**	**(0.1)**	(0.1)	(0.1)
OSEM	38.0	33.6	30.3	29.9	15.2	13.2
(10 iterations)	(6.2)	(6.0)	(5.0)	(5.2)	(2.5)	(2.2)
OSEM+GS	14.0	12.7	12.6	11.1	7.7	6.0
(10 iterations)	(1.2)	(1.2)	(1.0)	(1.0)	(0.6)	(0.5)
OS-MAP	11.9	11.5	10.7	9.8	6.3	6.0
($\beta = 0.0001$)	(3.4)	(3.0)	(2.9)	(2.4)	(0.3)	(0.3)
OS-MAP	10.0	9.9	8.2	7.0	4.6	4.4
($\beta = 0.001$)	(1.6)	(1.3)	(1.2)	(0.9)	(0.2)	(0.2)
OS-MAP	7.4	8.7	6.6	5.4	3.9	4.0
($\beta = 0.01$)	(0.3)	(0.3)	(0.2)	(0.2)	(0.2)	(0.2)
OS-MAP	6.9	8.7	6.8	5.4	**3.0**	**3.1**
($\beta = 0.1$)	(0.2)	(0.3)	(0.3)	(0.2)	**(0.4)**	**(0.4)**

Table 3.4. CNRs for each reconstruction of the NEMA NU2-like phantom. BVs are expressed as percentages, with errors shown in brackets. The highest CNR for each sphere is highlighted in bold. The acronym GS stands for Gaussian smoothing.

Reconstruction	Sphere Diameter (mm)					
Algorithm	10	13	17	22	28	37
OSEM	2.3	3.3	3.6	4.5	4.7	5.7
(4 iterations)	(0.1)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)
OSEM+GS	5.6	**9.2**	**15.2**	**17.4**	18.6	19.4
(4 iterations)	(0.1)	**(0.1)**	**(0.1)**	**(0.1)**	(0.1)	(0.1)
OSEM	1.3	2.0	2.1	2.5	5.5	6.7
(10 iterations)	(0.2)	(0.2)	(0.2)	(0.2)	(0.4)	(0.4)
OSEM+GS	3.4	4.8	4.7	6.6	10.4	13.9
(10 iterations)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)	(0.2)
OS-MAP	4.3	5.1	6.4	8.2	12.1	13.3
($\beta = 0.0001$)	(0.3)	(0.3)	(0.3)	(0.2)	(0.1)	(0.1)
OS-MAP	4.9	5.8	8.1	11.5	16.5	18.0
($\beta = 0.001$)	(0.2)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
OS-MAP	**6.0**	6.2	9.7	14.8	19.3	19.7
($\beta = 0.01$)	**(0.1)**	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
OS-MAP	5.9	5.9	8.9	14.7	**24.5**	**25.4**
($\beta = 0.1$)	(0.1)	(0.1)	(0.1)	(0.1)	**(0.1)**	**(0.1)**

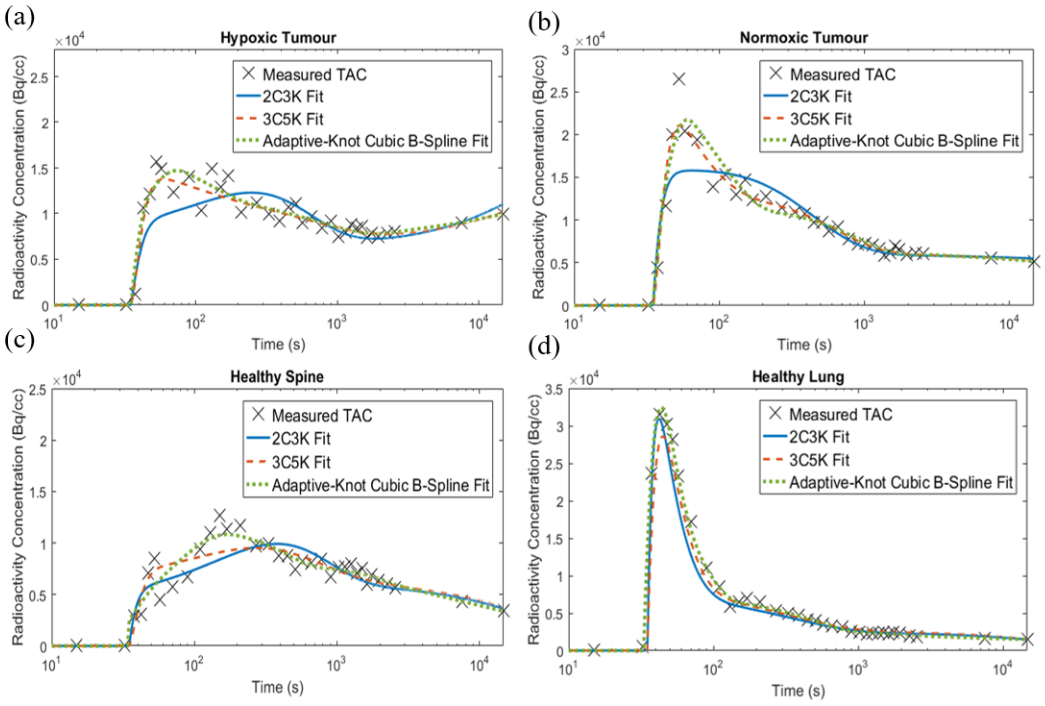


Figure 3.8. 2C3K, 3C5K and adaptive-knot cubic B-spline fits to the TACs obtained from (a) hypoxic tumour, (b) normoxic tumour, (c) healthy lung and (d) healthy spine ROIs in an FMISO dPET image sequence of a stage IV NSCLC patient.

3.4.2 FMISO NSCLC Phantom

3.4.2.1 Selection of TACs for the FMISO NSCLC Phantom

Figure 3.8 shows fits of the 2C3K, 3C5K and cubic B-spline models to the real TAC data obtained from normoxic and hypoxic tumour and healthy lung and spine regions of the imaged NSCLC patient. Runs test and leave-one-out cross validation weighted RSS scores are shown in Table 3.5.

Fits of the adaptive-knot spline and 3C5K models had the lowest cross-validation scores for the healthy tissue and tumour regions respectively, and were therefore used to represent the ground-truth TACs for these regions in the FMISO NSCLC phantom. All the model fits used in this phantom passed the runs test. The worst leave-one-out cross-validation scores were obtained for the 2C3K model fits, which only passed the runs test for the hypoxic tumour region.

There appears to be some unusual structure in the spine TAC shown in Figure

Table 3.5. Weighted RSS errors for model fits to patient TACs calculated using leave-one-out cross-validation. The lowest RSS value for each tissue region is shown in bold. Fits that passed the Wald–Wolfowitz runs test at the 5% significance level are underlined.

Model	Healthy Lung	Healthy Spine	Hypoxic Tumour	Normoxic Tumour
2C3K	6.64×10^5	5.65×10^5	4.57×10^5	3.94×10^5
3C5K	4.53×10^5	3.50×10^5	$\mathbf{1.84 \times 10^5}$	$\mathbf{1.31 \times 10^5}$
Adaptive-Knot Cubic B-Splines	$\mathbf{3.63 \times 10^5}$	$\mathbf{1.97 \times 10^5}$	1.87×10^5	1.48×10^5

3.8 (d). There are a number of potential causes for this structure, such as patient movement or the spine ROI containing a mix of different tissues with different tracer uptake characteristics. The remaining TACs shown in Figure 3.8 (a-c) do not show any clear discontinuities, which one would expect to see in motion corrupted TACs. In 4D-PET reconstruction it is important to have a flexible model to describe all voxel TACs in order to avoid spatial bias propagation from poorly described voxels during the reconstruction, as discussed in section 2.4.4. As PET voxels are of the order of millimetres, it is possible for voxels to contain mixtures of different tissues and thus have unusual TAC shapes. If motion correction was not applied during 4D-reconstruction, or if it was applied but failed in one region of the images, then this would also lead to unusual TAC shapes. Thus if the unusual shape of the spine TAC is due to either motion or tissue inhomogeneity then this is still useful for the phantom as it would simulate potential issues that could be encountered during 4D reconstruction of clinical dPET images.

3.4.2.2 Assessment of Reconstructed Image Quality

Reconstructed images showing a single noise realization of the 10 minute frame of the FMISO NSCLC phantom starting 4 hours post injection are shown in Figure 3.9 for each reconstruction algorithm.

Iteration-by-iteration plots of image bias versus noise, averaged over all time-frames and patient voxels, are shown in Figure 3.10 (a) for FMISO NSCLC images reconstructed using the MLEM algorithm and the MAP algorithms with each β value investigated in this work. Figure 3.10 (b) shows a plot of image bias and

noise after the final iteration of each reconstruction algorithm, to allow better visualization of how varying β affects the bias-noise trade-off in the final images. By 30 iterations one can see that updating the images continues to increase the noise but has little effect on the bias.

Figure 3.10 shows that increasing β results in reduced noise levels. Increasing β from 0 up to 0.1 also reduces the bias, though increasing β from 0.1 to 1 causes the bias to increase again. These results indicate that, of the MAP reconstructions tested on the FMISO NSCLC dataset, the optimal β value is 0.1 when $\delta = 0.1$, as smaller β values are both more biased and noisier, while higher ones are less noisy but over-smooth the images, causing the bias to increase.

3.5 Discussion

In this chapter the PET-SORTEO simulator was used to produce simulated PET detector data representative of data that would have been produced by a Siemens

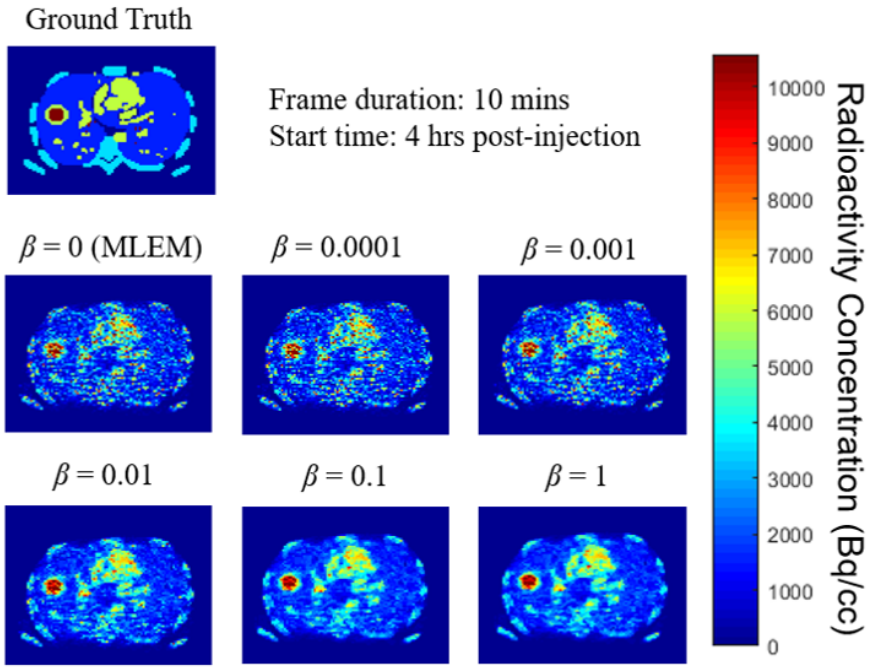


Figure 3.9. MAP-reconstructed images of a single noise realization of the final 10 minute time-frame of the FMISO NSCLC patient. This figure illustrates the effect of varying the β smoothing parameter. Note that the MAP reconstruction with $\beta = 0$ is the standard MLEM reconstruction.

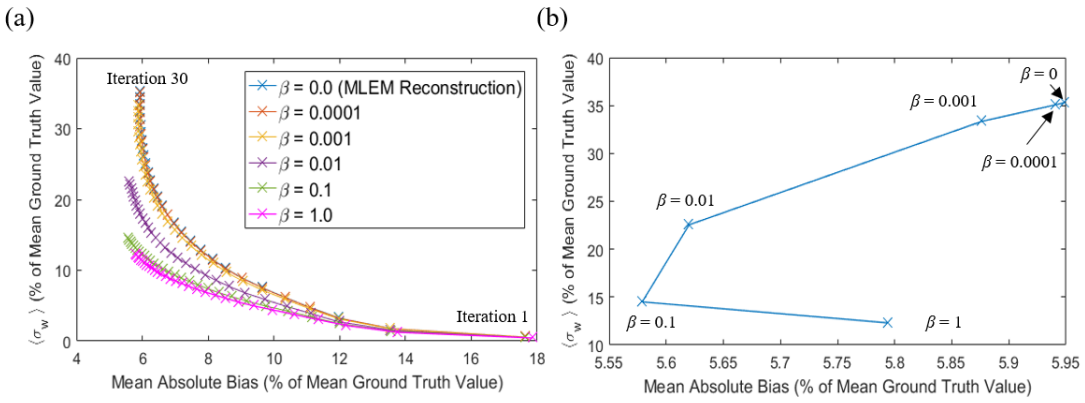


Figure 3.10. Iteration-by-iteration plots of image bias versus noise, averaged over all time-frames and patient voxels, for all MAP reconstructions of the FMISO NSCLC phantom (a), and plots of image bias versus noise in the final reconstruction iteration with varying β (b). MAP reconstructions with $\beta = 0$ correspond to MLEM reconstructions.

Biograph mMR scanner. The methods for performing corrections on the data and for calculating the system matrix, both of which will be used in the 4D-PET reconstruction algorithms proposed in later chapters, were also presented. To test these methods, simulated data from two digital phantoms, one based on a NEMA-NU2 image quality phantom and the other based on an FMISO dPET scan of a NSCLC patient, was reconstructed using both iterative maximum likelihood and MAP reconstruction algorithms.

OS-MAP reconstructions of the NEMA NU2-like phantom with $\beta = 0.01$ generally produced the best CNR values of all the OS-MAP reconstructions investigated, except for the two largest spheres where $\beta = 0.1$ performed better. OS-MAP with the Lange penalty and $\beta = 0.01$ and $\delta = 0.1$ was found to produce images with improved resolution, though slightly higher noise, than those produced with 4 iterations of OSEM, followed by 4 mm spatial Gaussian smoothing.

To check how realistic the simulation and reconstruction methods presented here are, the results from the NEMA NU2-like phantom scan can be compared to those from a study, published by Delso *et al.* [33], in which a NEMA NU2 image quality phantom was scanned with a real mMR scanner. Delso *et al.* [33] filled the phantom volumes with the same activities used here, and scanned the phantom in 3D-mode for 10 minutes, which was also done here. The results from their study

are shown in Table 3.8. Note that the results in this table were obtained from phantom images reconstructed using 4 OSEM iterations with 21 subsets, followed by 4 mm Gaussian smoothing post-reconstruction.

One cannot draw exact comparisons between the simulated and published results because the two studies used different methods to calculate the system matrix and correct the scanner data. The two studies also used different numbers of OSEM subsets and iterations. Nevertheless, the NEMA NU2-like results shown in Tables 3.2, 3.3 and 3.4 are still broadly similar to those published by Delso *et al.* [33], indicating that the simulation and reconstruction methods being used in this work are reasonably representative of those used for real PET scans.

When constructing the FMISO NSCLC dPET phantom, fits of the 2C3K, 3C5K and adaptive-knot cubic B-splines were used to fit each TAC from the clinical FMISO scan, apart from the AIF. The spline and 3C5K models gave the best fits to the healthy tissue and tumour TACs, respectively. This is despite the fact that the 2C3K model is the most commonlly used model for describing FMISO-dPET tumour TACs [180, 103]. However, McGowan *et al.* [103] have shown that for whole-tumour FMISO kinetics the 3C5K model often outperforms the 2C3K model, which is consistent with the results obtained here. Although the 3C5K model performed best for the tumour TACs, the weighted RSS errors of the spline fits are not substantially higher.

These results provide a good illustration of the fact that different models can work better for different tissues in a given dPET image. This is important

Table 3.6. Published results from an mMR scan of a real NEMA NU2 phantom [33]. CRCs, BVs and CNRs are shown, with the asociated errors given in brackets.

	Sphere Diameter (mm)					
	10	13	17	22	28	37
CRC (%)	32.5 (5.1)	50.0 (9.2)	62.9 (7.2)	70.8 (6.0)	65.1 (1.2)	72.3 (1.1)
BV (%)	5.3 (1.0)	4.8 (0.8)	4.2 (0.5)	3.7 (0.3)	3.3 (0.2)	3.0 (0.1)
CNR	6.1 (0.2)	10.4 (0.2)	15.0 (0.2)	19.1 (0.1)	19.7 (0.1)	24.1 (0.1)

for 4D-PET reconstruction, as bias from poorly modelled regions can spatially propagate to well modelled ones during the reconstruction, meaning that care must be taken when selecting the model to be fitted during the reconstruction, as discussed in section 2.4.4.

Of all the reconstructions of the FMISO NSCLC dPET phantom data investigated, the MAP reconstruction with $\beta = 0.1$ and $\delta = 0.1$ performed the best (see Figure 3.10). The 4D-PET reconstruction algorithms proposed in later chapters will be used to reconstruct images from the same simulated FMISO NSCLC dPET phantom data used here. To estimate the potential gains of using these 4D algorithms, they need to be compared to conventional image reconstruction methods as well. In addition to using standard MLEM reconstructions (which provide no regularization), spatially regularized MAP reconstructions with $\beta = 0.1$ and $\delta = 0.1$ will also be used.

The fact that the optimal β value for the FMISO NSCLC dPET phantom was found to be a factor of 10 higher than that of the NEMA NU2-like phantom is likely to be because 2D-mode data was simulated for the former phantom, while 3D-mode data was simulated for the latter. The improved sensitivity of the 3D-mode acquisition leads to a higher number of detected counts, and therefore a better SNR, meaning that lower levels of spatial smoothing are required.

3.6 Conclusion

The methods proposed in this chapter for simulating and reconstructing PET scanner data from digital phantoms were shown to produce images representative of those that would have been obtained real PET scan data. The methodologies described in this chapter for calculating the system matrix, producing simulated dPET data from digital phantoms and evaluating the quality of reconstructed dPET image sequences will be used for implementing and evaluating the 4D-PET reconstruction algorithms proposed in the next two chapters.

Chapter 4

4D-PET Reconstruction with Adaptive-Knot Cubic B-Splines[1]

4.1 Introduction

4D-PET reconstruction has the potential to substantially improve the SNR in both the reconstructed dPET image sequences and the estimated kinetic parameters. However, in cancer imaging studies many tissues can be present in the scanner FOV, and the tracer uptake characteristics of each tissue can be very different, leading to a wide range of potential TAC shapes to be fitted. Bias from poorly modelled regions has been shown to spatially propagate to well modelled regions during 4D-PET reconstruction [81], meaning that flexible temporal functions are desirable. However, the optimal choice of temporal function to use in such situations remains an open question [124, 129].

As was discussed in section 2.4.4, there are two main ways to handle the bias propagation issue. One method is to reconstruct the image sequences using flexible models such as linear kinetic models. If the linear model being used doesn't contain kinetic parameters related to the biological processes being studied, then the temporally smoothed images can be analysed post-reconstruction with a more physiologically relevant kinetic model. The other approach is to use adaptive kinetic

[1]Some of the results presented in this chapter were presented at the 2017 IEEE International Symposium on Biomedical Imaging in Melbourne, Australia. These were published in: G. P. Ralli, D. R. McGowan, M. A. Chappell, R. A. Sharma, G. S. Higgins and J. D. Fenwick. 4D-PET reconstruction of dynamic non-small cell lung cancer [18-F]-FMISO-PET data using adaptive-knot cubic B-splines. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 1189-1192, 2017.

modelling, which first fits a primary kinetic model containing the kinetic parameters of interest and then fits a more flexible secondary model to the residuals of the primary model fit to each voxel TAC [80, 101]. Both of these methods require a flexible temporal model: in the former to produce temporally smoothed images with low bias levels and in the latter to act as the secondary model.

In situations with diverse kinetics, two commonly used linear models are the spectral model of Cunningham and Jones [25] and spline functions [29], both of which were introduced in section 2.4.3. Although the spectral model can describe a wide range of TAC shapes, large numbers of spectral basis functions are often required to do so, potentially leading to over-fitting of the data [54]. Spline functions can describe a wide range of TAC shapes with comparatively fewer B-spline basis functions, potentially making them more effective at keeping bias levels low without over-fitting the data [109, 162]. To our knowledge, however, no studies have compared the performance of spline and spectral model-based 4D-PET reconstruction.

An open question with spline-based reconstructions is how to place the knots. Nichols *et al.* [109] proposed placing the knots such that they produce equal arc lengths along the head curve (a temporal histogram of all list-mode events). This approach doesn't account for the fact that TACs in different voxels can be very different, meaning that the optimal knot positions are likely to vary from TAC to TAC. Verhaeghe *et al.* [162, 163] proposed a method which starts with a large number of knots for each TAC and after fixed numbers of iterations pauses the reconstruction to adjust the number and positions of the knots. An open question with this method is how often to pause the reconstruction to re-adjust the knots. Furthermore, it has been shown in the general literature concerning splines that approaches which start with large numbers of knots and gradually reduce them often fail to converge to the optimal knot positions [60, 192].

In this chapter a 4D-PET reconstruction algorithm using a theoretically driven adaptive-knot placement algorithm is proposed. This method is compared to spectral model-based 4D-PET reconstruction, as well as the conventional (non-4D)

MLEM and MAP algorithms. Each of these methods is used to reconstruct the simulated FMISO NSCLC dPET phantom data described in chapter 3.

4.2 Adaptive-Knot Placement Algorithm

4.2.1 B-Spline Basis Functions

A p^{th} order spline function $S_p(t)$ is defined as a piecewise polynomial function of degree $p-1$. Splines can be expressed as a weighted sum of p^{th} order B-spline basis functions:

$$S_p(t) = \sum_{k=1}^{N_k} \theta_k B_{k,p}(t), \tag{4.1}$$

where N_k is the number of B-spline basis functions and θ_k is the coefficient of the k^{th} B-spline function $B_{k,p}(t)$.

The B-splines are defined on a knot vector $\boldsymbol{\tau} = [\tau_1, \tau_2, ..., \tau_{N_k+p}]$, where the knots are listed in ascending order. From the knot positions, the B-splines are calculated using the Cox-de Boor recurrence relations [29]:

$$B_{k,1}(t) = \begin{cases} 1 & \tau_k \leq t < \tau_{k+1} \\ 0 & \text{Otherwise,} \end{cases} \tag{4.2}$$

$$B_{k,p}(t) = \frac{t-\tau_k}{\tau_{k+p-1}-\tau_k} B_{k,p-1}(t) + \frac{\tau_{k+p}-t}{\tau_{k+p}-\tau_{k+1}} B_{k+1,p-1}(t). \tag{4.3}$$

Spline functions of order p are continuous at the knot sites, as are their derivatives up to the derivative of order $p-1$, provided the knots are distinct. If multiple knots are present at a given point (e.g if $\tau_k = \tau_{k+1}$) then the continuity of derivative order is reduced by 1 for each additional knot [29].

4.2.2 Good Knot Placement

For an order p spline defined on the interval $[a, b]$, the first and last knots are usually repeated with multiplicity p, such that $\tau_1 = ... = \tau_p = a$ and $\tau_{N_k+1} = ... = \tau_{N_k+p} = b$, to allow for discontinuities at the beginning and end of the curve. What remains is to choose the internal knots.

For a spline approximation of a given function $h(t)$, theorem XII.5 of [29] shows that selecting the m internal knot positions such that

$$\int_{\tau_i}^{\tau_{i+1}} |h^{(p)}(t)|^{\frac{1}{p}} dt = \frac{1}{m} \int_a^b |h^{(p)}(t)|^{\frac{1}{p}} dt, \tag{4.4}$$

where $h^{(p)}(t)$ is the p^{th} order derivative of $h(t)$, gives an approximation error of order m^{-p}, for noise free data. Adding more than a certain number of knots when fitting splines to noisy data results in over-fitting, and therefore a worse description of the data.

4.2.3 The Weighted Savitzky-Golay Filter

Calculating the knots for dPET data with (4.4) requires function derivatives to be estimated from discrete, noisy data points. In this work the Savitzky-Golay filter [134] (with weighted data points) was used to suppress noise amplification when calculating the derivatives.

In this methodology a subset of k adjacent data points $\{t_i, x_i\}$, with t_i and x_i being the independent variables and noisy data points, respectively, are fitted with a low degree polynomial of degree q, using weighted least squares. Note that k must be an odd number. In this work k was set to 11, as preliminary experiments showed that this gave good estimates of the derivatives while effectively suppressing noise. The dPET weighting factors defined in equation (3.1) were used:

$$w_m = \frac{\Delta T_m e^{-\lambda t_m}}{a_m}, \tag{4.5}$$

where a_m is the decay-corrected average activity concentration within the ROI during frame m, ΔT_m is the frame duration, t_m is the mid-point of the frame and λ is the radiotracer decay constant.

A change of variable, $z = t_i - t_c$, 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 function $X_q(z)$,

$$X_q(z) = \alpha_1 + \alpha_2 z + ... + \alpha_q z^{q-1}. \tag{4.6}$$

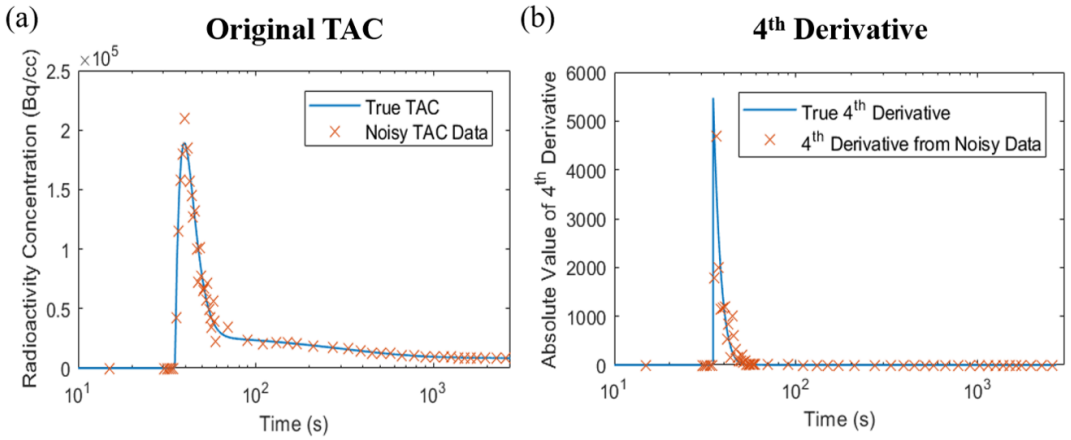


Figure 4.1. Illustration of the Savitzky-Golay method for calculating the magnitude of the 4^{th} derivative with $k = 11$ and $q = 5$. (a) Shows a simulated noisy TAC and (b) shows the magnitude of the 4^{th} derivative calculated from the noisy TAC data using the Savitzky-Golay method along with the 4^{th} derivative of the true TAC. Note that the simulated noisy TAC was generated using the methodology described in section 4.3.1.1.

For cubic (order 4) splines, the 4^{th} order derivative is required when using (4.4) to place the knots, meaning that $X_q(z)$ must be quartic or higher ($q \geq 5$). The value of the 4^{th} derivative at the central data point (i.e. $t = t_c, z = 0$) is given by:

$$\left.\frac{d^4 X_q}{dt^4}\right|_{t=t_c} = 24\alpha_5, \tag{4.7}$$

i.e. 24 times the fitted coefficient α_5.

This methodology does not allow the derivatives of the first and last $\frac{k-1}{2}$ data 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 [134].

In order to illustrate the performance of this method for calculating the magnitude of the 4^{th} derivative required by equation (4.4), Figure 4.1 shows an example of a simulated simulated noisy TAC and the magnitude of it's 4^{th} derivative calculated using the Savitzky-Golay method with $k = 11$ and $q = 5$, which were the setting used in the rest of this chapter. Note that the simulated TAC used in this figure was generated using the methodology described in section 4.3.1.1. Calculating high order derivatives of noisy data can lead to high levels of noise in the estimated derivatives, and using high levels of smoothing (in the

case of the Savitzky-Golay method this corresponds to a high value of k) can lead to a loss of temporal information. The example shown in Figure 4.1 shows the settings of the Savitzky-Golay filter used here is able to capture the shape of the 4^{th} derivative while keeping the noise at moderate levels.

4.2.4 Knot Placement Algorithm

The algorithm used in this work to calculate the knot positions for cubic spline fits to individual TACs is summarized in the four steps below. Note that this algorithm requires the number of knots to be specified in advance.

1. Place 4 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). Here the optimal rise time is determined by performing multiple cubic spline fits with the rise times varied in one second intervals from 0 to 60 s into the scan. For each rise time estimate, 4 knots are placed at this rise time and the remaining knots are distributed uniformly between this point and the end-point of the TAC. The rise time estimate that produces the lowest weighted RSS error in the spline fit is taken to be optimal.
3. Calculate the 4^{th} derivative of all points along the TAC after the initial rise point using the weighted Savitzky-Golay method ($k = 11, q = 5$).
4. Place the remaining knots according to (4.4).

4.3 Methods

4.3.1 Validation of the Adaptive Knot Placement Algorithm with Single TACs

4.3.1.1 Generation of Simulated TACs

Fitted tumour and AIF TACs were taken from a set of clinical FMISO dPET images of stage IV NSCLC patients. Each of these images was acquired using the same dPET imaging protocol described in section 3.2.2 for the clinical image

used to create the FMISO NSCLC dPET phantom. The AIF TACs were taken from cylindrical ROIs of diameter 10 mm located in the centre of the descending aorta on five consecutive PET axial slices. Tumour TACs were obtained from ROIs outlining primary tumours, involved nodes and metastases. All ROIs were drawn by an experienced radiologist and checked by a second radiologist. A total of 13 tumour TACs and 5 AIFs were used in this study.

The AIFs were fitted with the three-exponential model of Feng *et al.* [40] and the tumour TACs were fitted with the 3C5K compartment model (as with the FMISO tumour TACs in section 3.4.2.1, the 3C5K model was found to fit all tumour TACs better than the more commonly used 2C3K model). Each TAC was fitted using weighted least squares, with the weighting factors defined in (4.5), using the Levenberg-Marquardt algorithm provided by the MATLAB optimization toolbox.

Each of the fitted TACs was re-binned into the [1×30 s, 6×5 s, 6×20 s, 7×60 s, 10×120 s, 3×300 s] time-frame sequence used for the FMISO NSCLC dPET phantom. For this particular study only the first 45 minutes of the scan (and not the two 10 minute frames 2 and 4 hours post injection) was used. 1000 Gaussian noise realizations were added to each of the TACs using the dPET noise model of Chen *et al.* [20]:

$$\sigma_m^2 = c\frac{1}{w_m} = c\frac{a_m}{\Delta T_m e^{-\lambda t_m}}, \tag{4.8}$$

where σ_m^2 is the variance of time-frame m, c is a scale factor determined from the residuals of the model fits and w_m is the weighting factor of frame m defined in (4.5).

This study was repeated for two different noise levels. The first, 'low noise' level, had scale factor $c = 3$, and simulates the level of noise that may be seen during a reconstruction. The second, 'high noise' level, had scale factor $c = 9$, and simulates the level of noise that would be seen after a conventional (non-4D) reconstruction has been completed. The scale factor from the high noise level was determined from the residuals of fits to the AIF and tumour TACs taken from single voxels in the patient images, while the low noise level was set to be a factor of three lower than this.

4.3.1.2 Assessment of Spline Fit Quality

For each noise realization of both the high and low noise TACs, the knots were calculated using the algorithm described in section 4.2.4. Because this algorithm requires the number of knots to be specified in advance, this was repeated with the number of free knots varied from 1 to 20 for each TAC. Note that here 'free' knots refers to knots whose position is free to vary, which are used in addition to the 12 'fixed' knots (4 at the beginning, end and initial rise of each TAC).

A further study in which, for a given TAC, the knots were placed using the high noise data and fitted to the low noise data was also carried out. The purpose of this study was to investigate the feasibility of pre-calculating the knots using conventionally reconstructed dPET image sequences and then performing spline-based 4D-PET reconstruction using these knots. This would simplify the implementation of the 4D algorithm by not requiring the knots to be calculated during the reconstruction.

The three cases studied here are summarized in Figure 4.2. Linear non-negative weighted least squares, using the weighting factors defined in (4.5), was used to fit all of the splines in this study.

The fitted splines were binned into the same time-frame sequence as the true TACs. The accuracy of each individual spline fit was characterised using the

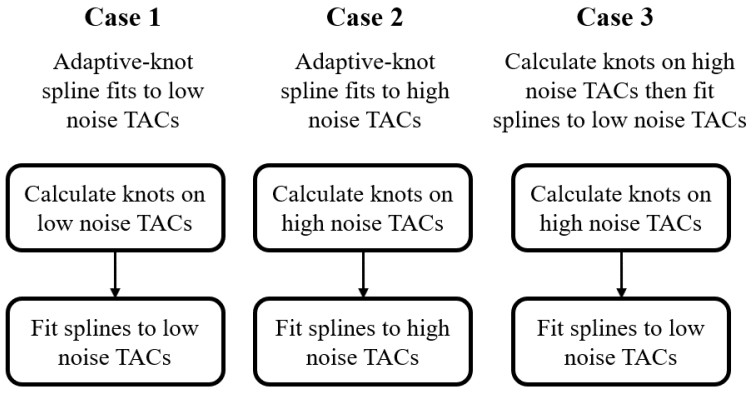


Figure 4.2. Summary of the three cases investigated in the single TAC study to validate the proposed knot placement algorithm.

root mean square (RMS) error

$$RMS = \sqrt{\frac{1}{N_T}\sum_{m=1}^{N_T}\left(x_m^{true} - x_m^{spline}\right)^2}, \tag{4.9}$$

where N_T is the number of time-frames and x_m^{true} and x_m^{spline} are the true and fitted TAC values in time-frame m, respectively.

The optimal number of knots for a given noise realization of an individual TAC was taken to be the number that produced a spline fit minimizing the Akaike information criterion, corrected for finite sample sizes (AICc) [18]:

$$AICc = \sum_{m=1}^{N_T}\frac{(x_m^{TAC} - x_m^{spline})^2}{\sigma_m^2} + 2N_{knots} + \frac{2N_{knots}(1+N_{knots})}{N_T - N_{knots} - 1}, \tag{4.10}$$

where N_{knots} is the total number of knots and x_m^{TAC} is the noisy value of the TAC data point at time-frame m. The second term in (4.10) penalizes over fitting, while the third corrects for finite sample sizes [18].

Once the optimal number of free knots, N_{opt}, for a given noise realization of a TAC had been determined, a further spline fit with the same 12 fixed knots and N_{opt} uniformly spaced knots between the initial rise and end points of the TAC was performed. The RMS error of the spline fit with uniform knots was then calculated, to assess the extent to which using adaptive knot placement gives improvements over using uniform knots.

For every TAC, the RMS error for spline fits with N_{opt} free knots was averaged over all 1000 noise realizations. The average RMS error for a given TAC was expressed as a fraction of the mean ground-truth TAC. These average RMS values were then averaged over all the tumour and AIF TACs to give overall fitting errors, $\langle RMS \rangle$, for both types of TAC. This was repeated for each of the three cases listed in Figure 4.2. The $\langle RMS \rangle$ values for the spline fits with N_{opt} uniform knots were also calculated for the tumour and AIF TACs in each of the three cases.

4.3.2 4D-PET Reconstruction Study

4.3.2.1 Image Reconstruction Algorithms

All 50 noise realizations of the simulated FMISO NSCLC dPET phantom dataset described in chapter 3 were reconstructed using the conventional MLEM and MAP

Table 4.1. Settings used for the each of the reconstruction algorithms used in this chapter.

Algorithm	Number of Iterations	Number of Sub-Iterations	β	δ
MLEM	30	-	0.0	-
MAP	30	-	0.1	0.1
Nested-EM	30	30	0.0	-
Nested-MAP	30	30	0.1	0.1

algorithms. The MAP reconstructions were carried out with the Lange penalty function using the algorithm described in section 3.3.3.1.

The nested-EM algorithm, described in section 2.4.2, was used to perform 4D reconstructions. Recall that each iteration of this algorithm involves two steps: the images are first updated using the MLEM algorithm and then the linear kinetic model is fitted to each voxel TAC using multiple sub-iterations of equation (2.55).

To investigate the effectiveness of incorporating spatial regularization into the 4D reconstruction process, an additional set of nested reconstructions where the image update steps were carried out using one iteration of the MAP algorithm described in 3.3.3.1 were performed. These reconstructions are referred to here as nested-MAP reconstructions.

Table 4.1 shows the settings used for each of the image reconstruction algorithms. The values of β and δ (defined in (2.33)) were both set to 0.1 for both the MAP and nested-MAP reconstructions, as the results in chapter 3 show that these provide good results for the FMISO NSCLC dPET phantom data. The number of sub-iterations was set to 30 because preliminary experiments in which splines were fitted to voxel TACs from MAP-reconstructed phantom images with (2.55) showed that this number was sufficient to get good fits. For each algorithm, a sequence of uniform images with the radioactivity concentration set to 100 Bq/cc was used for the initial dPET sequence estimates. All image sequences were reconstructed onto a grid of $3.1\times3.1\times2.0$ mm^3 voxels.

4.3.2.2 Linear Kinetic Models

1) Adaptive-knot cubic B-splines.

The results presented in section 4.4.1 for the single TAC study show that 11 free knots provides good descriptions of both the tumour and AIF TACs. The results in section 4.4.1 also show that splines fitted to the low noise TACs using knots calculated on the high noise TACs give good descriptions of the true underlying TAC, meaning that pre-calculating the knot positions using conventionally reconstructed images is potentially feasible. Therefore, for each noise realization of the phantom data, voxel specific knots were calculated by using the knot placement algorithm described in section 4.2.4 to place 11 free knots along the voxel TACs derived from the images reconstructed with the MAP algorithm (using the settings described in Table 4.1).

2) Spectral model.

Recall that the spectral model describes TACs as weighted sums of pre-defined exponential decays convolved with the AIF, along with an extra term to account for the finite blood volume in the region that the TAC is extracted from:

$$f(t, \boldsymbol{\theta}) = \theta_0 C_p(t) + C_p(t) \otimes \sum_{k=1}^{N_k} \theta_k \exp(-\alpha_k t), \tag{4.11}$$

where $f(t, \boldsymbol{\theta})$ is the TAC, $C_p(t)$ is the AIF, θ_k is the coefficient of the k^{th} exponential decay function, with decay constant α_k, and $\otimes$ denotes a convolution. Because the α_k values are defined in advance, only the coefficients θ_k need to be calculated when fitting the model.

In this work 100 spectral basis functions were used. The α_k values were logarithmically spaced between 1.1×10^{-4} s^{-1} (the decay constant of ^{18}F) and 0.01 s^{-1}. The AIF for each noise realization was estimated by drawing a large ROI in the left ventricular region of the MAP reconstructed images and fitting the TAC from this ROI with the three exponential model of Feng *et al.* [40]. Note that decay correction was not applied to the TACs prior to the spectral model fit. The images produced by the spectral model-based reconstructions were decay-corrected post-reconstruction.

A further, idealized, set of spectral model-based reconstructions were formulated by firstly fitting the spectral basis functions to TACs from the ground truth phantom images and then removing any basis functions with $\theta_k < 0.1$ from a given voxel. Removing basis functions with $\theta_k < 0.1$ was not found to noticeably affect the fits to the ground-truth TACs. Once the fitted basis functions with $\theta_k < 0.1$ were removed, 5 basis functions remained for both the healthy lung and normoxic and hypoxic tumour TACs and 7 basis functions remained for the spline TAC. This method could not be used practically (since the true images are not known in reality), but it allows the proposed adaptive-knot spline-based reconstruction algorithm to be compared to 4D reconstructions with both un-optimized and idealized spectral model models.

4.3.2.3 Image Analysis

The bias and noise measures defined in section 3.3.4.2 were used to calculate the bias and noise in each voxel of the images produced at every iteration of the reconstruction algorithms listed in Table 4.1. As before, overall bias and noise were characterised as the mean absolute bias and mean image noise ($\langle\sigma_w\rangle$) averaged over all voxels within the patient (the 'whole patient' region). The same measures were also averaged over the tumour region alone, as this is usually the primary focus of oncological dPET studies.

As with most dPET protocols, short time-frames are used for the early time-points of the dPET phantom image sequence to capture the faster variation of the TACs early on. Fewer counts are therefore available in the early time-frames, making them particularly noisy. To compare the performance of the 4D algorithms at the early time-frames, values of the overall bias and noise in the whole patient and tumour regions were also calculated for just the first 120 s of the scan.

4.4 Results

4.4.1 Validation of the Adaptive-Knot Placement Algorithm with Single TACs

The results for each of the cases listed in Figure 4.2 are given in Table 4.2. For each case, the range of N_{opt} values and the $\langle RMS \rangle$ values for spline fits using N_{opt} free knots, 11 free knots and N_{opt} uniform knots are shown for both the tumour and AIF TACs. Note that in case 3 spline fits with uniform knots were not performed.

The reason for showing $\langle RMS \rangle$ values with 11 free knots is that the results in Table 4.2 show that the N_{opt} values vary from TAC to TAC. Despite this, using 11 free knots for each TAC doesn't appear to substantially increase the fitting errors, except when fitting to the high noise TACs (case 2), which implies that using 11 free knots for all TACs in the adaptive-knot spline-based 4D-PET reconstruction could still give reasonable fits to all the TACs, while simplifying the overall implementation.

Table 4.2 also shows that the case 3 results (where the splines are fitted to the low noise TACs using knots calculated on the high noise TACs) are very similar to the case 1 results (where the knot calculations and spline fits are carried out using

Table 4.2. Results for each of the cases shown in Figure 4.2 for the validation of the proposed knot placement algorithm. For each case the range of N_{opt} values and the $\langle RMS \rangle$ values for spline fits using N_{opt} free knots, 11 free knots and N_{opt} uniform knots are shown for both the tumour and AIF TACs. Spline fits with uniform knots were not performed in case 3.

TAC type	Range of N_{opt} values	$\langle RMS \rangle$ for spline fits with N_{opt} free knots	$\langle RMS \rangle$ for spline fits with 11 free knots	$\langle RMS \rangle$ for spline fits with N_{opt} uniform knots
Case 1				
Tumour	4-7	0.04	0.04	0.11
AIF	8-9	0.11	0.12	0.62
Case 2				
Tumour	4-6	0.05	0.07	0.19
AIF	7-9	0.17	0.18	0.73
Case 3				
Tumour	4-6	0.04	0.04	-
AIF	8-9	0.11	0.12	-

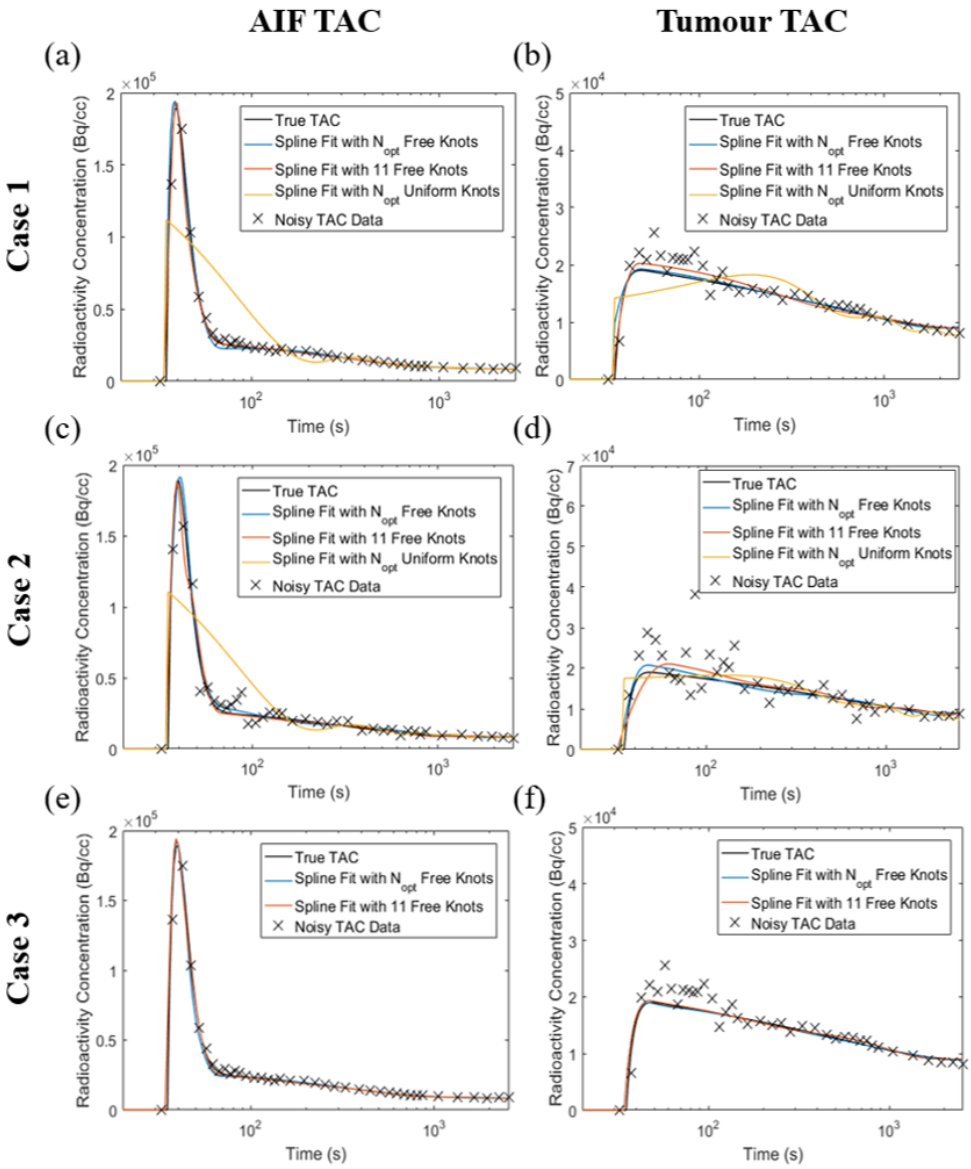


Figure 4.3. Single noise realization of spline fits with N_{opt} free knots, 11 free knots and N_{opt} uniform knots to an AIF TAC (a,c,e) and a tumour TAC (b,d,f), for each of the three cases listed in Figure 4.2. Note that the y-scale of the high noise (case 2) tumour TAC in plot (d) is higher than for the other tumour TACs. Spline fits with uniform knots were not performed in case 3.

the low noise TACs). This implies that it may be feasible to carry out adaptive-knot spline-based 4D reconstructions using knots that have been pre-calculated on conventionally-reconstructed dPET image sequences.

To further illustrate the results in Table 4.2, Figure 4.3 shows examples of a single noise realization of both a tumour and AIF TAC from each of the three cases.

Spline fits using N_{opt} free knots, 11 free knots and N_{opt} uniform knots are shown. The fits with 11 and N_{opt} free knots look very similar to each other, except in case 2. The faster TAC variation at early time-points means that dense knot sequences are required in these regions. The uniform knot sequences are not sufficiently dense at early time-points, leading to high fitting errors in the early TACs.

4.4.2 4D-PET Reconstruction Study

Reconstructed images of a single noise realization of a 5 minute frame of the FMISO NSCLC dPET phantom, starting 40 minutes into the scan, are shown in Figure 4.4 for each image reconstruction algorithm. Mean and variance images, calculated over all 50 noise realizations, are shown for the same time-frame and reconstruction algorithms in Figures 4.5 and 4.6, respectively.

Figure 4.4 appears to show a loss of anatomical detail in the reconstructions based on the un-optimised spectral model, particularly in the spine region. Figure 4.5 shows that the images appear quite blurred in the areas around the bone regions of the un-optimised spectral model-based reconstructions. Furthermore, Figure 4.6 shows that the noise in both the tumour region and bone regions of the un-optimised spectral model-based reconstructions is lower than in the same regions of the other images. The results in Figures 4.5 and 4.6 suggest that the un-optimised spectral model-based reconstructions have not converged in all regions of the images. The un-optimised spectral model consists of 100 basis functions, which is substantially higher than the number of basis functions used by the splines and optimised spectral model. At each 4D reconstruction iteration the models were fitted using 30 iterations of equation (2.55). This number of iterations was chosen as it gave good spline fits to the TAC data, however it is possible that since the unoptimized spectral model contains more parameters more iterations of equation (2.55) would have given better fits. However, the overall images produced by the un-optimised spectral model are already noisier than the other 4D-reconstructions, and using further iterations would likely amplify this noise further.

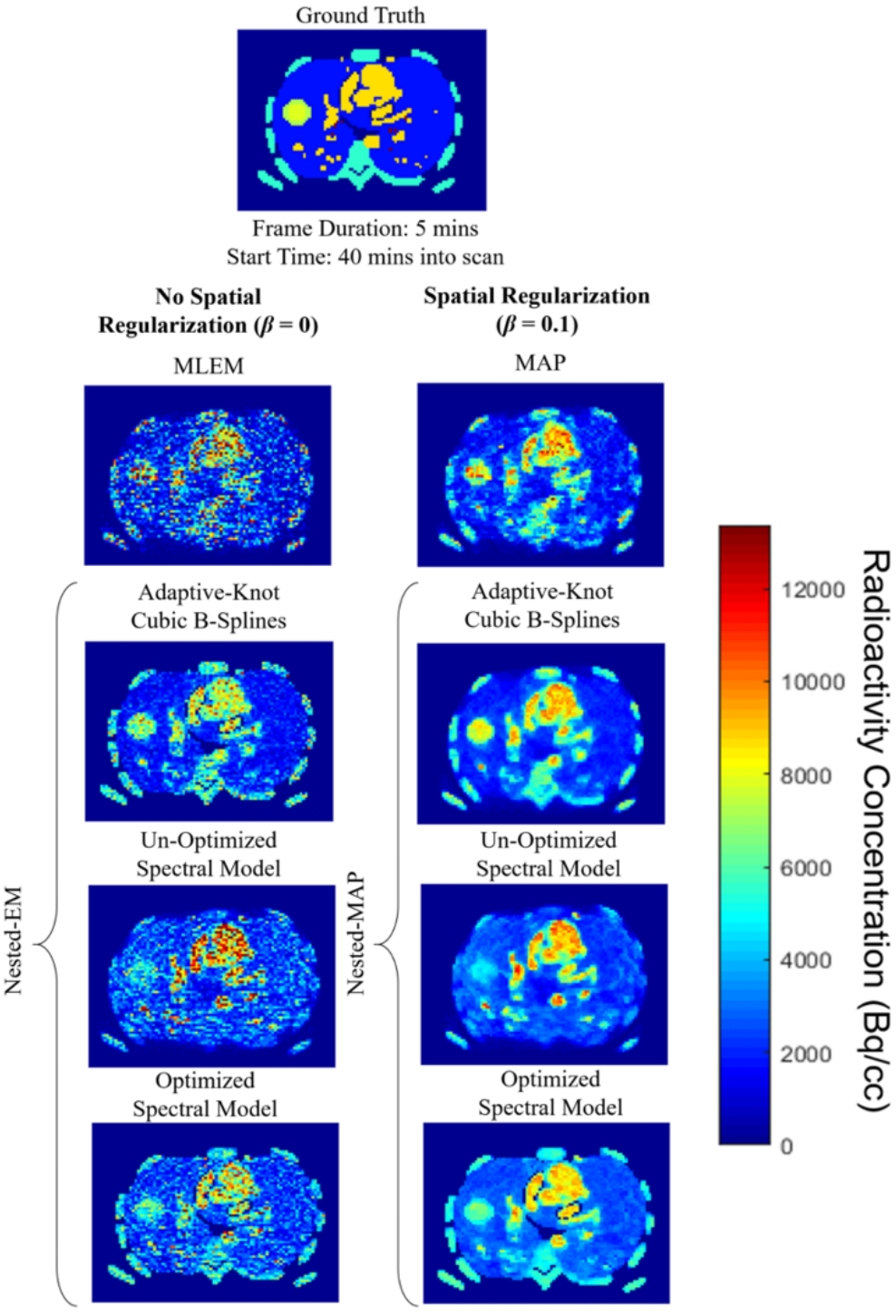


Figure 4.4. Reconstructed images of a single noise realization of a 5 minute time-frame of the FMISO NSCLC dPET phantom, starting 40 minutes into the scan. Images produced by every reconstruction algorithm tested in this chapter are shown.

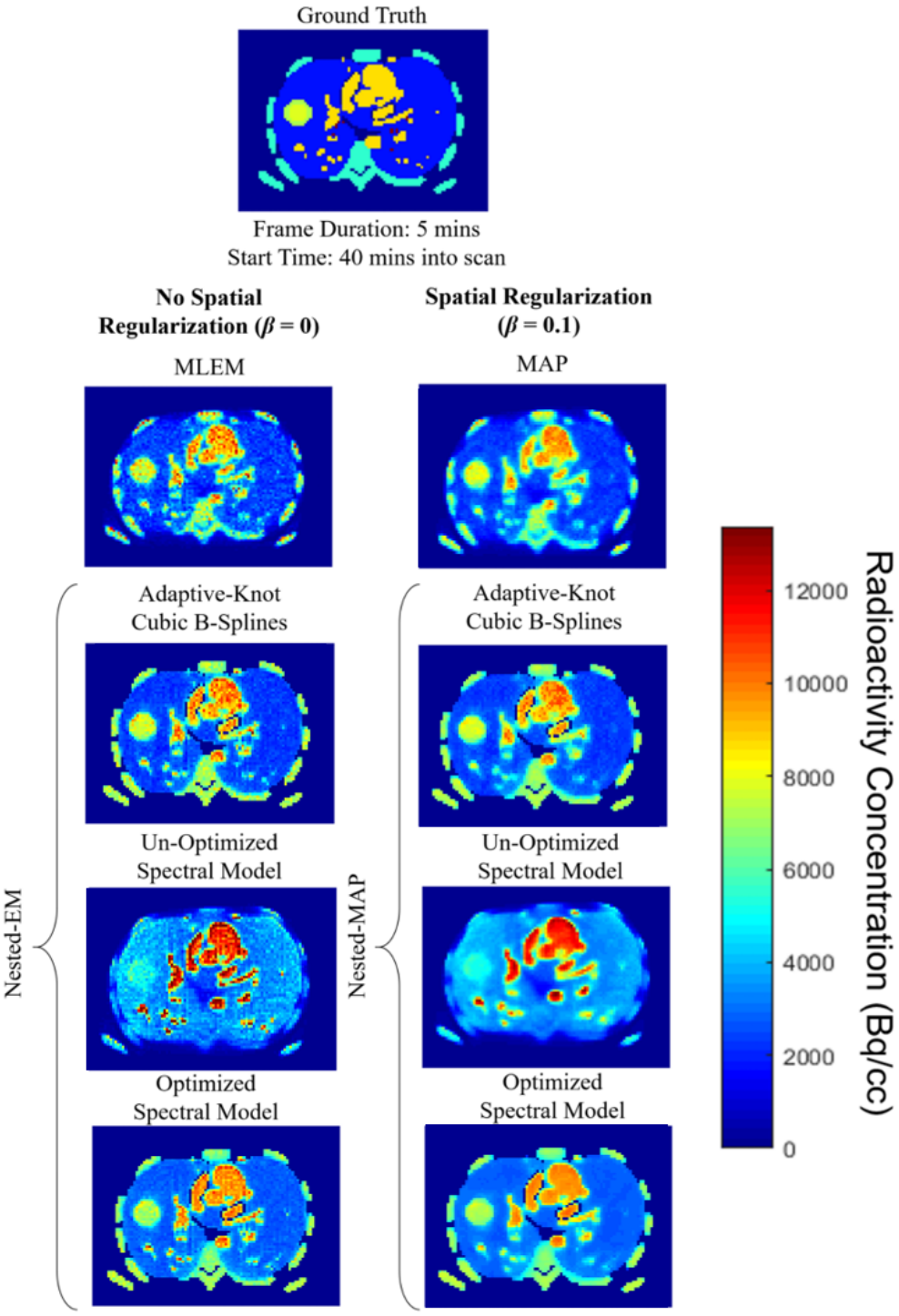


Figure 4.5. Mean of reconstructed images of a 5 minute time-frame of the FMISO NSCLC dPET phantom, averaged over all 50 noise realizations. Images produced by every reconstruction algorithm tested in this chapter are shown.

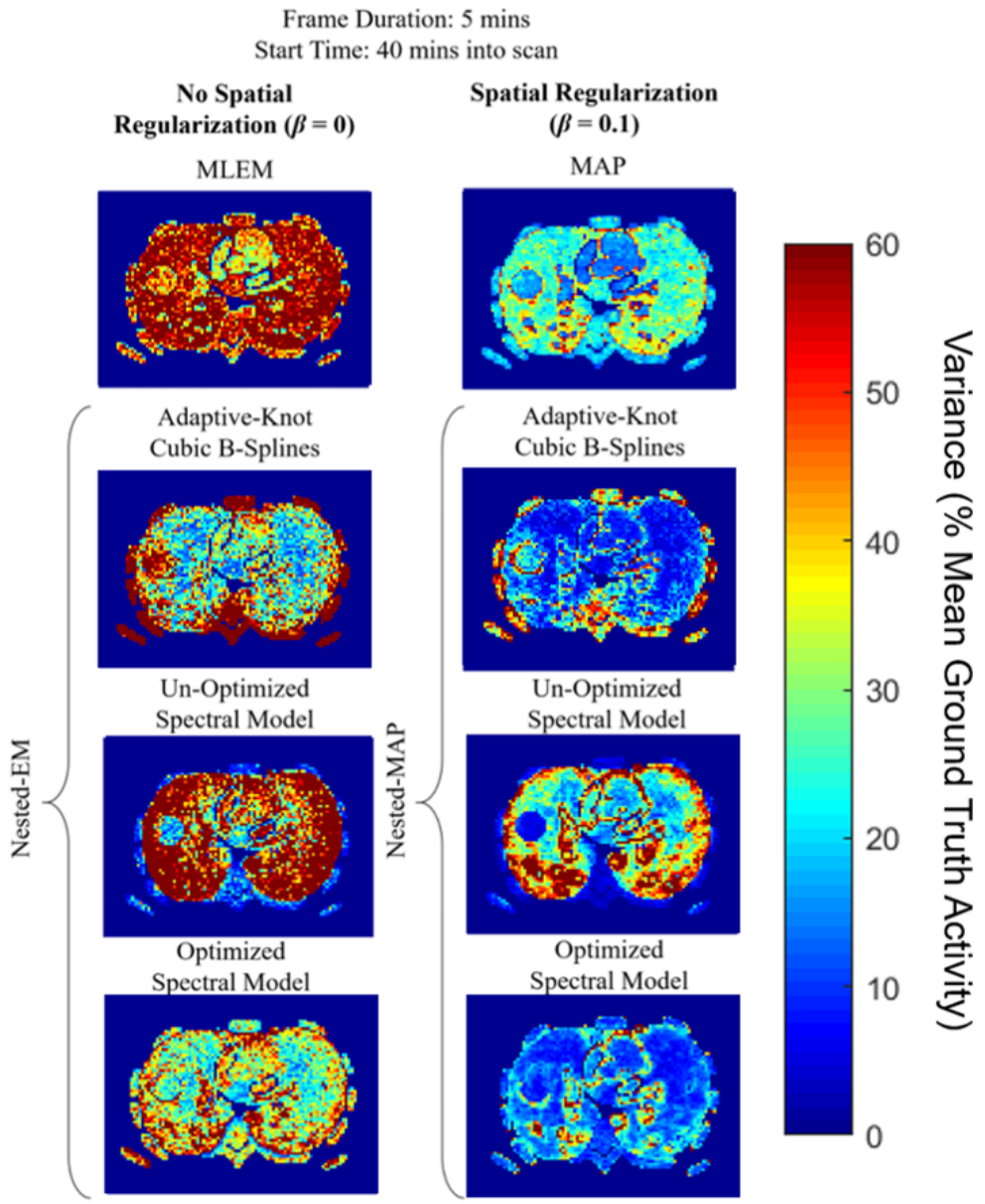


Figure 4.6. Images showing the voxel variance of reconstructed images of a 5 minute time-frame of the FMISO NSCLC dPET phantom, calculated over all 50 noise realizations. The variance values are expressed as a percentage of the mean radioactivity concentration of the ground-truth phantom image of this time-frame.

Iteration-by-iteration plots of image bias versus noise for each image reconstruction algorithm, averaged over all time-frames and patient voxels, are given in Figure 4.7. Reconstructions with and without spatial regularization are shown in separate plots. Plots of image bias versus noise averaged over just the tumour region are also shown.

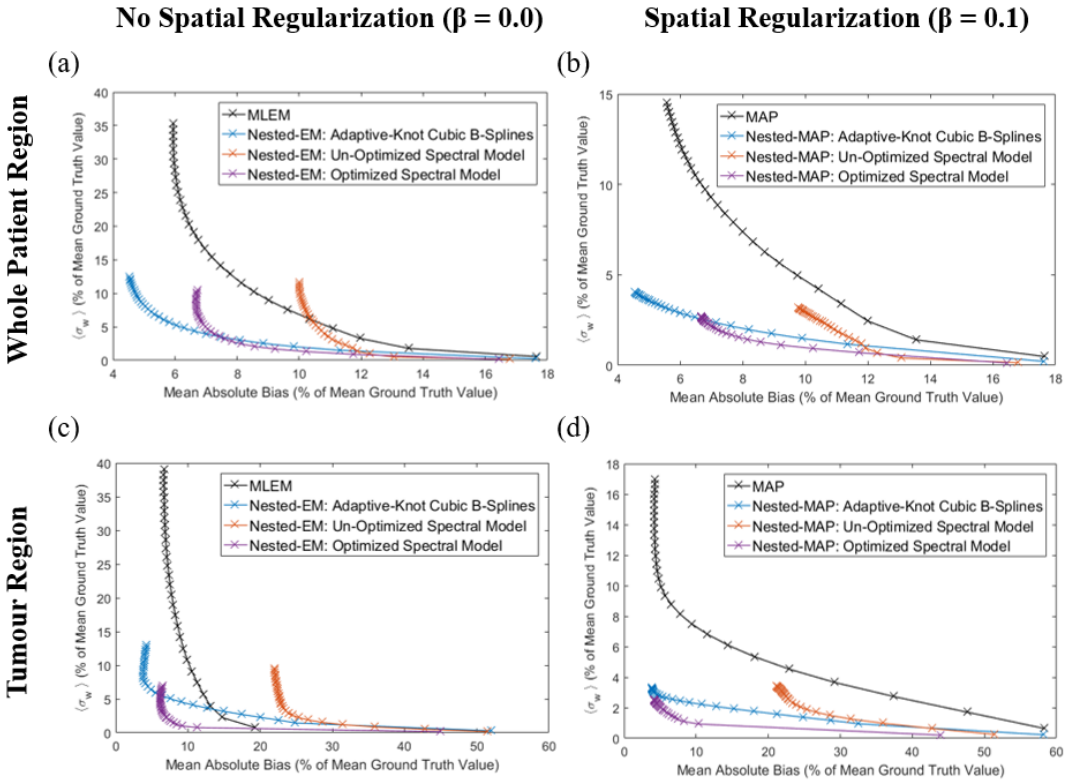


Figure 4.7. Iteration-by-iteration plots of image bias versus noise (averaged over all time-frames) for each image reconstruction algorithm, averaged over all patient voxels (a,b) and just the tumour region (c,d). Results are shown for reconstructions with (b,d) and without (a,c) spatial regularization. Note that the scales on the plots are different.

To give a better illustration of how using spatial regularization affects the quality of the 4D-reconstructed images, the bias-noise curves for the whole patient and tumour regions shown in Figure 4.7 are re-plotted in Figure 4.8 so that the results from 4D reconstructions with and without spatial regularization appear on the same graph. This figure shows that the images produced with spatially regularized 4D reconstructions had lower noise and similar levels of bias to the 4D reconstructed images with no spatial regularization. An exception to this is the tumour region in images produced by the spectral model-based reconstructions, which had lower bias when spatial regularization was used.

In the whole patient region the adaptive-knot spline-based nested-MAP reconstruction gives the best bias-noise trade-off, though for the tumour region the spectral model-based reconstruction performed slightly better. To further illustrate

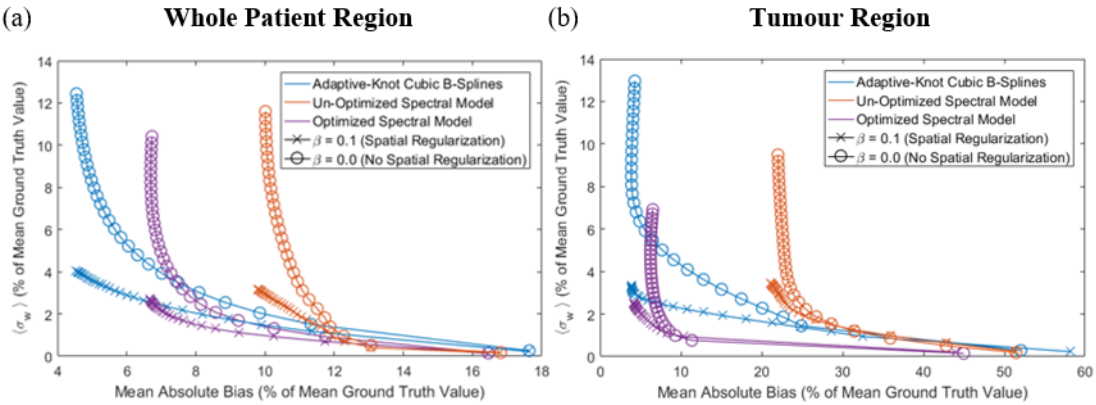


Figure 4.8. Iteration-by-iteration plots of image bias versus noise (averaged over all time-frames) for each image 4D reconstruction algorithm, averaged over all patient voxels (a) and just the tumour region (b). Results are shown for reconstructions with and without spatial regularization. Note that the scales on the plots are different.

this, plots of a single noise realization of a tumour TAC from images produced by the spline and optimized spectral model-based nested-MAP reconstructions, as well as the conventional MAP algorithm, are shown in Figure 4.9. This plot shows tumour TACs taken from an ROI drawn around the entire hypoxic tumour region and a single voxel at the centre of the hypoxic tumour region.

While the whole ROI TACs from the spline and optimized spectral model-based reconstructions in Figure 4.9 look very similar, the single voxel TAC from the optimized spectral appears to give a better description of the TAC, particularly at early time-frames. To further investigate this, Figure 4.10 shows iteration-by-iteration plots of image bias versus noise, averaged over the first 120 s of the scan, for both the whole patient and tumour regions in the nested-MAP reconstructions. This figure shows that for the first 120 s the optimized spectral model gives the best bias-noise trade-off for both the tumour and the whole patient regions.

4.5 Discussion

A 4D-PET reconstruction algorithm using cubic B-spline basis functions with a theoretically driven knot placement algorithm was proposed. This methodology was compared to spectral model-based 4D-PET reconstruction as well as the conventional

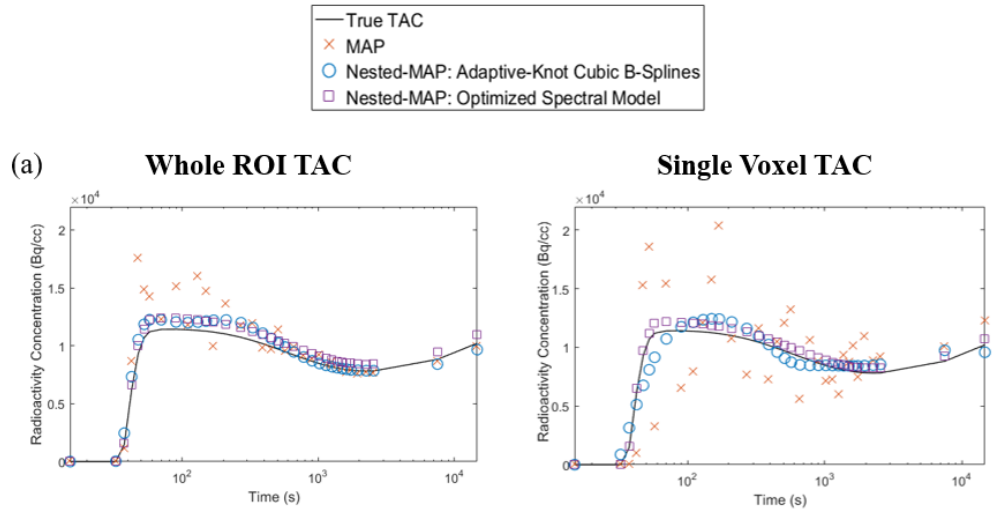


Figure 4.9. Single noise realization of a measured hypoxic tumour TAC, taken from images reconstructed with the MAP and spline and optimized spectral model-based nested-MAP reconstructions. Tumour TACs taken from the whole hypoxic tumour ROI (a) and a single voxel at the centre of the hypoxic tumour region (b) are shown.

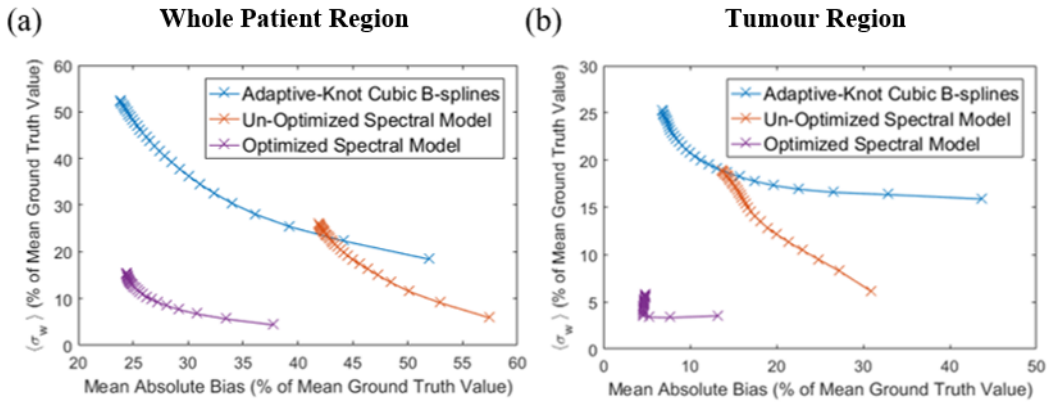


Figure 4.10. Iteration-by-iteration plots of image bias versus noise (averaged over first 120 s of the scan) for each image nested-MAP reconstruction, averaged over all patient voxels (a) and just the tumour region (b).

MLEM and MAP algorithms, using the simulated FMISO NSCLC dPET phantom data described in chapter 3. The proposed knot placement algorithm was found to give good descriptions of single simulated FMISO NSCLC and AIF TACs, provided an appropriate number of knots was used. Overall, the adaptive-knot spline-based reconstruction produced the least biased images of the reconstruction algorithms being compared, and the spline and optimized spectral model-based reconstructions

outperformed the conventional MAP and MLEM algorithms.

When averaging the bias and noise over just the first 120 s of the scan, Figure 4.10 shows that the optimized spectral model provides the best bias-noise trade-off in both the whole patient and tumour regions. The superior performance of the spectral model at early time-frames is likely to be due to the more limited range of TAC shapes that can be generated at early time-points with this model (convolutions of the input function with a weighted sum of exponential decays), whereas B-spline basis functions can in principle create arbitrarily-shaped TAC shapes early on. The spectral model is therefore less likely to over-fit the data in the early frames, leading to both increased noise suppression and to a better description of the TACs in the early time-frames.

Although the optimized spectral model performed best in the early time-frames, the spline-based reconstruction produced the lowest bias out of all the reconstructions, when averaging the bias and noise over the whole patient region and all time-frames. This implies that while the spectral model is better at describing the early time-points, the more flexible splines give better descriptions of TACs across the whole imaged time-course.

The optimized spectral model gave the best bias-noise trade-off within the tumour region, especially when spatial regularization was used in the reconstruction (see Figure 4.7). The ground-truth tumour TACs of this phantom were generated with the 3C5K compartment model, meaning that the spectral model (which is based on the general solution to the compartment model equations) is likely to have an advantage in the tumour region compared to the rest of the phantom, as the non-tumour TACs in the phantom were not generated with compartment models.

An ideal choice of temporal function for 4D-PET reconstruction would combine the superior performance of the spectral model in early time-frames with the lower overall bias of the adaptive knot splines. A new linear kinetic model which is designed to do this is proposed in the next chapter.

One limitation of this study is that the method used to create the optimized spectral model is highly artificial. In the next chapter, the effectiveness of

incorporating temporal roughness penalties into the 4D reconstruction objective function, thus replacing the simplistic method used here, is investigated. Another limitation is that this study is focused entirely on the quality of the reconstructed images, though in dPET studies one is usually interested in the quality of the fitted kinetic parameters as well. The next chapter investigates the effects of 4D-reconstruction on the quality of both the reconstructed images and the kinetic parameters.

The results in Figure 4.8 show that incorporating spatial regularization into the 4D-reconstructions results in improvements in the bias-noise trade-off, provided that the spatial regularization penalty function and hyperparameters are well chosen. The remainder of the 4D-reconstruction work presented in this thesis will therefore use spatially-regularized reconstructions.

4.6 Conclusion

The proposed adaptive-knot cubic spline-based 4D-PET reconstruction algorithm produced images with the lowest bias of all the algorithms being compared in this study, when averaging the image bias over the entire scan time. However, the optimized spectral model gave the best bias-noise trade-off in the early time-frames.

The results in this chapter also indicate that using spatial regularization in 4D-PET reconstruction can lead to further improvements in image quality, provided the parameters are well chosen.

In the next chapter a new linear kinetic model is proposed. This new model is designed to combine the superior performance of the spectral model in the early time-frames with the ability of the splines to provide good descriptions of TACs over the entire imaged time-course.

Chapter 5

4D-PET Reconstruction using a Spline-Residue Model[1]

5.1 Introduction

In the previous chapter a 4D-PET reconstruction algorithm using adaptive-knot cubic splines was proposed. The performance of this algorithm was compared to that of spectral model-based 4D-PET reconstruction. While the spline-based reconstruction produced images with lower bias, when averaging the bias over the entire scan time, the spectral model-based reconstruction produced a better bias-noise trade-off in the early time-frames.

TACs vary most rapidly early on during imaging, and so short time-frames are used for the early time-points, making them particularly noisy. The temporal basis functions of the spectral model are exponential decays convolved with the AIF. Due to their convolution with the AIF, the shapes of spectral basis functions are much more constrained at these early time-points than those of B-splines; but for the same reason they can still describe the early parts of TACs well. This may explain why the spectral-model based reconstruction was found to produce better results at the early time-points in chapter 4. On the other hand, B-spline basis functions can describe a wider range of TAC shapes than spectral basis functions, and thus may introduce less bias than the spectral model into the

[1]The results presented in this chapter have been published in: G. P. Ralli, M. A. Chappell, D. R. McGowan, R. A. Sharma, G. S. Higgins and J. D. Fenwick. 4D-PET reconstruction using a spline-residue model with spatial and temporal roughness penalties. *Physics in Medicine & Biology*, 63(9):095013, 2018.

reconstruction process. However, the more flexible splines can also over-fit the data, potentially limiting the degree of noise suppression achievable compared to the spectral model. A temporal model that combines the noise suppressing capabilities of the spectral model, particularly in early time-frames, with the greater flexibility of spline functions might therefore be advantageous.

In this chapter an alternative linear model is proposed. This model, whose basis functions are B-spline basis functions convolved with the AIF, is termed the spline-residue model. Like the spectral model, convolution with the input function constrains the spline-residue basis function shapes in early time-frames, potentially enhancing the noise suppression achieved in the early frames by 4D reconstruction. But by using B-splines instead of exponential functions, the spline-residue model can describe a wider range of TAC shapes across the whole imaged time-course than can the spectral model, potentially reducing bias. O'Sullivan *et al.* [111] developed a similar (though not identical) linear model for post-reconstruction non-parametric analysis of radiotracer uptake kinetics in dPET images. To our knowledge, however, use of such a model for 4D-PET reconstruction has not previously been proposed or evaluated.

The spline-residue model-based 4D-PET reconstruction algorithm used here includes both spatial and temporal roughness penalties. The performance of spline-residue-based 4D-PET reconstruction is compared to that of a conventional (non-4D) MAP reconstruction algorithm, and to 4D-PET reconstruction based on the spectral model, adaptive-knot cubic B-splines and an irreversible two-tissue compartment model (2C3K) commonly used to analyse FMISO dPET data [180, 103]. Performance is measured using the image quality metrics defined in section 3.3.4.2, and parametric maps describing voxel-by-voxel compartment model fits to the image data.

5.2 Nonparametric Spline-Residue Description of dPET TACs

A TAC can quite generally be modelled as the convolution of the AIF with a residue function,

$$f(t) = K \int_0^t C_p(s) R(t-s) ds, \tag{5.1}$$

where $C_p(t)$ is the AIF, K is a proportionality constant interpreted as overall flow and $R(t)$ is the residue function, which describes the fraction of tracer remaining in the region at time t after entering it, and so provides information about the kinetics of radiotracer transport and metabolism processes [55].

The spline-residue model represents the residue function as a weighted sum of N_S cubic B-spline functions, $\zeta_k(t)$, with a Dirac delta function added to account for the finite blood volume in the region:

$$R(t) = \mu_0 \delta(t) + \sum_{k=1}^{N_s} \mu_k \zeta_k(t), \tag{5.2}$$

where μ_k is the coefficient of the k^{th} B-spline function $\zeta_k(t)$, the coefficient μ_0 equals V_b/K and V_b is the fractional blood volume. Thus from (5.1) the TAC is given by a weighted sum of spline-residue basis functions $\rho_l(t)$:

$$f(t, \boldsymbol{\theta}) = \sum_{l=0}^{N_S} \theta_l \rho_l(t), \tag{5.3}$$

where

$$\rho_l(t) = \begin{cases} C_p(t) & l = 0 \\ C_p(t) \otimes \zeta_l(t) & l > 0, \end{cases} \tag{5.4}$$

$\otimes$ denotes a convolution, and the coefficients θ_l are given by

$$\theta_l = \begin{cases} V_b & l = 0 \\ K\mu_l & l > 0. \end{cases} \tag{5.5}$$

Recall from section 2.4.3 that the compartment and spectral models describe the residue function as a weighted sum of exponential decays, with the former allowing both the decay constants and weighting factors to vary when fitting the model and the latter only allowing the weighting factors to vary.

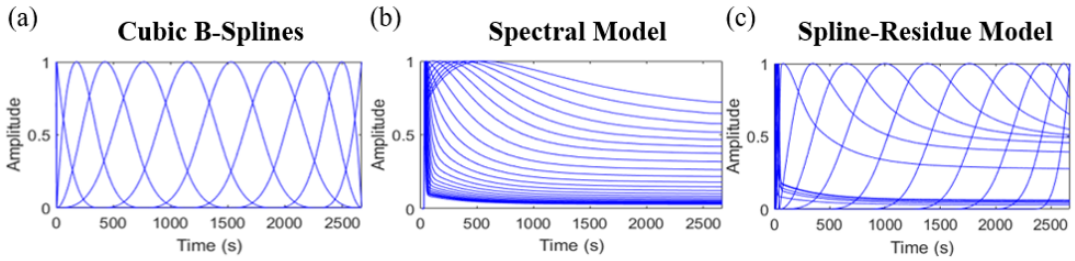


Figure 5.1. Comparison of typical shapes of normalized (a) cubic B-spline, (b) spectral model and (c) spline-residue model basis functions.

Spline-residue, cubic B-spline and spectral model basis functions are compared in Figure 5.1. The early spectral and spline-residue model basis functions are very similar, while the later basis functions are rather different.

5.3 Nested-MAP2: Temporally Regularised 4D-PET Reconstruction Algorithm for Linear Kinetic Models

Recall from section 2.4.1 that linear kinetic models represent the number of positron annihilation events in a given voxel j at time-frame m, x_{jm}, as a linear combination of N_k pre-defined basis functions:

$$x_{jm}(\boldsymbol{\theta}_j) = \sum_{k=1}^{N_k} \theta_{jk} B_{km}, \tag{5.6}$$

where the θ_{jk} is the coefficient of the k^{th} basis function $B_k(t)$ in voxel j, whose value at time-frame m, B_{km}, is given by

$$B_{km} = \int_{T_{ms}}^{T_{mf}} B_k(t) \exp(-\lambda t) dt, \tag{5.7}$$

where λ is the radiotracer decay constant and T_{ms} and T_{mf} are the start and end times of time-frame m.

To perform temporally regularized 4D-PET reconstructions the following objective function is used:

$$\Psi(\boldsymbol{\Theta}) = L(\boldsymbol{Y}|\boldsymbol{\Theta}) - \gamma\Gamma(\boldsymbol{\Theta}), \tag{5.8}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \dots \boldsymbol{\theta}_{N_V}]^T$, N_V is the number of voxels, $L(\boldsymbol{Y}|\boldsymbol{\Theta})$ is the log-likelihood function of the scanner data $\boldsymbol{Y}$ given $\boldsymbol{\Theta}$, defined in section 2.4.1, γ is a tunable hyperparameter controlling the trade-off between temporal smoothness and accurate TAC description, and $\Gamma(\boldsymbol{\Theta})$ is a temporal roughness penalty, defined as

$$\Gamma(\boldsymbol{\Theta}) = \sum_{j=1}^{N_V} \Lambda(\boldsymbol{\theta}_j), \tag{5.9}$$

with $\Lambda(\boldsymbol{\theta}_j)$ being the penalty function for voxel j.

To maximize $\Psi(\boldsymbol{\Theta})$ with respect to $\boldsymbol{\Theta}$, a methodology based on optimization transfer [85] is used. Recall from section 2.4.2 that in optimization transfer the determination of $\boldsymbol{\Theta}$ at iteration n is transferred to a surrogate function $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n)$ which minorizes the objective function:

$$Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n) \leq \Psi(\boldsymbol{\Theta}), \tag{5.10}$$

with equality if and only if $\boldsymbol{\Theta} = \boldsymbol{\Theta}^n$. Choosing $\boldsymbol{\Theta}^{n+1}$ as the value of $\boldsymbol{\Theta}$ maximizing $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n)$ implies that $\Psi(\boldsymbol{\Theta}^{n+1}) \geq \Psi(\boldsymbol{\Theta}^n)$, as discussed in section 2.4.2.

The surrogate function for $\Psi(\boldsymbol{\Theta})$ is obtained by subtracting $\gamma\Gamma(\boldsymbol{\Theta})$ from the surrogate function defined in equation (2.49), which was proposed by Wang and Qi [175] for the minorization of $L(\boldsymbol{Y}|\boldsymbol{\Theta})$:

$$Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n) = \sum_{j=1}^{N_V} \left(\sum_{i=1}^{N_{LOR}} P_{ij} \right) \left(\sum_{m=1}^{N_T} \left[\hat{x}_{jm}^{n+1} \ln(x_{jm}(\boldsymbol{\theta}_j)) - x_{jm}(\boldsymbol{\theta}_j) \right] - \gamma' \Lambda(\boldsymbol{\theta}_j) \right), \tag{5.11}$$

where the P_{ij} values are the system matrix elements, N_{LOR} is the number of detector pairs, N_T is the number of time-frames, $\gamma' = \frac{\gamma}{\sum_{i=1}^{N_{LOR}} P_{ij}}$ and $\hat{x}_{jm}^{n+1}$ is the MLEM update applied to the activity image in frame m given by the current parameter estimate, $x_{jm}(\boldsymbol{\theta}_j^n)$:

$$\hat{x}_{jm}^{n+1} = \frac{x_{jm}(\boldsymbol{\theta}_j^n)}{\sum_{i=1}^{N_{LOR}} P_{ij}} \sum_{i=1}^{N_{LOR}} P_{ij} \frac{y_i}{\sum_{k=1}^{N_V} \left[P_{ik} x_{km}(\boldsymbol{\theta}_k^n) + r_i + s_i \right]}. \tag{5.12}$$

Because $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^n)$ is separable in voxels, the parameter values $\boldsymbol{\theta}_j$ can be obtained using

$$\boldsymbol{\theta}_j^{n+1} = \max_{\boldsymbol{\theta}_j} \left\{ \sum_{m=1}^{N_T} \left[\hat{x}_{jm}^{n+1} \ln(x_{jm}(\boldsymbol{\theta}_j)) - x_{jm}(\boldsymbol{\theta}_j) \right] - \gamma' \Lambda(\boldsymbol{\theta}_j) \right\}. \tag{5.13}$$

Instead of maximizing (5.13), an adaptation of a method proposed by Matthews *et al.* [99] for the unregularized case, which uses a WLS approach to calculate $\boldsymbol{\theta}_j$, is used here. For the temporally regularized case, $\boldsymbol{\theta}_j$ is obtained by minimizing the penalized WLS error

$$\boldsymbol{\theta}_j^{n+1} = \min_{\boldsymbol{\theta}_j} \left\{ \sum_{m=1}^{N_T} w_{jm} \left[\hat{x}_{jm}^{n+1} - x_{jm}(\boldsymbol{\theta}_j) \right]^2 + \gamma' \Lambda(\boldsymbol{\theta}_j) \right\}, \tag{5.14}$$

with weighting factors

$$w_{jm} = \frac{1}{x_{jm}(\boldsymbol{\theta}_j^n)}. \tag{5.15}$$

Wang and Qi have noted in [176] that monotonic convergence to the ML solution is not guaranteed when the model fitting step is converted to a WLS problem, however convergence has been observed in practice [99].

Temporal roughness penalties have been used in previous 4D-PET reconstruction studies, but there is no consensus on what form $\Lambda(\boldsymbol{\theta}_j)$ should take. Two penalty functions are explored in this work, the first being

$$\Lambda(\boldsymbol{\theta}_j) = |\boldsymbol{\theta}_j|^2, \tag{5.16}$$

which corresponds to L_2 regularization, and the second

$$\Lambda(\boldsymbol{\theta}_j) = \int_0^{T_{scan}} \left(\frac{\partial^2 f(t, \boldsymbol{\theta}_j)}{\partial t^2} \right)^2 dt, \tag{5.17}$$

where T_{scan} is the scan duration and $f(t, \boldsymbol{\theta}_j)$ is the fitted TAC. The second penalty function is often used to fit splines to noisy data and has been applied to spline-based 4D-PET reconstruction [109, 89], though not in an optimization transfer framework.

Both penalty functions can be expressed as

$$\Lambda(\boldsymbol{\theta}_j) = \boldsymbol{\theta}_j^T \boldsymbol{\Omega} \boldsymbol{\theta}_j. \tag{5.18}$$

For the penalty function in (5.16) $\boldsymbol{\Omega}$ is the identity matrix $\boldsymbol{I}$, while for (5.17) the elements of $\boldsymbol{\Omega}$, Ω_{ab}, are

$$\Omega_{ab} = \int_0^{T_{scan}} \ddot{B}_a(t) \ddot{B}_b(t) dt, \tag{5.19}$$

where $\ddot{B}_k(t)$ is the second-order time-derivative of basis function k. Using this notation the cost function in equation (5.14) can be re-written as a Tikhonov regularization problem

$$\boldsymbol{\theta}_j^{n+1} = \min_{\boldsymbol{\theta}_j} \left\{ \sum_{m=1}^{N_T} w_{jm} \left[\hat{x}_{jm}^{n+1} - x_{jm}(\boldsymbol{\theta}_j) \right]^2 + \gamma' \boldsymbol{\theta}_j^T \boldsymbol{\Omega} \boldsymbol{\theta}_j \right\}, \tag{5.20}$$

and solved using the equation [154]

$$\boldsymbol{\theta}_j^{n+1} = (\boldsymbol{B}^T \boldsymbol{W} \boldsymbol{B} + \gamma' \boldsymbol{\Omega})^{-1} \boldsymbol{B}^T \boldsymbol{W} \hat{\boldsymbol{x}}_{jm}^{n+1}, \tag{5.21}$$

where $\boldsymbol{B}$ has the B_{km} elements defined in equation (5.7), $\boldsymbol{W}$ is a diagonal matrix $\text{diag}(w_{j1}, w_{j2}, ..., w_{jN_T})$ and $\hat{\boldsymbol{x}}_{jm}^{n+1}$ is an $N_T \times 1$ vector with m^{th} element $\hat{x}_{jm}^{n+1}$.

To determine the best value of γ' to use in (5.20) for voxel j, γ'_j, a range of γ' values can be defined in advance, and the one producing the fit with the lowest generalized cross validation (GCV) score taken as optimal for this voxel [167]:

$$\gamma'_j = \min_{\gamma'} \{GCV(\gamma')\} = \min_{\gamma'} \left\{ \frac{\sum_{m=1}^{N_T} w_{jm} \left[\hat{x}_{jm}^{n+1} - x_{jm}(\boldsymbol{\theta}_j) \right]^2}{\left[\text{Tr}(\boldsymbol{I} - \boldsymbol{B}(\boldsymbol{B}^T \boldsymbol{W} \boldsymbol{B} + \gamma' \boldsymbol{\Omega})^{-1} \boldsymbol{B}^T \boldsymbol{W}) \right]^2} \right\}, \tag{5.22}$$

where Tr(...) denotes the matrix trace. Because regularised linear model fitting suppresses the contribution of the individual model parameters, the regularised model behaves as though it has a lower (non-integer) number of degrees of freedom to the un-regularized version of the model. These are often referred to as the effective degrees of freedom, and can be calculated using the expression in the denominator of equation (5.22) [154, 167]. GCV is thus closely related to reduced-χ^2 analysis for model selection, in which the model that minimizes the ratio of the RSS fitting error to the number of degrees of freedom is taken to be optimal. The model fitting step is much faster than the image update step, and so selection of the γ' value for each voxel TAC in this semi-automatic manner does not greatly slow down the reconstruction.

Spatial regularization can be built into this reconstruction methodology by replacing the image update step in (5.12) with one iteration of a 2D- or 3D-MAP reconstruction algorithm.

The temporally regularized 4D-PET reconstruction algorithm for linear kinetic models, subsequently referred to as nested-MAP2 to distinguish it from the earlier nested-MAP algorithm used in chapter 4, can be summarized as follows.

1. Start with an initial dPET image sequence estimate, in this work a sequence of uniform images with the radioactivity concentration in each voxel set to 100 Bq/cc.
2. Update each image with one iteration of the MAP algorithm. Here the MAP updates were performed using the algorithm described in section 3.3.3.1.
3. Fit a temporal model to each voxel TAC via (5.21), using either a fixed value of γ' or a range of γ' values and subsequently selecting the best value via GCV.
4. Return to step 2 using the image voxel values predicted by the fitted model, $x_{jm}(\boldsymbol{\theta}_j^n)$, as the seed for next MAP update.

5.4 Methods

5.4.1 Digital Phantom Data

The simulated data from the FMISO NSCLC dPET phantom, described in chapter 3, was used to compare each of the reconstruction algorithms used in this work. A second phantom with the same patient geometry as the original FMISO NSCLC dPET phantom, but using the 2C3K model fits to the patient data (shown in section 3.4.2.1) as the ground-truth tissue TACs, was also generated. This second 'simplified' phantom was binned into the same time-frame sequence as the 'realistic' FMISO NSCLC dPET phantom used in chapters 3 and 4, and 50 noise realizations of simulated dPET scanner data were generated from this phantom using the same methodology described in chapter 3.

For the remainder of this chapter these two phantoms are referred to as the simplified and realistic phantoms. The purpose of using these two different phantoms was to study the performance of linear model-based versus 2C3K-based 4D-PET reconstruction when all the underlying TACs take realistic shapes (realistic phantom), and when they are all described by fits of the 2C3K model (simplified phantom).

5.4.2 Image Reconstruction

From each simulated realization of FMISO patient phantom PET scanner data, dPET image sequences were reconstructed in 3.1×3.1×2.0 mm^3 voxels using both the conventional MAP algorithm introduced in section 3.3.3.1 and the 4D nested-MAP2 algorithm proposed in section 5.3. Each algorithm was run for 30 iterations. Spatial regularization was used for all reconstructions, as the results in chapter 4 show that it is advantageous to do so, provided the parameters are well chosen. As in chapter 4, the spatial regularization parameters β and δ used in the Lange penalty function defined in (2.33) were both set to 0.1.

Nested-MAP2 4D reconstructions were performed using the non-linear 2C3K model, and the linear adaptive-knot cubic B-spline, spectral and spline-residue models. For each linear model, reconstructions were carried out using the temporal regularisation penalties of equations (5.16) and (5.17), and with no temporal regularisation. For 2C3K model-based reconstructions no temporal regularization was used because this model is much more constrained than the linear ones. For nested-MAP2 reconstruction based on the 2C3K model, the non-linear model fitting step was performed using the Levenberg-Marquardt algorithm instead of (5.21).

Spectral, spline-residue and 2C3K model-based 4D reconstructions require image-derived or blood-sampled AIFs, which were obtained here by fitting the three-exponential model of Feng *et al.* [40] to TACs obtained from conventionally (not 4D) reconstructed MAP images, for ROIs placed in the left ventricle.

Spectral model-based reconstructions were carried out using 100 basis functions with exponential decay constants spaced logarithmically between 1.1×10^{-4} s^{-1} (the decay constant of ^{18}F) and 0.01 s^{-1}. It should be noted that the model fitting formula (5.21) used in this chapter does not have any non-negativity constraints. Strictly speaking the spectral model is defined such that all the basis function weights are ≥ 0 [25]. In potentially allowing the weights to go negative, the 'spectral model' used in this chapter is slightly different to the spectral model that is used in both chapter 4 and more commonly in the wider literature. However, there are some situations, such as when the AIF is replaced by a TAC from a reference region,

where it is possible to get negative spectral coefficients, and applications which do not impose positivity on the coefficients have been proposed [164].

For spline-based reconstructions, voxel-specific knot locations were selected using the adaptive-knot placement algorithm proposed in section 4.2.4, which for cubic splines places knots along equal segments of the integral of the 4^{th} root of the 4^{th} derivative of a TAC. For each voxel 11 free knots were positioned by applying the algorithm to the TAC obtained for that voxel from conventional MAP-reconstructed images.

Basis functions for the spline-residue model were obtained by placing 4 knots at the beginning and end of each TAC, as well as the point where the TAC starts to rise, to handle discontinuities, and positioning an additional 6 uniformly spaced knots between the initial rise and end points of the TAC. The B-splines associated with these knots were convolved with the AIF to calculate the spline-residue basis functions. Preliminary work fitting the spline-residue model to synthetic noisy FMISO TACs led to the choice of 6 additional knots, this number generating fits that best matched the ground-truth. Uniformly spaced knots performed well, perhaps because the residue function for a given TAC varies considerably less than the TAC itself.

For each linear model and temporal roughness penalty used in the 4D reconstruction algorithms, the range of γ' values $\{\gamma' = 0.001, 0.002, \ldots, 0.01\}$ all produced good fits to TACs from conventional MAP-reconstructed patient phantom images. At each iteration of the temporally regularized nested-MAP2 reconstructions, therefore, the γ' value used for each voxel was individually selected from those ten as the one that minimized the GCV score of the model fit to that voxel's TAC, as described in section 5.3.

Allowing the γ'_j values, and thereby the objective function, to vary at each iteration could affect the convergence. To check this, for each linear model the temporally regularized reconstructions were re-run with fixed γ'_j values, taken as those obtained in the final iteration of the corresponding reconstruction where γ'_j was free to vary.

5.4.3 Image Analysis

5.4.3.1 Image Quality Metrics

As in chapter 4, the image bias and noise measures defined in section 3.3.4.2 were used to calculate the bias and noise in each voxel of the images produced at every iteration of each of the reconstruction algorithms. Overall bias and noise were characterised as the mean absolute bias and mean image noise ($\langle\sigma_w\rangle$) averaged over all voxels within the patient (the 'whole patient' region). The same measures were also averaged over the tumour region alone. To assess the performance of the algorithms in the early time-frames, values of the overall bias and noise in the whole patient and tumour regions were also calculated for just the first 120 s of the scan.

Using WLS to fit the models during the reconstruction, and allowing the γ'_j values to vary during the temporally regularized reconstructions, may adversely affect the algorithm convergence. To assess the convergence of the reconstruction process, the mean square error (MSE) was calculated for each image voxel j at time-frame m and iteration n of the nested-MAP2 reconstructions

$$MSE^n_{jm} = \frac{1}{N_r}\sum_{k=1}^{N_r}(a^{true}_{jm} - a^n_{jm,k})^2, \tag{5.23}$$

where N_r is the number of noise realizations, a^{true}_{jm} is the true activity in voxel j at time-frame m and $a^n_{jm,k}$ is the reconstructed activity of the k^{th} noise realization in voxel j at iteration n and time-frame m. At each iteration a weighted sum of MSE^n_{jm} over all time-frames was calculated for every individual voxel, and these values were summed over all image voxels to create a single total MSE measure, TMSE:

$$TMSE^n = \sum_{j=1}^{N_V}\sum_{m=1}^{N_T}\frac{MSE^n_{jm}\Delta T_m}{a^{true}_{jm}}, \tag{5.24}$$

where N_V and N_T are the numbers of voxels and time-frames, respectively, and ΔT_m is the duration of time-frame m. To check that the images did not change substantially after the first 30 reconstruction iterations explored throughout most of this study, each nested-MAP2 reconstruction was run for a further 10 iterations. Fractional changes in TMSE from one iteration n to the next $n+1$ were calculated as

$$[\text{Fractional TMSE change at iteration } n] = \frac{TMSE^n - TMSE^{n+1}}{TMSE^n}, \tag{5.25}$$

and plotted as a function of iteration number for $n = 1$ to $n = 39$, positive TMSE changes corresponding to reductions in the total error, and negative changes to increases. Additionally, the fractional change between iterations 30 and 40 was calculated as

$$[\text{Fractional TMSE change from iteration 30 to 40}] = \frac{TMSE^{30} - TMSE^{40}}{TMSE^{30}}. \tag{5.26}$$

5.4.3.2 Parametric Map Quality Metrics

FMISO uptake kinetics are often determined using the 2C3K model, shown schematically in Figure 3.2 [180, 103]. Following this approach, the 2C3K model was fitted to voxel TACs obtained from all the image sequences reconstructed using the different algorithms. The voxels studied were those lying within the phantom sub-volume shown in Figure 5.2, which contains the tumour-like region and surrounding lung. The kinetics analysis was restricted to this region as in onclolgical dPET studies it is generally the tumour kinetics (and in certain applications, such as radiotherapy treatment planning, those of the immediate surroundings of the tumour) that are of the greatest interest. The accuracy and precision of uptake kinetics as characterised by the 2C3K model fits to the reconstructed images versus fits of the same model to the ground-truth phantom TACs were then determined, studying the 2C3K model rate-constants shown in Figure 3.2, together with the flux constant

$$k_{flux} = \frac{K_1 k_3}{k_2 + k_3}. \tag{5.27}$$

The bias and noise of each fitted parameter $q = K_1, k_2, k_3, k_{flux}$ were calculated for each voxel j lying within the image sub-volume:

$$[\text{Parameter Bias}]_j = |\bar{q}_j - q_j^{true}|, \tag{5.28}$$

$$[\text{Parameter Noise}]_j = \sigma_{q_j}, \tag{5.29}$$

where q_j^{true} is the ground-truth value of kinetic parameter q in voxel j, $\bar{q}_j$ is the mean of the q_j values obtained for voxel j from each of the 50 noise realizations

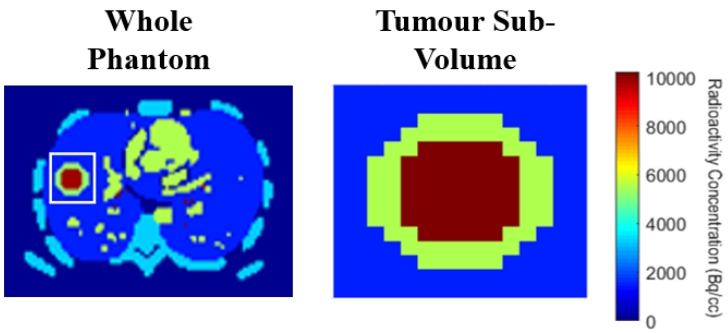


Figure 5.2. Image of the final time-frame of the realistic phantom, with a white box marking the sub-volume of the phantom for which parametric maps were calculated.

and σ_{q_j} is the standard deviation of these 50 q_j values. The $[\text{Parameter Bias}]_j$ and $[\text{Parameter Noise}]_j$ measures were then averaged over all the voxels within the analysed phantom sub-volume to characterize the overall performance of each reconstruction algorithm. These metrics were also averaged over a region containing hypoxic tumour alone, to assess the performance of the algorithms specifically within the tumour region.

5.5 Results

5.5.1 Image Quality Metrics

Iteration-by-iteration plots of image bias versus noise, averaged over all time-frames and patient voxels, are shown in Figure 5.3 for realistic and simplified phantom image sequences reconstructed using temporally regularized nested-MAP2 4D algorithms based on the cubic spline, spectral and spline-residue linear models. The purpose of this figure is to compare the performance of the two temporal roughness penalties for each linear model separately, and so each plot compares results obtained for one phantom and one reconstruction algorithm using either no temporal roughness penalty or the penalty functions of equations (5.16) or (5.17). The γ'_j hyperparameters were free to vary during each of the reconstructions shown.

For spectral and spline-residue-based 4D reconstructions, the $|\boldsymbol{\theta}|^2$ temporal roughness penalty of equation (5.16) produced substantially less noisy images than

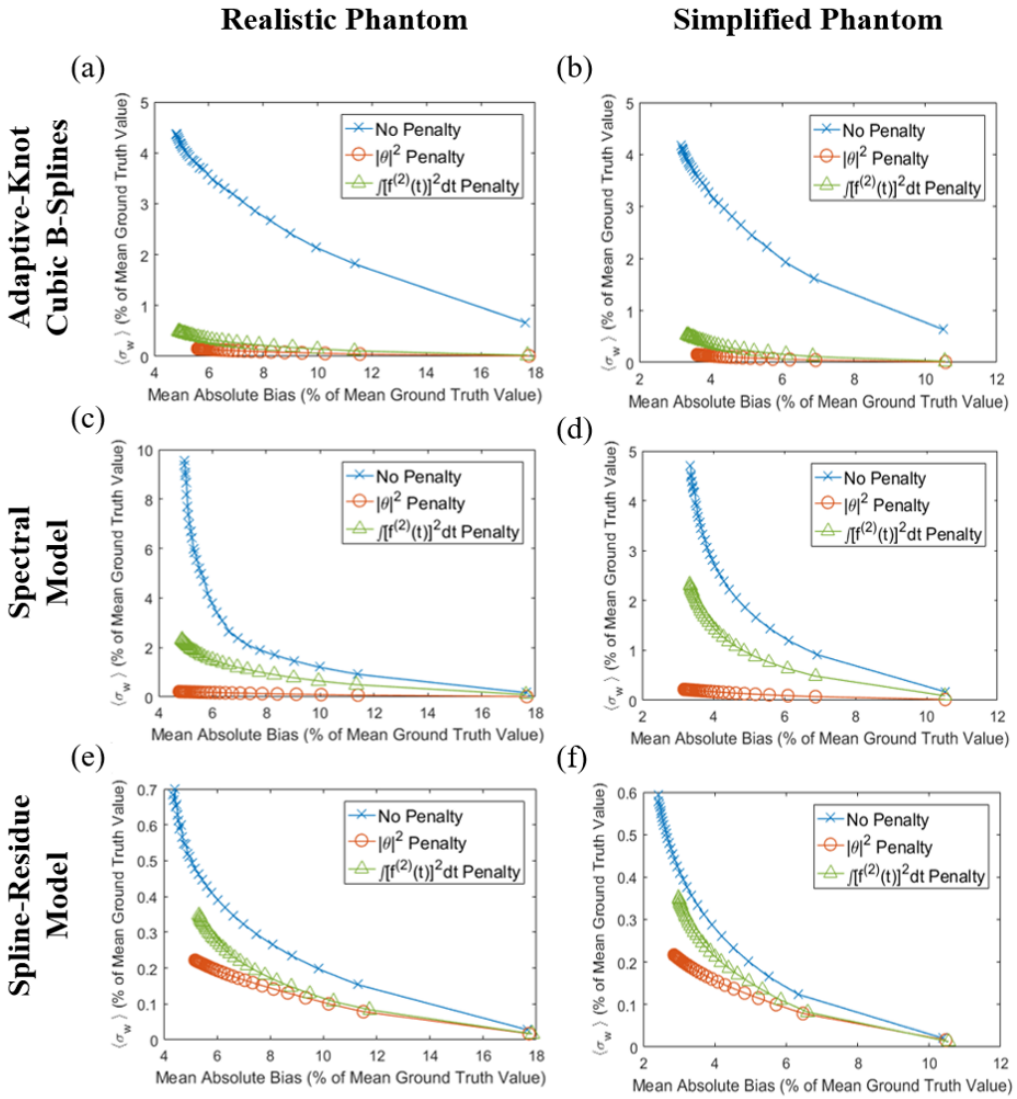


Figure 5.3. Comparison of image reconstruction results obtained for the realistic (first column) and simplified (second column) phantoms using 4D algorithms based on the (a) and (b) cubic splines, (c) and (d) spectral and (e) and (f) spline-residue models with different temporal roughness penalties. The plots show image noise (weighted standard deviation, σ_w) versus absolute bias, averaged over all time-frames and patient voxels, at each reconstruction iteration, and have different scales to allow clear visualisation of performance differences between the different penalty functions. The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown.

the other penalty options, at similar levels of bias, and was therefore considered the best penalty function for these algorithms. For cubic spline-based 4D reconstructions, however, the integrated square derivative penalty function of equation (5.17) was viewed as the best penalty function, since it produced the least biased images at

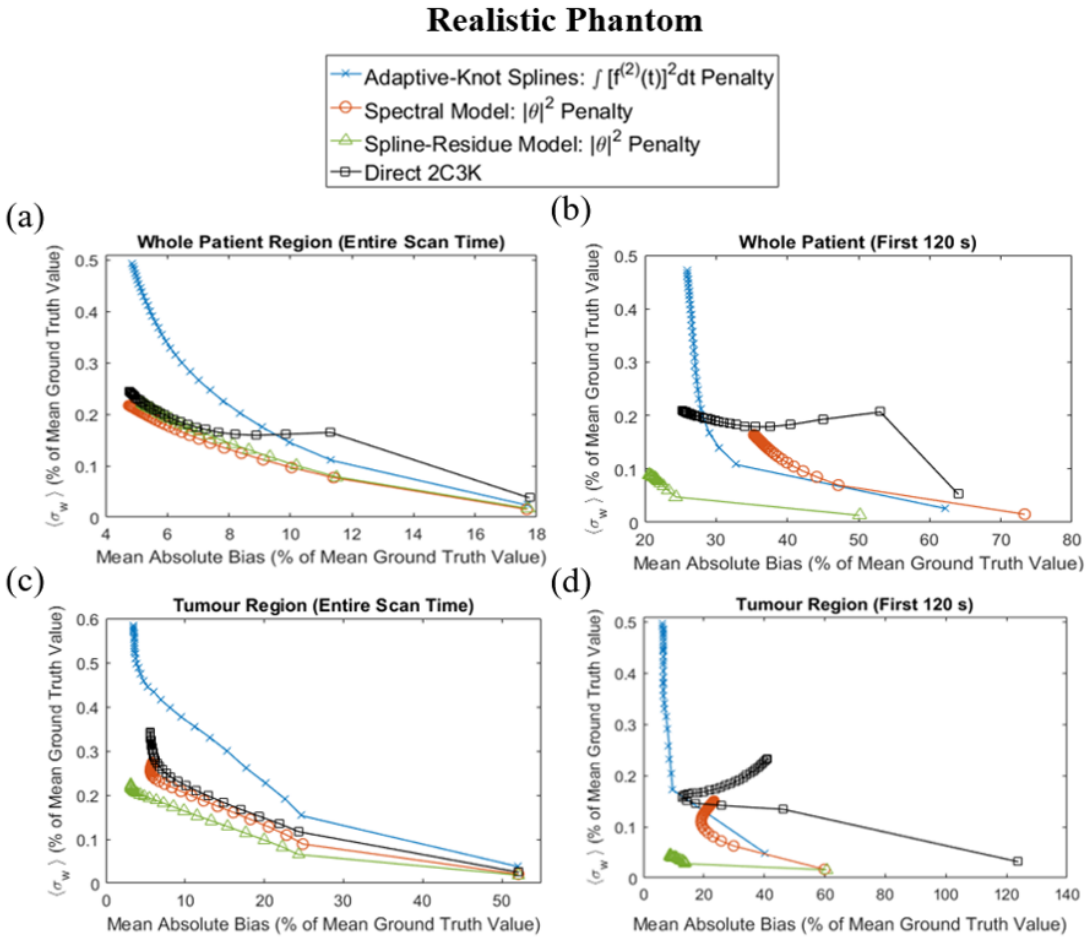


Figure 5.4. Comparison of image quality metrics for 4D reconstructions of the realistic phantom based on linear models (with their optimal temporal roughness penalty) and on the 2C3K model. The noise and bias metrics are averaged over the whole patient (a) and (b) and tumour (c) and (d) phantom regions, and over the entire scan time (a) and (c) and first 120 s (b) and (d). The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown.

noise-levels only slightly higher than obtained using the $|\boldsymbol{\theta}|^2$ penalty.

In Figure 5.4, noise is plotted against bias for images of the realistic phantom reconstructed using each linear model-based 4D algorithm and its associated optimal temporal roughness penalty, and using with the 2C3K-based 4D algorithm. Separate plots are shown for noise and bias measures averaged over the whole patient or the tumour regions, and averaged over the whole scan time or just the first 120 s. The γ'_j hyperparameters were free to vary during each of the reconstructions shown. Image quality curves for reconstructions based on the spline-residue model were

substantially better than those obtained for reconstructions based on cubic splines or the 2C3K model. And at early time-points and within the tumour region, the image quality curves for the spline-residue-based algorithm were also better than those for reconstructions based on the spectral model. However, the spectral model had a slight edge when the image quality metrics were averaged across the whole phantom and scan duration.

Corresponding data are shown in Figure 5.5 for reconstructions of the simplified phantom. For this phantom the spline-residue model produced slightly better results than the spectral model when considered across the whole phantom and scan duration; but 4D reconstructions based on the 2C3K model achieved the lowest bias values, unsurprisingly since the phantom kinetics are 2C3K-based.

Figure 5.6 shows fractional changes in TMSE as a function of iteration number for the 2C3K- and linear model-based reconstructions of the realistic phantom, using the optimal temporal roughness penalties for the linear models and selecting the optimal γ'_j hyperparameter at each reconstruction iteration using GCV. After 30 iterations the fractional change in TMSE per iteration was very small for reconstructions based on the spectral, spline-residue and 2C3K models. Total fractional changes in TMSE between iterations 30 and 40, calculated with (5.26), were 0.013 for the spectral model, 0.017 for the spline-residue model, 0.007 for the 2C3K model and -0.030 for cubic splines. Thus continuing reconstruction beyond 30 iterations led to small improvements at best, and in the case of the spline-based reconstruction a 3% worsening in TMSE, most likely due to noise amplification at the later iterations.

Figure 5.7 compares the fractional changes in TMSE as a function of iteration number for the linear model-based reconstructions of the realistic phantom (using the optimal penalty function for each model) with γ'_j fixed and free to vary. The results in this figure show that beyond 5 iterations the convergence of reconstructions obtained with fixed and varying γ'_j was identical.

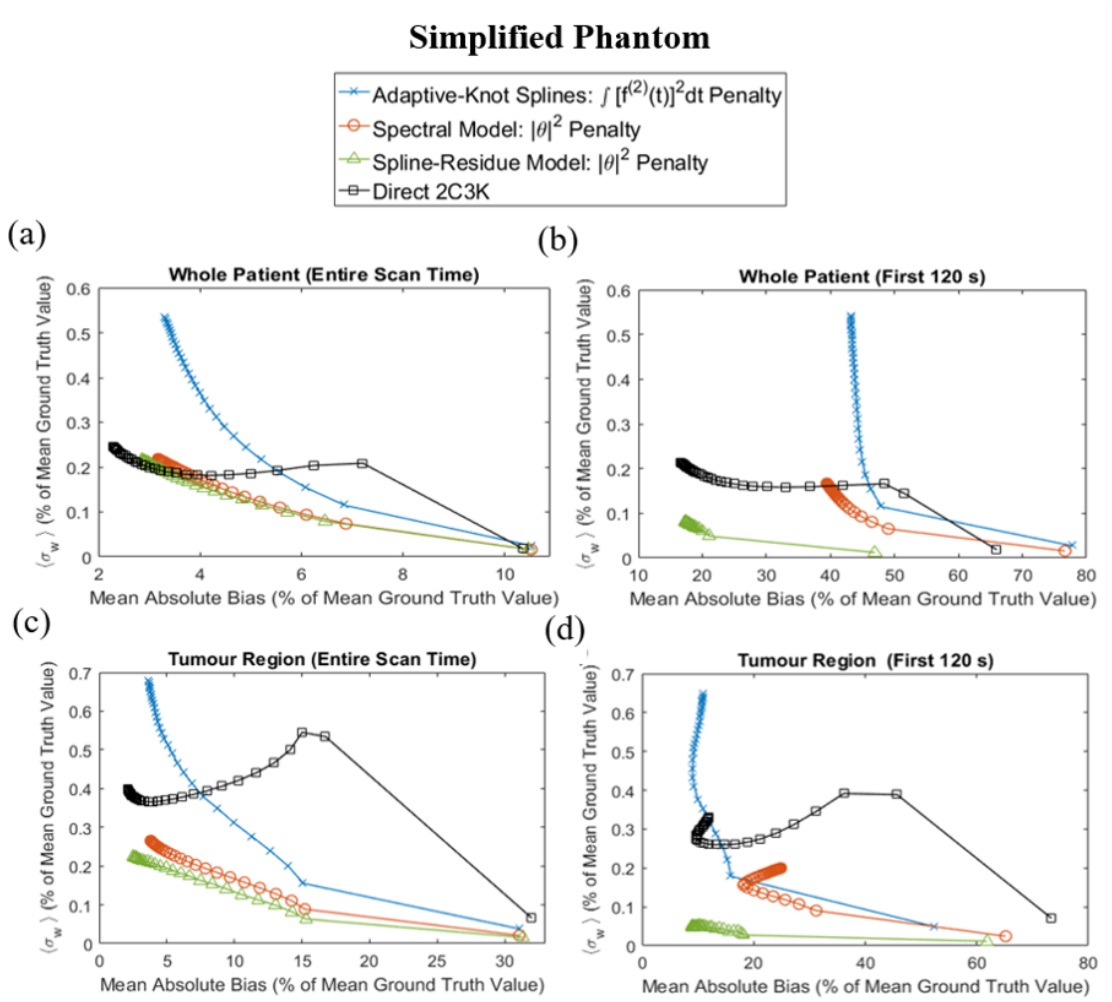


Figure 5.5. Comparison of image quality metrics for 4D reconstructions of the simplified phantom based on linear models (with their optimal temporal roughness penalty) and on the 2C3K model. The noise and bias metrics are averaged over the whole patient (a) and (b) and tumour (c) and (d) phantom regions, and over the entire scan time (a) and (c) and first 120 s (b) and (d). The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown.

5.5.2 Parametric Maps

Figure 5.8 shows a single noise realisation of the k_{flux} parametric maps of the realistic phantom sub-region shown in Figure 5.2, obtained from image sequences reconstructed using the conventional (non-4D) MAP algorithm and nested-MAP2 4D algorithms based on the 2C3K model and linear models used with their optimal temporal roughness penalties. Figure 5.9 shows voxel-by-voxel spatial plots of absolute bias and noise in k_{flux} parametric maps, calculated over all noise

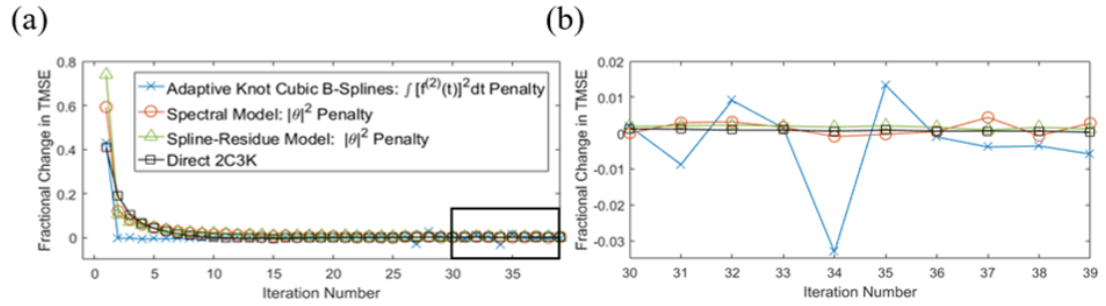


Figure 5.6. (a) Fractional iteration-to-iteration change in total mean square error (TMSE) versus iteration number, plotted for 2C3K and linear model-based nested-MAP2 reconstructions (with the optimal temporal roughness penalties and the γ'_j hyperparameters free to vary) of the realistic phantom. (b) A magnification of the rectangular region outlined in (a).

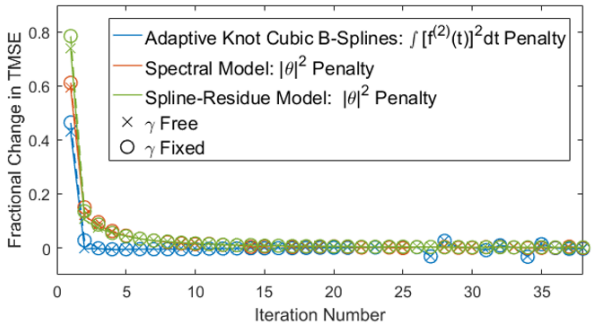


Figure 5.7. Iteration-by-iteration fractional change in total mean square error (TMSE) plotted as a function of nested-MAP2 iteration number, for linear-model based reconstructions of the realistic phantom (using the optimal temporal roughness penalty for each linear model) with γ'_j fixed and free to vary.

realisations of the realistic phantom. The voxel-by-voxel noise and bias plots calculated for the simplified phantom were similar. The boundary between the hypoxic core and normoxic perimeter of the tumour appears to have a high bias in all of the reconstructions.

To further illustrate the performance of the compartment model fits to the reconstructed TACs, Figure 5.10 shows examples of 2C3K model fits to a single noise realisation of voxel TACs taken from the reconstructed images of the realistic phantom. Fits to two voxel TACs are shown, one from the centre of the hypoxic tumour region and a second from the healthy lung region. This figure shows that the early time-points are where the greatest bias in the model fits occurs. This bias is particularly high for the fits to the TACs from the conventional MAP-reconstructed

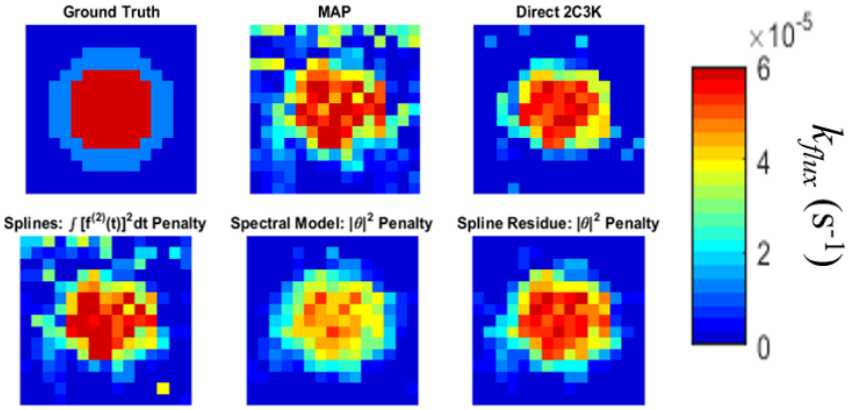


Figure 5.8. Single noise realisation of k_{flux} parametric maps of the subregion of the realistic phantom obtained from image sequences reconstructed using the conventional (non-4D) MAP algorithm, as well as using the 2C3K and linear model-based nested-MAP2 reconstructions (with their optimal temporal roughness penalties).

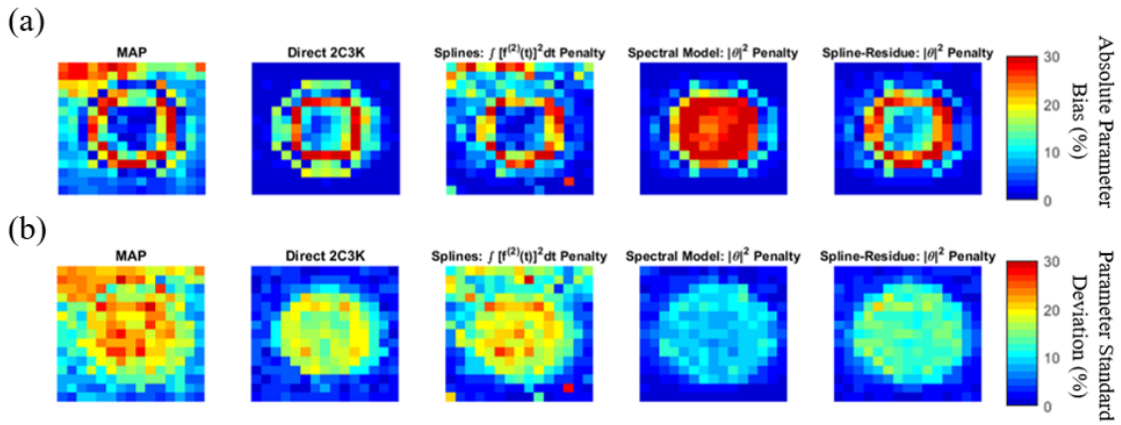


Figure 5.9. Voxel-by-voxel plots of (a) the absolute bias and (b) the noise (standard deviation) in the k_{flux} parametric maps of the subregion of the realistic phantom obtained from image sequences reconstructed using the conventional (non-4D) MAP algorithm, as well as using the 2C3K and linear model-based nested-MAP2 reconstructions (with their optimal temporal roughness penalties). Bias and standard deviation values are expressed as percentages of the average ground-truth k_{flux} value across the phantom sub-region.

images. The 2C3K fits to the voxel TACs from the spline-residue and 2C3K model-based reconstructions appear to give the best descriptions of the true 2C3K models.

The average normalised absolute bias and noise (standard deviation) in 2C3K model parameters, for fits to the TACs of every voxel of the realistic phantom sub-region of Figure 5.2 in images reconstructed using the different algorithms, is shown in Figure 5.11 for the individual 2C3K rate constants and the composite k_{flux}

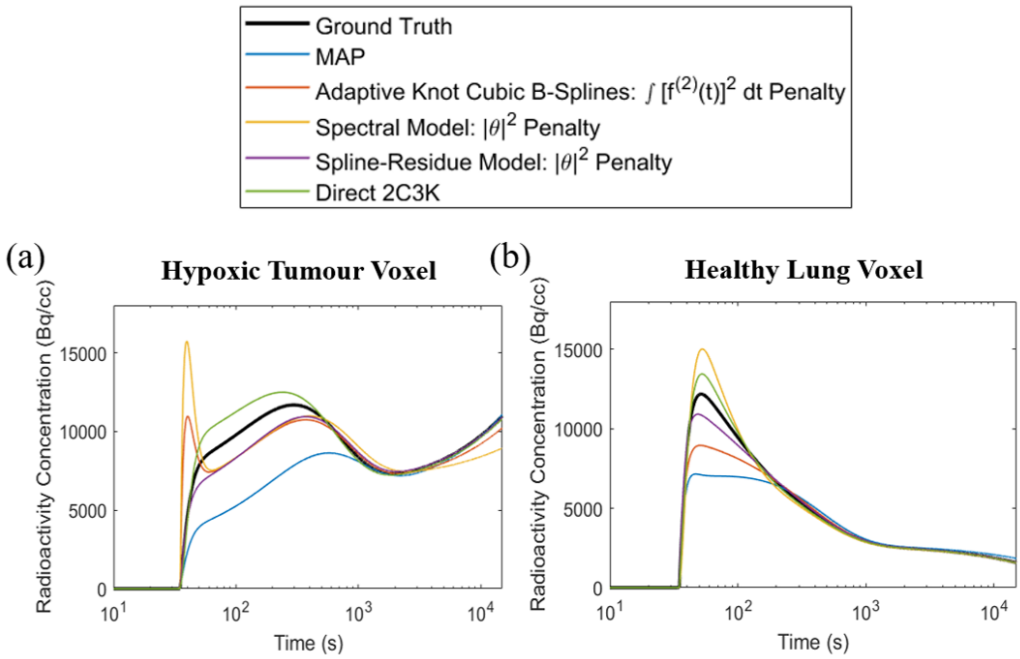


Figure 5.10. Fits of the 2C3K model to voxel TACs taken from reconstructed images of the realistic phantom. Voxel TACs taken from the centre of the hypoxic tumour region (a) and the healthy lung region (b) are shown. The 2C3K fits to the true phantom voxel TACs are also shown for comparison.

parameter. For each parameter, these values are normalised as fractions of the values achieved by conventional (non-4D) MAP reconstruction for the same parameter, thus showing the extent to which each nested-MAP2 reconstruction improves on the conventional MAP reconstruction for each kinetic parameter. Equivalent plots for the simplified phantom are shown in Figure 5.12.

Hypoxic tissue is know to be more resistant to both chemotherapy and radiotherapy, and thus the accurate characterisation of hypoxic tumour tissue is currently an active area of research [160]. The phantoms used in this chapter are based on the FMISO, which was developed to image tumour hypoxia. Tumour regions with high FMISO uptake are hypoxic but not necrotic, and are thus likely to be more resistant to treatment. It is thus useful to compare the performance of the different reconstructions in the hypoxic region of the phantom tumours. Absolute bias and noise values of the kinetic parameters averaged over the hypoxic tumour sub-volume alone are shown in Figure 5.13 for the realistic and simplified phantoms.

For the realistic phantom it can be seen from Figure 5.11 that for 5 of the 8 combinations of bias/noise and kinetic parameters analysed, bias or noise averaged across the whole phantom was reduced more than 50% by using 4D reconstruction based on the spline-residue model rather than conventional (non-4D) MAP-reconstruction. Furthermore, for 5/8 combinations spline-residue 4D reconstruction produced better results than any of the other reconstruction algorithms studied, while spectral-based 4D reconstruction produced the best results for 2/8. Compared to conventional reconstruction, spline-residue-based 4D reconstruction did not notably increase bias or noise for any of the 8 combinations, whereas the bias and noise in fitted K_1 parameter values rose substantially above

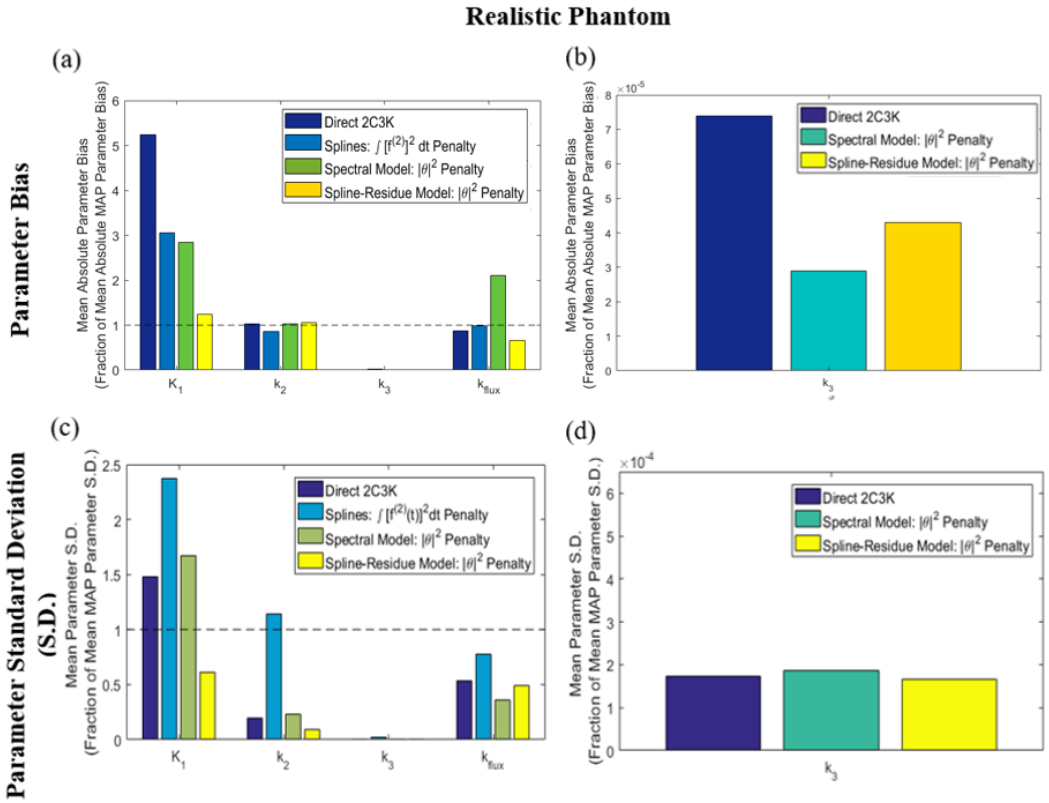


Figure 5.11. (a) Absolute bias and (c) noise (standard deviation) of fitted parameter values in parametric maps derived from the 4D-PET reconstructed images of the realistic phantom, averaged over the entire parametric map of the image sub-region shown in Figure 5.2. Values are plotted for each of the kinetic parameters as fractions of those obtained for the same parameter from analysis of conventional (non-4D) MAP-reconstructed images. Results for the k_3 parametric maps obtained from the 2C3K, spectral and spline-residue model based 4D reconstructions are re-plotted on larger scales in (b) and (d) to make them visible.

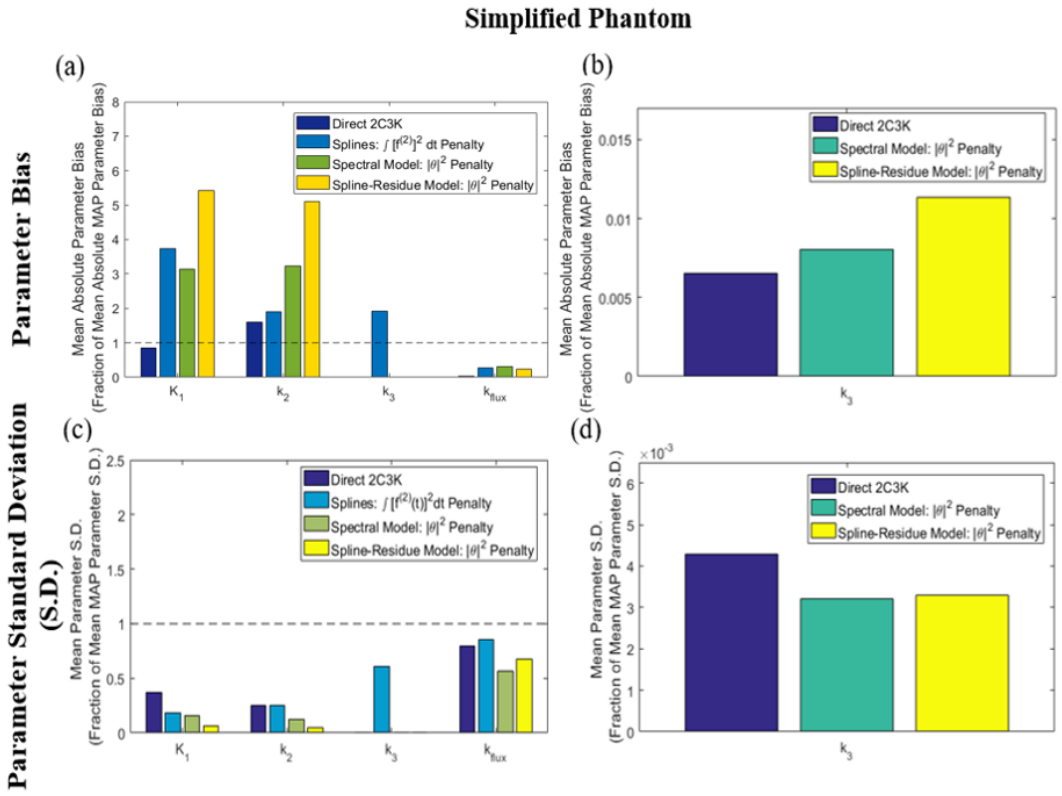


Figure 5.12. (a) Absolute bias and (c) noise (standard deviation) of fitted parameter values in parametric maps derived from the 4D-PET reconstructed images of the simplified phantom, averaged over the entire parametric map of the image sub-region shown in Figure 5.2. Values are plotted for each of the kinetic parameters as fractions of those obtained for the same parameter from analysis of conventional (non-4D) MAP-reconstructed images. Results for the k_3 parametric maps obtained from the 2C3K, spectral and spline-residue model based 4D reconstructions are re-plotted on larger scales in (b) and (d) to make them visible.

the conventional MAP levels when 4D reconstruction was carried out using the other temporal models. 4D reconstruction based on the 2C3K model generated the most biased kinetic parameters, and higher levels of noise than spline-residue-based reconstruction. From Figure 5.13 it can be seen that for the hypoxic tumour region alone, the spline-residue- and spectral-based 4D reconstructions each achieved the best results for 3/8 of the bias/noise and kinetic parameter combinations analysed, and cubic spline-based 4D reconstruction the best results for 2/8.

For the simplified phantom, whose kinetics entirely follow the 2C3K model, 4D reconstruction based on the 2C3K model unsurprisingly achieved the lowest levels of bias for fitted 2C3K model kinetics parameters (see Figures 5.12 and 5.13).

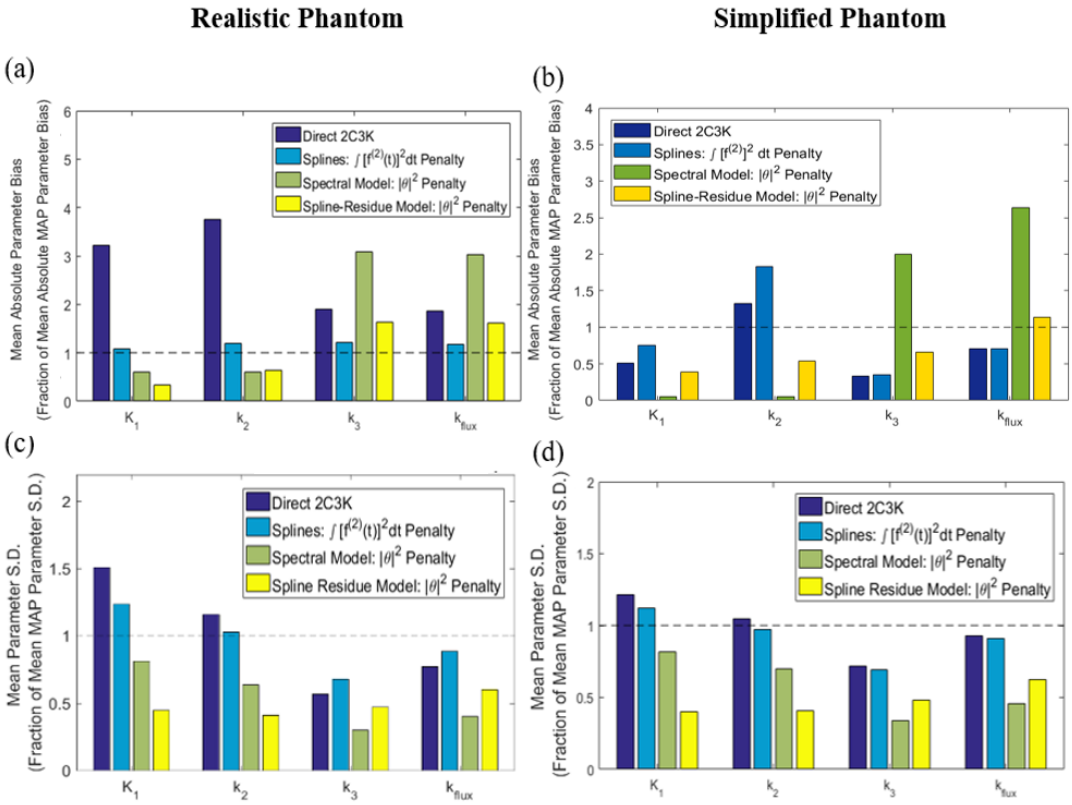


Figure 5.13. (a, b) Absolute bias and (c, d) noise (standard deviation) of fitted parameter values in parametric maps derived from the 4D-PET reconstructed images of the realistic phantom, averaged over the hypoxic tumour region alone. For each of the kinetic parameters, the values plotted are fractions of those obtained for the same parameters from analysis of conventional (non-4D) MAP-reconstructed images.

This algorithm also achieved useful reductions in average noise levels for 2C3K parameter values throughout the simplified phantom compared to conventional MAP reconstruction, although not within the hypoxic tumour region. 4D reconstruction based on the spline-residue and spectral models achieved the greatest reductions in noise, but outside of the hypoxic tumour region the bias in fitted K_1 and k_2 values was larger for these algorithms than for conventional MAP reconstruction, particularly so for the spline-residue model.

5.6 Discussion

A spline-residue model was proposed for use in 4D-PET reconstruction. This was compared to 4D-PET reconstruction based on adaptive-knot cubic B-splines, the

spectral model and the non-linear 2C3K model. Results were obtained for a 'realistic phantom' (the FMISO NSCLC dPET phantom used in chapters 3 and 4) in which noise-free ground-truth TACs were represented with statistically acceptable fits of cubic splines and a 3C5K compartment model to TAC data obtained from a patient with stage IV NSCLC. Further results were obtained for a 'simplified phantom' in which ground-truth TACs were represented by fits of the simple 2C3K compartment model, which did not describe the patient FMISO TACs well (see section 3.4.2.1).

For the realistic phantom, 4D reconstruction based on spline-residues generated less bias or noise in parameter maps of fitted kinetic values than did any of the other algorithms studied, in 5/8 of the combinations of bias/noise and kinetic parameters analysed here. Additionally, the spline-residue algorithm reduced bias or noise by over 50% compared to conventional (non-4D) MAP reconstruction in 5/8 combinations, and notably increased bias or noise in none. 4D reconstructions based on the 2C3K model generated the most biased kinetic parameters, and also generated higher levels of noise than did spline-residue-based reconstruction.

If anything, this analysis should favour 4D reconstruction based on the simple 2C3K compartment model. This is the model most commonly used in the literature to characterise FMISO kinetics [180, 103], and consequently the one that was fitted voxel-by-voxel to TACs obtained from reconstructed images when characterising the accuracy and precision of the tracer kinetics in the images. Despite this, the parametric map results obtained from 4D reconstructions based on the 2C3K model were poorer than those from spline-residue- and spectral-based 4D reconstructions. Results obtained from cubic spline-based 4D reconstructions were also generally worse than those from the spectral and spline-residue-based reconstructions.

The results presented in section 3.4.2.1 indicate that the 3C5K compartment model gives a better description of the patient TACs used to create the realistic phantom than the 2C3K model. Although the 3C5K model is not commonly used to analyse FMISO dPET data, and to our knowledge 3C5K-based 4D-PET reconstruction has not been proposed before, it is possible that for the realistic

phantom a 3C5K-based reconstruction would give better results than the 2C3K-based one. To test this, 30 iterations of the nested-MAP2 algorithm in which the 3C5K model was fitted at each iteration using the Levenberg-Marquardt algorithm, were run on the realistic phantom data. Figure 5.14 shows the image quality results in Figure 5.4 along with an additional bias-noise curve for the 3C5K-based reconstruction. In the early time-frames the 3C5K reconstruction marginally reduced image bias compared to the spline-residue approach but considerably increased noise-levels, while over the whole imaged time-course the 3C5K-based approach generated more bias and considerably more noise than the spline-residue model. Because of the poor performance of the 3C5K-based reconstruction, it is not considered further in this work.

For the simplified phantom 4D reconstruction based on the 2C3K model performed much better, achieving lower levels of bias in fitted kinetic parameter maps than any of the other reconstruction algorithms studied, and useful noise reductions compared to the conventional MAP algorithm. Thus, the 2C3K-based 4D algorithm might be expected to provide good results in situations where the radiotracer kinetics are accurately described by the 2C3K model throughout the imaged FOV.

The results obtained for parametric maps largely concord with those of the more direct analysis of the accuracy and precision of reconstructed images (Figures 5.4 and 5.5). For the realistic phantom, 4D reconstructions based on the spline-residue and spectral models produced higher quality images than reconstructions based on the 2C3K model or cubic splines. Across all time-frames and phantom regions, the quality of images reconstructed using the spectral model-based 4D algorithm was slightly better than that of images reconstructed using the spline-residue-based algorithm. However the spine-residue-based algorithm generated much higher quality images at early time-points, perhaps because it comprises far fewer basis functions and is therefore less likely to overfit data; and presumably this gain at early time-points led to the K_1 and k_2 parametric maps generated from spline-residue-based reconstructions being of a higher quality overall than those obtained from spectral-based reconstructions.

Images produced by linear model-based 4D reconstruction algorithms were much noisier when temporal roughness was not penalised (see Figure 5.3). For the phantoms studied here, linear model-based 4D reconstructions produced less biased results than 2C3K-based 4D reconstructions at comparable or lower noise levels. However, it should be noted that this advantage only became apparent when temporal roughness penalties were built into the linear model-based reconstructions. Most linear model-based 4D-PET algorithms described in the literature do not

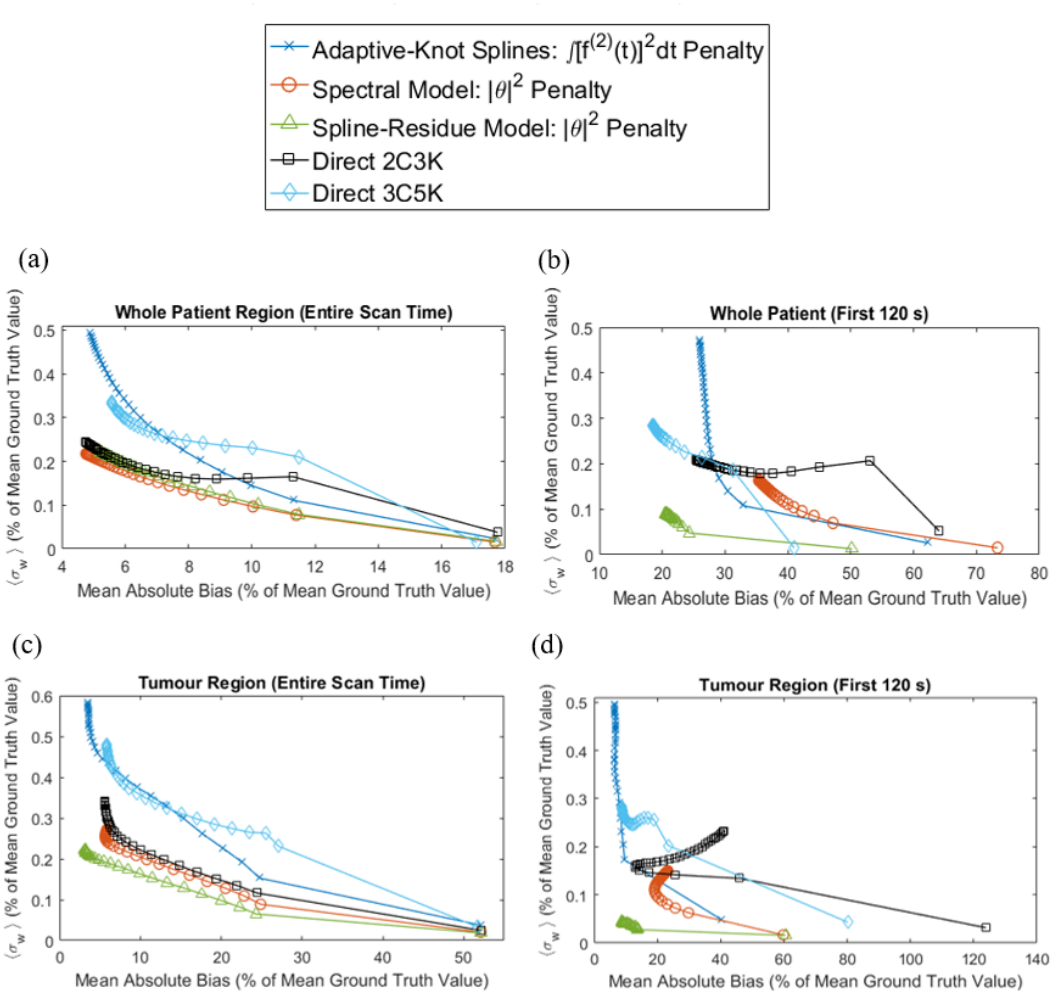


Figure 5.14. Comparison of image quality metrics for 4D reconstructions of the realistic phantom based on linear models (with their optimal temporal roughness penalty) and on the 2C3K and 3C5K models. The noise (weighted standard deviation, σ_w) and bias metrics are averaged over the whole patient (a, b) and tumour (c, d) phantom regions, and over the entire scan time (a, c) and first 120 s (b, d). The γ'_j hyperparameters were free to vary during each of the linear model-based reconstructions shown.

incorporate such penalties, possibly because this makes the algorithms more complex. However, this may put them at a disadvantage when compared to 4D algorithms based on less highly parametrised non-linear kinetic models.

Algorithm convergence is not guaranteed when model fitting step (5.14) is instead accomplished via the penalized WLS approach taken here. However, the results in Figure 5.6 show that the nested-MAP2 reconstructions did approach convergence within 30 reconstruction iterations, consistent with observations of convergence reported by Matthews *et al.* [99] for 4D reconstruction algorithms using WLS. Also, allowing γ'_j to vary during the reconstruction didn't adversely affect the convergence (see Figure 5.7).

L_1 regularization ($\Lambda(\boldsymbol{\theta}) = |\boldsymbol{\theta}|$) has been proposed for post-reconstruction fitting of the spectral model [54], with the aim of producing sparse solutions. This is useful if one intends to calculate physiologically relevant kinetic parameters directly from the $\boldsymbol{\theta}$ values. However, here linear models are used simply to produce smooth descriptions of the reconstructed TACs during 4D reconstruction, the resulting reconstructed images subsequently being analysed with a compartment model. Consequently, the priority here was to use regularization to limit overfitting, and thus L_2 regularization ($\Lambda(\boldsymbol{\theta}) = |\boldsymbol{\theta}|^2$) was explored, as it is specifically designed for this purpose, and is more computationally efficient than L_1 regularization.

The focus of this study was a phantom with kinetics taken from an FMISO dPET scan of an NSCLC patient, FMISO being of interest in oncology as a tracer of tumour hypoxia [180, 21, 103]. However most dPET studies use other tracers, particularly FDG. Examples of patient tissue TACs that were poorly described by non-linear kinetic models have been reported in the literature for tracers other than FMISO [41, 101, 111], and so 4D-PET reconstruction based on the more flexible linear spline-residue and spectral models may be useful for a wider range of tracers than FMISO alone.

Finally, O'Sullivan *et al.* [111] have shown that fits of a residue model based on splines provided better descriptions of TACs taken from previously reconstructed FDG-dPET brain image sequences than did fits of the 2C3K model, allowing

more robust estimation of kinetic parameters such as K_1, k_{flux}, median radiotracer transit time and fractional blood volume. Thus spline-residue-based 4D-PET reconstruction may prove useful for direct determination of some physiologically relevant kinetic parameters, as well as being a means to limit noise amplification during reconstruction before kinetic analysis of the resulting smoother images using a compartment model. The ability of the spline-residue model to produce kinetic parameters that give good descriptions of the patient biology is investigated in the next chapter.

5.7 Conclusion

Of all the reconstruction algorithms studied in this work, the proposed spline-residue-based 4D-PET reconstruction algorithm overall produced the highest quality (least biased or noisy) parametric maps of FMISO uptake kinetics for a phantom with a thoracic geometry and realistic FMISO uptake kinetics derived from those in a dynamic scan of an NSCLC patient. Specifically, reconstruction of images using this algorithm rather than the conventional (non-4D) MAP algorithm led to reductions of 50% or more in bias or noise for a majority of the combinations of kinetics parameters and bias/noise measures that were analysed, and did not generate notably worse results for any combination. 4D reconstruction based on the simple irreversible two-tissue compartment model 2C3K produced the most biased parametric maps for this phantom overall, and generated notably higher levels of bias and noise in the K_1 kinetics parameter than those obtained from conventional image reconstruction.

For a simplified phantom in which all ground-truth kinetics followed the poor descriptions of real data provided by fits of the 2C3K model, 4D reconstruction based on this model achieved the lowest levels of bias in fitted kinetic parameter maps. Thus, the 2C3K-based 4D algorithm can provide good results when the kinetics of radiotracer uptake throughout the FOV are described well by the 2C3K model.

When the underlying kinetics of tracer uptake in all imaged tissues are not well described by a simple compartment model, as for the FMISO kinetics studied in this work, it is advantageous to perform 4D-PET reconstruction using more

flexible linear kinetic models, the spline-residue model proving the best of the models studied here. Temporal roughness penalties improve the performance of 4D-PET reconstruction algorithms based on linear kinetic models, the optimal penalty function depending on the linear model being used.

Chapter 6

Radiogenomic Comparison of the Compartment and Spline-Residue Models

6.1 Introduction

In chapter 5 a spline-residue model was proposed for use in 4D-PET reconstruction. This model was found to provide a good description of all the simulated TACs present in the two FMISO dPET phantoms used in chapter 5. The spline-residue model was proposed in that work because of its potential to combine the superior noise suppressing capabilities of the spectral model in early time-frames with the ability of splines to accurately describe the entire imaged time-course. Images reconstructed with spline-residue based 4D-PET reconsruction were then analysed post-reconstruction using a compartment model.

The spline-residue model falls under a wider class of 'nonparametric' kinetic models. Recall from section 5.2 that a general TAC can be described as a convolution of the AIF with a residue function, where the residue function describes the fraction of tracer remaining in the region at a given time, and so provides information about the kinetics of the radiotracer transport and metabolism processes [55]. Nonparametric kinetic models deconvolve the AIF from the measured TAC, thus giving an estimate of the tissue residue function. O'Sullivan *et al.* [111] have shown that estimates of kinetic parameters such as the flux and flow rate constants can be derived from an estimate of the residue function, meaning that nonparametric models

have the potential to provide physiologically relevant kinetic parameters. This suggests that, rather than analysing images produced by the spline-residue model-based 4D-PET reconstruction algorithm proposed in chapter 5 with a compartment model, it may be possible to obtain clinically useful parameters directly from the spline-residue model fits carried out during the reconstruction.

A number of studies have shown that, due to their inherent flexibility, nonparametric models provide improved descriptions of TACs compared to more commonly used compartment models (which assume that the residue function consists of a weighted sum of exponential decays) [41, 59, 71, 110, 111]. The two most common approaches to nonparametric kinetic modelling are to represent the residue function with spline [111] or piecewise constant functions [41], though since it is reasonable to assume that real residue functions are smooth, the former should give a better representation. To our knowledge, no studies have explored whether the kinetic parameter estimates obtained from nonparametric models provide an improved description of the underlying patient biology compared to those derived from compartment model fits.

In this chapter a clinical dataset comprising 41 primary breast cancer patients is analysed. These patients underwent FDG-dPET scans and tumour biopsies before and after the administration of the drug Metformin. Fits of the spline-residue model are compared to those of irreversible two- and three-tissue compartment models using goodness-of-fit measures. To investigate which model best reflects the underlying biology, the kinetic parameters obtained from the model fits are correlated with gene expression data measured from the patient biopsies.

6.2 Deriving Kinetic Parameters from Fits of the Spline-Residue Model

Recall from section 5.2 that a TAC $f(t)$ can be modelled as the convolution of the AIF with a residue function,

$$f(t) = K \int_0^t C_p(s) R(t-s) ds, \tag{6.1}$$

where $C_p(t)$ is the AIF, K is a proportionality constant interpreted as overall flow and $R(t)$ is the residue function. Because the residue function describes the fraction of tracer remaining in the region at time t after entering it, it is a normalized impulse response function. The spline-residue model describes the residue function as a weighted sum of N_S cubic B-spline functions, $\zeta_k(t)$, with a Dirac delta function added to account for the finite blood volume in the region:

$$R(t) = \mu_0\delta(t) + \sum_{k=1}^{N_s} \mu_k\zeta_k(t), \tag{6.2}$$

where μ_k is the coefficient of the k^{th} B-spline function $\zeta_k(t)$, the coefficient μ_0 equals V_b/K and V_b is the fractional blood volume. The TAC is therefore represented as a weighted sum of spline-residue basis functions $\rho_l(t)$:

$$f(t, \boldsymbol{\theta}) = \sum_{l=0}^{N_S} \theta_l\rho_l(t), \tag{6.3}$$

where

$$\rho_l(t) = \begin{cases} C_p(t) & l = 0 \\ C_p(t) \otimes \zeta_l(t) & l > 0, \end{cases} \tag{6.4}$$

$\otimes$ denotes a convolution, and the coefficients θ_l are given by

$$\theta_l = \begin{cases} V_b & l = 0 \\ K\mu_l & l > 0. \end{cases} \tag{6.5}$$

Replacing the μ_k coefficients in equation (6.2) with the fitted θ_k coefficients gives an estimate of the impulse response function (IRF). The IRF is equal to the normalized residue function $R(t)$ multiplied by K. $R(t)$ is therefore obtained by normalizing the IRF, and the normalization constant used to do so is equal to K.

The values of V_b and K can therefore be derived directly from fits of the spline-residue model. Additionally, O'Sullivan *et al.* [111] showed that the flux constant k_{flux} is given by $KR(T_{scan})$, where T_{scan} is the end-time of the scan. Table 6.1 shows how the spline-residue kinetic parameters used in this chapter can also be obtained from fits of the 2C3K and 3C5K kinetic models, which are shown in Figure 6.1 along with their associated rate constants. Because the spline-residue flow constant K is equivalent to the compartment model constant K_1, for the remainder of this chapter all flow constants are represented as K_1.

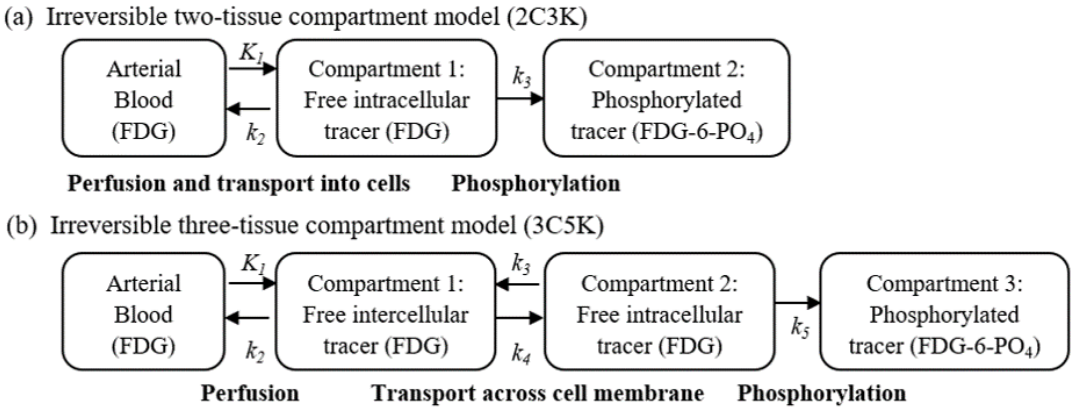


Figure 6.1. Schematic diagrams showing irreversible (a) two- and (b) three-tissue compartment models of FDG uptake. Flows between compartments are defined by rate-constants (k values) and compartment tracer concentrations.

Table 6.1. Illustration of how the spline-residue kinetic parameters can be obtained from fits of the 2C3K and 3C5K compartment models. Note that for the 2C3K and 3C5K models the asymptotic (i.e. $T_{scan} \to \infty$, where T_{scan} is the duration of the scan) expressions for the flux constant k_{flux} are shown.

Model	Fractional Blood Volume	Flow	Flux
Spline-residue	V_b	K	$KR(T_{scan})$
2C3K	V_b	K_1	$\frac{K_1 k_3}{k_2 + k_3}$
3C5K	V_b	K_1	$\frac{K_1 k_3 k_5}{k_2 k_5 + k_3 k_5 + k_2 k_4}$

6.3 Methods

6.3.1 Patient Data

A dataset comprising 41 primary breast cancer patients was analysed in this work. The patients underwent FDG dPET scans and tumour biopsies before and after the administration of the drug Metformin. The original purpose of the study for which these patients were recruited was to investigate how the Metformin drug affects tumour metabolism. Details of that study are available in [94]. In this work the data from this study is re-analysed to compare the performance of the spline-residue and compartment models.

Metformin is primarily used as a treatment for type 2 diabetes, though it has

been found that tumour cell proliferation is suppressed in cancer patients being treated with Metformin [186]. One potential reason for this is through the effects of Metformin on the mitochondrial metabolism. In the original study for which the data used in this chapter was collected for, Lord *et al.* [94] reported that Metformin inhibits oxidative phosphorylation, thus increasing the dependence on glycolysis. Compared to oxidative phosphorylation, glycolysis is a less efficient mode of glucose utilization, meaning that increases in glycolytic metabolism are likely to correspond to increased FDG uptake.

Patients were recruited from the medical oncology breast cancer clinic over a period of 30 months between May 2011 and November 2013 in three UK centres, Oxford, Luton and Dundee. All patients at the point of recruitment had been referred with a view to neoadjuvant chemotherapy, had histologically confirmed breast cancer and gave informed consent. In all cases the primary breast cancer was *in situ* and no patients had received any prior treatment for breast cancer.

Metformin was given in the Glucophage XR^{TM} formulation in an escalating dose once daily for a minimum of 13 days and a maximum of 21 days (500 mg for days 1–3, 1000 mg for days 4–6 and 1500 mg thereafter). For each patient, pre- and post-treatment biopsy samples were taken under ultrasound guidance from the periphery of the primary tumour. Within 1 minute of this procedure the biopsy material was snap frozen in liquid nitrogen prior to storage at -80°C.

Table 6.2 summarizes the inclusion and exclusion criteria used when recruiting the patients in this dataset. Table 6.3 summarizes the main patient and tumour characteristics, as well as those of the analysed patient samples.

6.3.2 PET-CT Protocol

A dynamic acquisition of the breast tumour (and any lymph nodes within the PET field of view) was initiated with the patient imaged supine. Prior to scanning, patients fasted overnight for at least 8 hours but could drink water. Patients' blood glucose was checked just prior to the scan with a portable blood glucose monitor to ensure it was < 7 mmol/L. Patients were injected with FDG (3 MBq/kg, up to

Table 6.2. List of key inclusion and exclusion criteria used when recruiting the patients in this dataset.

Inclusion Criteria	Exclusion Criteria
Women with a histology proven *in situ* primary breast cancer ≥ 2 cm in diameter.	Radiotherapy, major surgery, significant traumatic injury, endocrine therapy, immunotherapy, chemotherapy or experimental therapy during four weeks prior to starting or during trial.
Eastern Cooperative Oncology Group (ECOG) performance status 0–1.	Pregnancy or breast feeding.
Age ≥ 18 years	History of type 1 or type 2 diabetes.
Fasting or random serum glucose less than 7.0 mmol/L.	Treatment with Metformin in the past year.
No prior treatment for breast cancer and scheduled to commence neoadjuvant chemotherapy in < 3 weeks time.	Estimated glomerular filtration rate < 45 mL/min.
Have given written informed consent and are capable of cooperating with protocol.	Acute or chronic metabolic acidosis.
Adequate bone marrow, renal and liver function.	Known hypersensitivity to Metformin.

a maximum of 400 MBq) 30 seconds into PET imaging, which continued for 45 minutes. The 45 minutes of data were then binned into the following time-frame sequence: [1×30 s, 12×5 s, 6×10 s, 5×30 s, 10×60 s, 6×300 s]. 50 minutes after injection, a 30 minute static PET scan was performed from skull base to mid-thigh, acquiring data for four minutes at each bed position. Thus the primary breast cancer was scanned at approximately 60 minutes post injection.

All PET scans were carried out in 3D-mode, using either a GE Discovery 690 or a Siemens Biograph mCT-128 PET-CT system. The PET images were reconstructed onto a grid of 5.5×5.5×3.3 mm^3 voxels, using FORE+FBP for the dPET time-frames and 2 iterations of a TOF OSEM algorithm (with 24 subsets) for the static images. 6.4 mm spatial Gaussian smoothing was applied to each reconstructed image. Prior to each PET acquisition a CT scan was performed for localization and PET attenuation correction.

Static PET images of one of the patients before and after the administration of

Table 6.3. Patient and tumour characteristics, as well as a summary of the patient data acquired. The patient estrogen receptor (ER) and human epidermal growth factor receptor-2 (HER2) statuses is also shown. Note that here 'paired' means that data was acquired both before and after the patient was administered with Metformin.

	Number of Patients
Recruitment and Samples analysed	
Total patient recruitment	41
Number of paired PET-CT scans available for analysis.	36
Number of paired tumour samples with available mRNA sequencing data.	36
ER/HER2 status	
ER positive	32
ER negative	9
HER2 positive	8
HER2 negative	33
Triple negative (ER negative and HER2 negative)	8
Tumour type	
Ductal carcinoma	32
Lobular carcinoma	7
Mixed ductal and lobular carcinoma	2
Grade 1	2
Grade 2	24
Grade 3	15
Patient characteristics	
Median tumour size (on MRI scan)	49 mm (range 30-147 mm)
Median age at study entry	49 years (range 27-67 years)
Median body mass index	28.1 (range 19.6-45.3)

Metformin are shown in Figure 6.2. These images show that Metformin noticeably increased the overall FDG uptake levels. While the scales make it hard to see the difference in uptake in the primary tumour, the post-Metformin FDG uptake levels of the axillary lymph nodes (circled in Figure 6.2 (b) are noticeably higher.

6.3.3 Analysis of Data from Tumour ROIs

Tumour ROIs were delineated on the 60-minute static PET scans by a nuclear medicine radiologist. These tumour ROIs were transferred to the corresponding

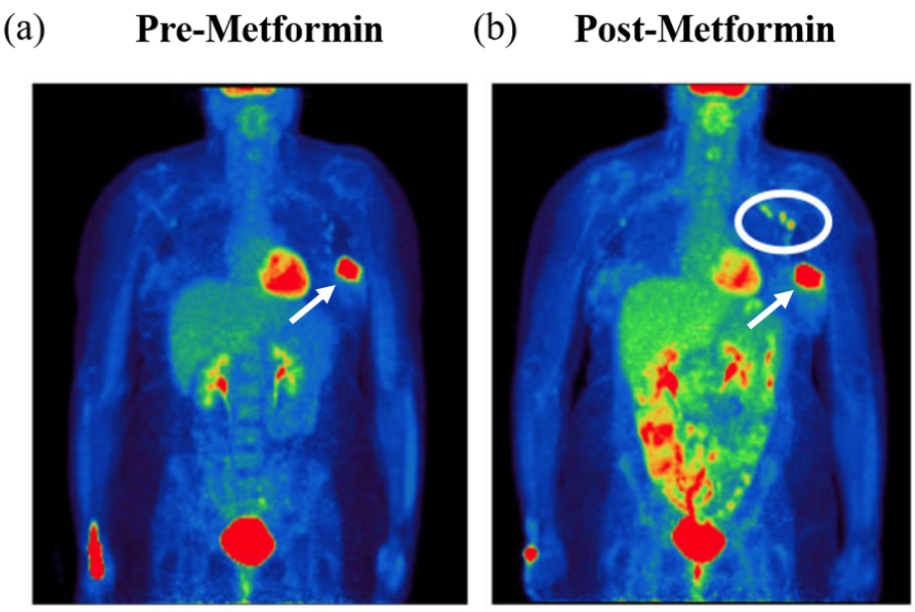


Figure 6.2. Coronal maximum intensity projections of static SUV images of a patient before (a) and after (b) the administration of the drug Metformin. Note that the scale in both images is the same. The white arrow in both images points to the primary breast cancer tumour. The post-Metformin FDG uptake levels of the axillary lymph nodes (circled in (b)) are notably higher.

dPET scan by co-registering the two image sets. Arterial blood TACs were obtained from ROIs defined in the descending aorta (average 42 ± 4 slices with mean volume of 32 ± 11 cm^3).

Tumour TACs were available from 35/36 of the patients, as the TACs from the remaining patient showed pronounced discontinuities, likely due to movement during scanning. For the remaining patients the pre- and post-Metformin tumour TACs were available, giving a total of 70 TACs to analyse. The tumour TACs were fitted with the 2C3K, 3C5K and spline-residue models. The three-exponential function of Feng *et al.* [40] was used to fit the arterial blood TACs to get the AIFs.

The three-exponential Feng AIF model is widely used in dPET studies, and is given by the following equation:

$$C_p(t) = \begin{cases} 0 & t < \tau \\ (A_1(t-\tau) - A_2 - A_3)e^{(-\lambda_1(t-\tau))} + A_2 e^{(-\lambda_2(t-\tau))} + A_3 e^{(-\lambda_3(t-\tau))} & t \geq \tau, \end{cases} \tag{6.6}$$

where A_i and λ_i $(i = 1 - 3)$ are free parameters that define the shape of the fitted

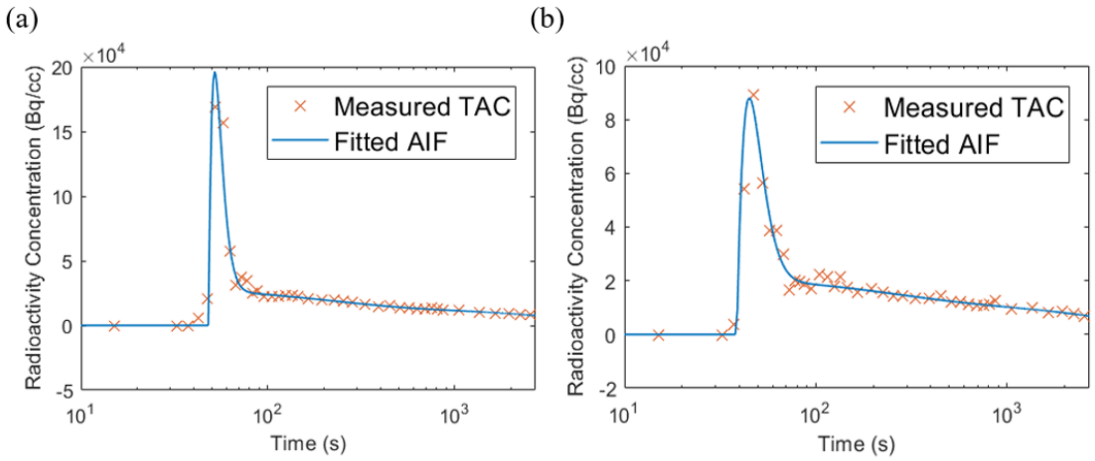


Figure 6.3. Examples of fits of the three-exponential Feng AIF model given in equation (6.6) to measured arterial blood TACs from two separate patients.

TAC and τ is a free parameter that accounts for the time delay between the start of the scan and the initial rise time of the TAC. Figure 6.3 shows two examples of arterial blood TACs being fitted with the Feng model.

WLS was used for all model fitting, using the dPET weighting factors

$$w_m = \frac{\Delta T_m e^{-\lambda t_m}}{a_m}, \tag{6.7}$$

where a_m is the decay-corrected average activity concentration within the ROI during frame m, ΔT_m is the frame duration, t_m is the mid-point of the frame and λ is the radiotracer decay constant. The Feng three-exponential AIF, 2C3K and 3C5K models were fitted using the Levenberg-Marquardt algorithm available in the MATLAB optimization toolbox.

Model fit quality was assessed for each TAC using leave-one-out cross-validation, calculating the weighted RSS error of the fit with the weighting factors defined in (6.7). The Wald-Wolfowitz runs test was used to check whether any significant structure remained in the residuals of each model fit at the 5% significance level.

The spline-residue model was fitted using L_2-regularized WLS. The same range of L_2-regularization penalty coefficients γ that were used in chapter 5 were also used here, with the optimal γ being selected using GCV. 12 fixed knots (4 at the beginning, end and initial rise time of each TAC) were used for each spline-residue fit.

For the spline-residue model fits, the initial rise time of each TAC was determined by performing multiple fits with the initial rise time varied in one second intervals from 30 to 60 s into the scan and selecting the fit producing the lowest WLS error. To determine the optimal number of 'free' knots, which here were placed uniformly between the initial rise time and end point of the TAC, the spline-residue model was fitted to a subset of 10 patient TACs with the number of free knots varied from 1 to 10, and the weighted RSS error of each fit was calculated using leave-one-out cross-validation. The number of free knots minimizing the weighted RSS error for each TAC was taken to be optimal.

The mean standardised uptake value (SUV_{mean}) within the tumour ROI was calculated for each of the static PET images. SUV_{mean} is defined as

$$SUV_{mean} = a_{mean}\frac{BW}{ID}, \tag{6.8}$$

where a_{mean} is the mean measured radioactivity concentration in the ROI, BW is the patient body weight and ID is the injected radioactivity dose. The mean tumour to blood ratio (TBR_{mean}) within the tumour ROI, defined as

$$TBR_{mean} = \frac{SUV_{mean,tumour}}{SUV_{mean,blood}}, \tag{6.9}$$

where $SUV_{mean,tumour}$ and $SUV_{mean,blood}$ are the mean SUVs in the tumour and arterial blood ROIs, respectively, was also calculated for each static PET image.

6.3.4 Genomic Data

The individual biochemical processes that occur within biological cells are regulated by genes. Gene expression is the process through which information from a gene is used in the synthesis of a functional gene product, such as a protein, which is in turn used in the relevant biochemical reaction. In gene expression, the segment of a deoxyribonucleic acid (DNA) molecule that defines a given gene is 'transcribed' into a corresponding messenger ribonucleic acid (mRNA) molecule. This mRNA molecule is exported from the cell nucleus into the cytoplasm, and there it is 'translated' into the corresponding protein or other functional gene

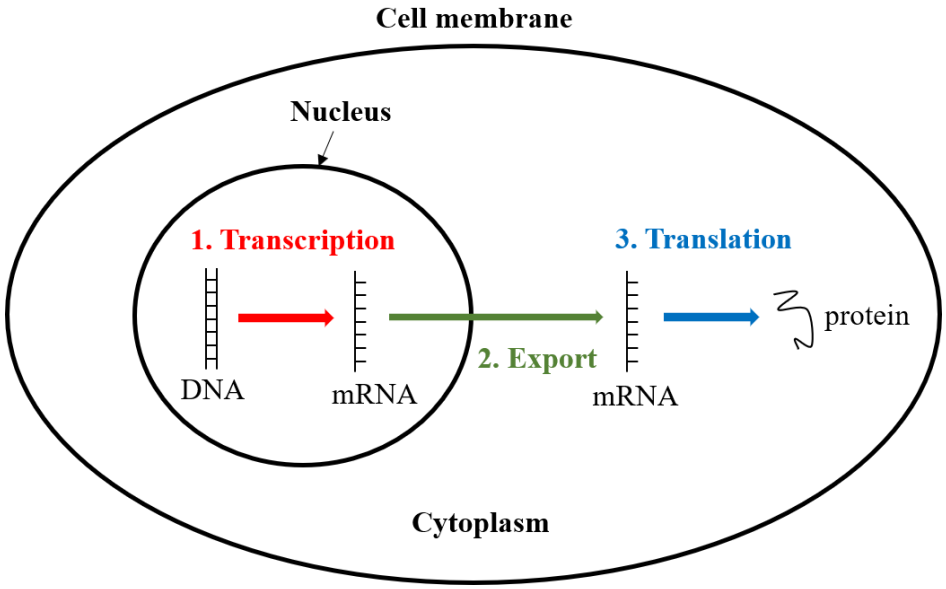


Figure 6.4. Schematic diagram summarizing the gene expression process. A more detailed description is given in [114].

product. This process is summarized in Figure 6.4, and a more detailed description of gene expression is given in [114].

The level of gene expression in a given sample of cells can be measured using RNA sequencing. Wang *et al.* [181] provide a detailed review of the most common methods used to perform RNA sequencing, the details of such methods being beyond the scope of this thesis. To measure the gene expression levels from the tumour biopsy samples, RNA sequencing was carried out by the Oxford Genomics Centre core facility at the Welcome Trust Centre for Human Genetics. Full details of how the RNA sequencing was carried out are available in [94].

The transport of FDG across the cell membrane is regulated by the glucose transporter (GLUT) genes, and the phosphorylation of FDG into FDG-6-PO_4 is regulated by the Hexokinase (HK) genes [47]. One might therefore expect an increase in the expression of these genes to correspond to higher levels of FDG uptake [147]. The baseline expression and pre-post Metformin expression changes of 11 glucose transporters (GLUT1, GLUT3, GLUT4, GLUT5, GLUT6, GLUT8, GLUT9, GLUT10, GLUT11, GLUT12, GLUT14) and three hexokinases (HK1, HK2, HK3) were available from the RNA sequencing data. The units of the measured

gene expression levels are fragments per kilobase million (FPKM).

Given a dataset containing the expression levels of many individual genes, measurements of certain biological features can be obtained using gene signature scores, or metagenes. Gene signature scores represent a given biological feature as a single number, and are calculated using pre-defined functions of the expressions of specific genes, of specific genes, the functions being found to be strongly correlated with the biological feature. Baseline and pre-post Metformin changes in hypoxia, angiogenesis and proliferation signature scores were available from this dataset. Details of the signatures used here are available in [13] (hypoxia signature), [97] (angiogenesis signature) and [35] (proliferation signature). Note that the hypoxia signature contains both GLUT1 and HK2, and the proliferation signature contains GLUT1, meaning that these signatures are not entirely independent of the individual expressions of these genes. These signature scores were calculated by Dr. Wei-Chen Cheng.

6.3.5 Correlating Imaging and Genomic Parameters

Of the 41 patients in the dataset, both pre- and post-Metformin dPET and gene expression data was available for 31.

The baseline and pre-post changes in the kinetic parameters from each model studied here were respectively correlated with the baseline and pre-post changes in GLUT1 and HK2 expression, as these are known to be particularly important for FDG uptake [47]. Correlations between the kinetic parameters and the baseline and pre-post changes in the hypoxia, angiogenesis and proliferation signature scores were also measured. The correlations were performed using the Pearson correlation coefficient. The same analysis was also performed for the SUV_{mean} and TBR_{mean} parameters derived from the static PET images.

To test the extent to which the combination of all available glucose transporters and that of all available hexokinases affect the kinetic parameters, two linear multiple regression models, one representing a given kinetic parameter as a weighted sum of all 11 available glucose transporters and the other representing it as a weighted

sum of the 3 available hexokinases, were fitted to each kinetic parameter. This was done for both the baseline and pre-post change data. For each fit, the sum of square deviations from the mean for kinetic parameter q was calculated:

$$SS_q = \sum_{i=1}^{N_p} (q_i - \bar{q})^2, \tag{6.10}$$

where N_p is the number of patients for which both dPET and gene expression data were available (31) and $\bar{q}$ is the mean value of kinetic parameter q. The sum of square residuals for each regression fit were also calculated:

$$SSR_q = \sum_{i=1}^{N_p} (q_i - \hat{q})^2, \tag{6.11}$$

where $\hat{q}$ is the value of q predicted by the regression fit. The $\mathcal{F}_q$-statistic,

$$\mathcal{F}_q = \frac{\frac{SS_q - SSR_q}{N_{gene}}}{\frac{SSR_q}{N_p - N_{gene} - 1}}, \tag{6.12}$$

where N_{gene} is the number of genes in the model, was calculated for each of the regression fits. Under the null hypothesis that the regression model does not provide a significantly better prediction of the kinetic parameters than when all the weightings in the regression model are set to zero, $\mathcal{F}_q$ should be F-distributed with degrees of freedom (DOF) $DOF_1 = N_{gene}$ and $DOF_2 = N_p - N_{gene} - 1$. The F-test, which calculates the probability of getting $\mathcal{F}_q$ under this null hypothesis, was used to test the dependence of the kinetic parameters on the expression of all available glucose transporters and hexokinases.

Figure 6.5 summarizes the available imaging and genomic measures, and how they were compared to each other. Overall, 17 PET parameters are studied alongside 7 sets of genomic data (GLUT1 and HK2 expressions, the three sets of signature scores, all available GLUT expressions and all available HK expressions). The total number of hypotheses tested is therefore $17 \times 7 \times 2 = 238$, with the extra factor of 2 being due to the fact that the analyses were done both at baseline and for the pre- to post-Metformin changes.

The Benjamini-Hochberg procedure [8] was used to control the false discovery rate (FDR) following the multiple hypothesis testing. In this methodology the

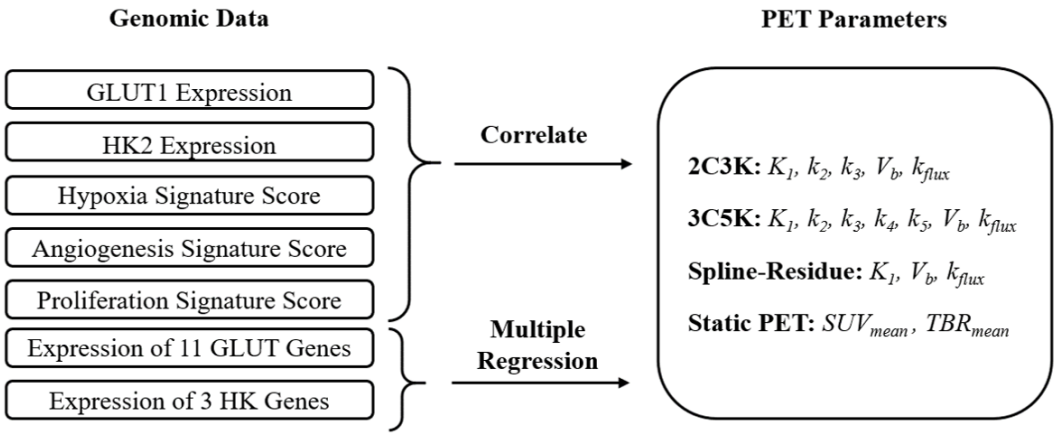


Figure 6.5. Schematic diagram summarizing the available Genomic and PET imaging parameters, and how they were compared to each other. Note that the analysis shown in this figure was carried out for both the baseline data and the pre- to post-Metformin changes in the data.

the p-values from each hypothesis are ranked in ascending order. The largest r value satisfying

$$p_r \leq \frac{r}{m}\alpha, \tag{6.13}$$

where p_r is the p-value of the r^{th} hypothesis in the ordered list, m is the total number of hypotheses (238 in this work) and α is a tunable parameter corresponding to the level to which it is desired to limit the FDR, is defined as $\hat{r}$. All hypotheses up to and including the $\hat{r}^{th}$ hypothesis in the ordered list are accepted, provided their p-values are ≤ 0.05, and the remainder are rejected. This is an exploratory study, and further studies with more patients are required to confirm the hypotheses that are generated in this work. The cost of a false negative (missing a potentially significant result) is therefore high, and so a fairly high FDR value of 0.2 was used here. This FDR value has been used in other published exploratory studies that used the Benjamini-Hochberg procedure to account for multiple hypothesis testing (e.g. [165]).

To test whether the correlation coefficients of different PET imaging parameters with a given genomic measure were significantly different from each other, an approach based on Fisher's Z-transform was used. For a given Pearson correlation

coefficient R, the Z-transform is defined as:

$$Z = \text{arctanh}(R), \tag{6.14}$$

where Z is the transformed value of R [42]. The motivation for the Z-transform is that the distribution of R can be skewed when it is calculated from a small sample, while even for small sample sizes Z is approximately Gaussian-distributed, with variance

$$\text{var}(Z) = \frac{1}{N-3}, \tag{6.15}$$

where N is the number of samples in the dataset used to calculate R [43]. When comparing the correlation coefficients R_1 and R_2 of two different imaging parameters with a given genomic measure, the z_c-statistic is defined as

$$z_c = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}}, \tag{6.16}$$

where Z_1 and Z_2 are the Z-transformed values of R_1 and R_2, and for this work $N_1 = N_2 = N$, as the correlation coefficients are all calculated from the same dataset. The p-value for testing whether R_1 and R_2 are significantly different is calculated using

$$p = 2P(z \geq |z_c|) = 2(1 - \text{cdf}_{norm}(|z_c|)), \tag{6.17}$$

where $\text{cdf}_{norm}(z)$ is the Gaussian cumulative distribution function.

6.4 Results

6.4.1 Assessment of Model Fit Quality

For the subset of 10 patient TACs for which the number of free knots were varied, the average weighted RSS error for fits of the spline-residue model to each TAC using a given number of free knots, calculated using leave-one-out cross-validation, are shown in Figure 6.6. This figure shows that the fitting error for this subset of patient TACs was minimized when 7 free knots were used. For the remainder of the work presented in this chapter all spline-residue fits were carried out with 7 free knots.

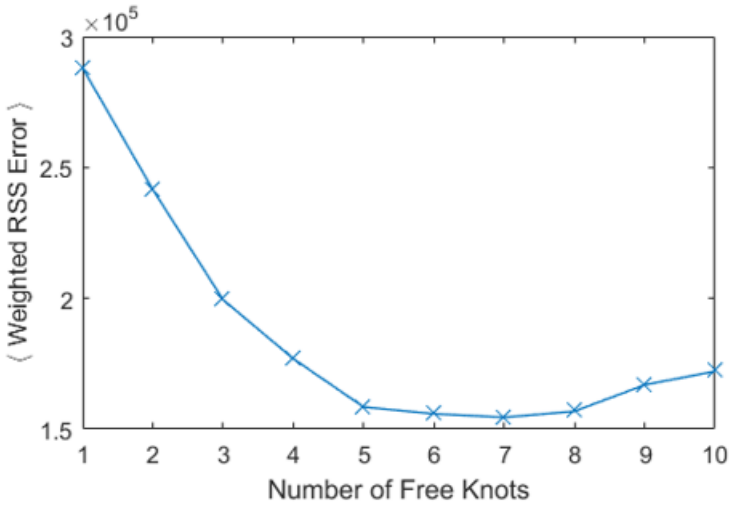


Figure 6.6. Mean weighted RSS error (calculated using leave-one-out cross-validation) for fits of the spline-residue model to a subset of 10 patient TACs as a function of the number of free knots used.

The goodness-of-fit measures for each of the kinetic models are summarized in Table 6.4. It can be seen from Table 6.4 that, of all the models tested, the spline-residue model produced the most fits that passed the runs test and also achieved the lowest median weighted RSS error. The fits of the 3C5K model had the highest median weighted RSS error overall. Figure 6.7 shows examples of fits of each of the kinetic models to two pre-drug tumour TACs, derived from separate patients, and the corresponding residue functions for each model fit. This figure shows that for some TACs each model gave quite similar results, while for other TACs the fits look somewhat different.

Figure 6.8 shows boxplots of the rate constants from each kinetic model fit along with the SUV_{mean} and TBR_{mean} parameters from the static PET images. To better

Table 6.4. Median weighted RSS errors for model fits to patients TACs, calculated using leave-one-out cross validation, and the number of TACs for which each model passes the Wald-Wolfowitz runs test at the 5% significance level. The best result for each metric is highlighted in bold.

Model	Number of fits that pass the runs test	Median weighted RSS error
Spline-residue	**65/70**	$\mathbf{1.45 \times 10^5}$
2C3K	56/70	2.07×10^5
3C5K	52/70	2.57×10^6

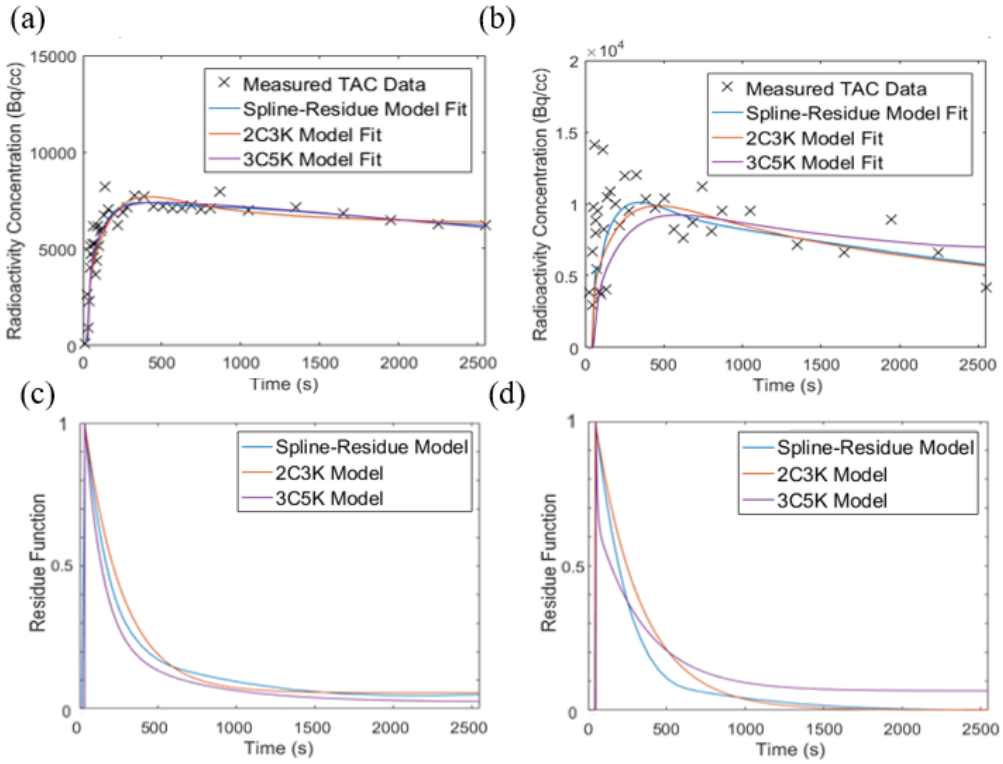


Figure 6.7. Examples of 2C3K, 3C5K and spline-residue model fits to pre-drug tumour TACs from two different patients. (c) and (d) show the corresponding residue functions from each model fit to the TACs shown in (a) and (b), respectively.

facilitate comparisons between the different models, the boxplots for the K_1 and k_{flux} parameters are re-plotted in Figure 6.9, with separate plots comparing the distributions of K_1 and k_{flux} produced by each kinetic model.

6.4.2 Correlations Between Imaging and Genomic Measures

Of the 238 hypotheses tested in this work, the ones that passed the Benjamini-Hochberg procedure to account for multiple hypothesis testing are shown in Table 6.5. All of the F-tests for the multiple regression analyses were filtered out by the Benjamini-Hochberg procedure. In general, stronger correlations between the pre- to post-Metformin changes in the imaging and genomic parameters than those between the baseline values were observed, which explains why none of the baseline correlations passed the Benjamini-Hochberg procedure. The results from each of the individual hypotheses that were tested in this work are shown in appendix A.

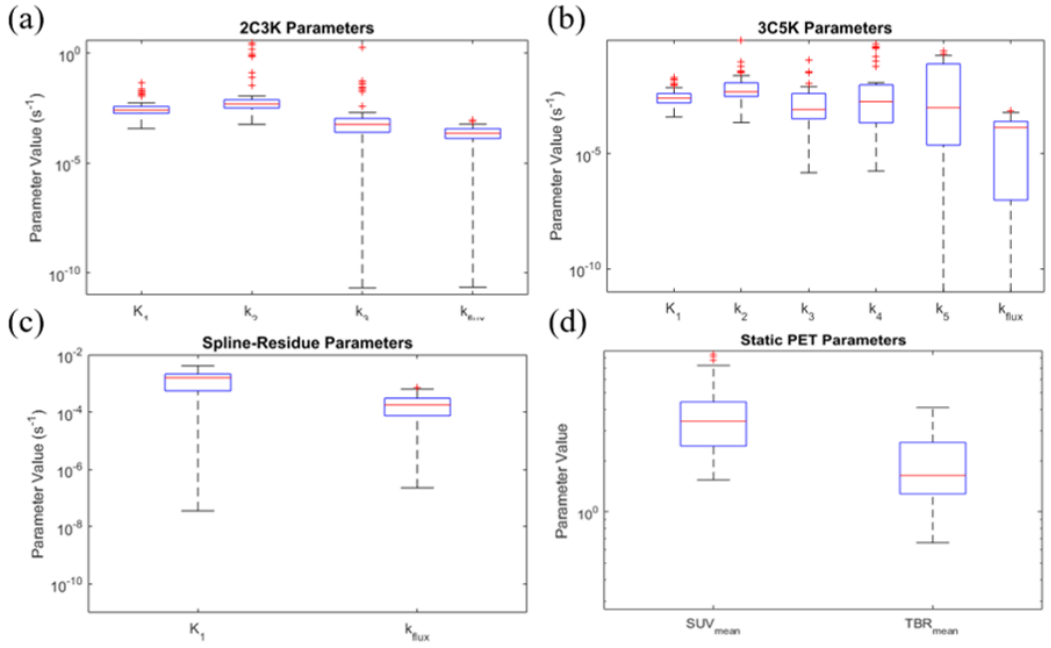


Figure 6.8. Boxplots of the rate and flux constants from the fitted 2C3K (a), 3C5K (b) and spline-residue model (c), as well as the parameters from the static PET images (d). Any data point that lies further than 1.5x the interquartile range from the first of third quartiles (the top and bottom of each of the boxes, respectively) are shown with red crosses. Note that a logarithmic scale is used to improve the visibility of both the boxes and the outliers.

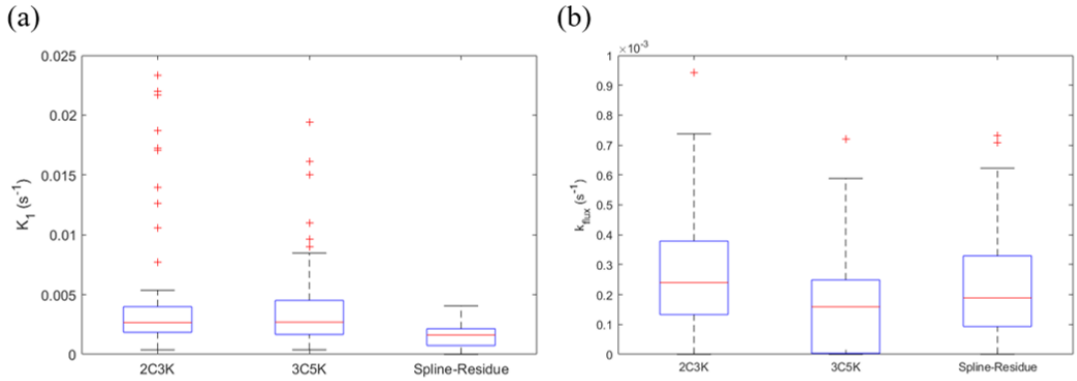


Figure 6.9. Boxplots of the K_1 (a) and the k_{flux} (b) parameters produced by each of the different model fits. Any data point that lies further than 1.5x the interquartile range from the first of third quartiles (the top and bottom of each of the boxes, respectively) are shown with red crosses.

In Figure 6.10 the correlations between the pre- to post-Metformin spline-residue k_{flux} changes and those of the GLUT1 and HK2 expressions, and the proliferation and hypoxia signature scores, are plotted. The pre- to post-Metformin changes in 3C5K k_2 and 2C3K k_3 correlated with the change in angiogenesis signature score,

Table 6.5. Statistically significant results remaining after the Benjamini-Hochberg procedure (with $\alpha = 0.2$) was used to account for multiple hypothesis testing. Note that all of the F-tests for the multiple regression analyses were filtered out the Benjamini-Hochberg procedure. Results are shown in order of ascending p-value.

Hypothesis	Pearson Correlation Coefficient
Pre-post change in proliferation score correlates with spline-residue k_{flux}	0.613 (p = 0.0003)
Pre-post change in GLUT1 expression correlates with spline-residue k_{flux}	0.608 (p = 0.0004)
Pre-post change in hypoxia score correlates with spline-residue k_{flux}	0.541 (p = 0.0020)
Pre-post change in angiogenesis score correlates with 3C5K k_2	-0.521 (p = 0.0031)
Pre-post change in angiogenesis score correlates with 2C3K k_3	-0.520 (p = 0.0032)
Pre-post change in HK2 expression correlates with spline-residue k_{flux}	0.519 (p = 0.0033)

and these correlations are plotted in Figure 6.11.

The circled points in Figure 6.11 (a) and (b) appear to be anomalous. These data points correspond compartment model fits to a pre-drug TAC shown in Figure 6.12. All three model fits to this TAC had the highest fitting errors (when calculated using leave-one-out cross validation) of all the TACs studied. The weighted RSS fitting errors $wRSS$ of each model fit to this TAC were calculated using

$$wRSS = \sum_{i=1}^{N_T} w_i (x_i^{(model)} - x_i^{(TAC)})^2, \tag{6.18}$$

where N_T is the number of time-frames, the w_i weighting factors are defined in (6.7) and $x_i^{(model)}$ and $x_i^{(TAC)}$ are the values of the fitted model and measured activity concentration in time-frame i, respectively. The $wRSS$ errors were also calculated for the TAC with the second highest model fitting errors (as calculated using leave-one-out cross validation). Both sets of $wRSS$ values are χ^2-distributed with $DOF = N_T - N_{param}$, where N_{param} is the number of model parameters, meaning that an F-test can be used to test whether the ratio of $wRSS$ errors for fits of a given model to both TACs is significantly different from 1. For fits of the spline-residue, 2C3K and 3C5K to the TAC in Figure 6.12 and the TAC with the

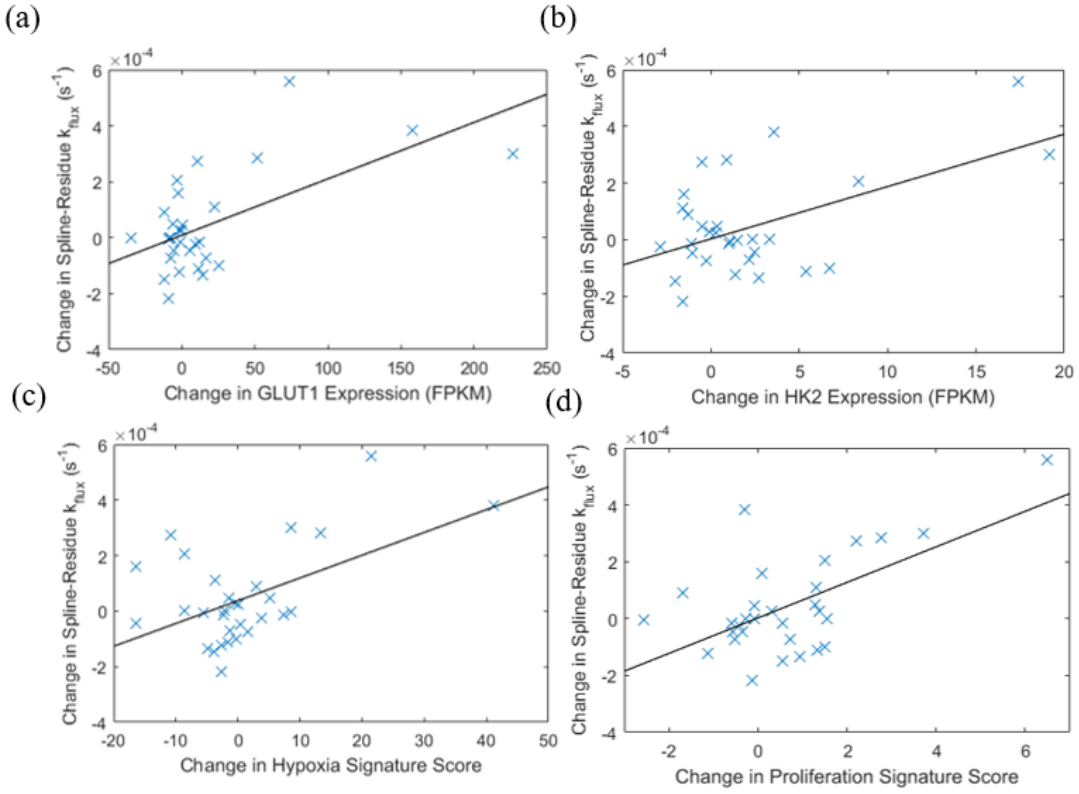


Figure 6.10. Plots of pre- to post-Metformin change in spline-residue k_{flux} versus changes in GLUT1 expression (a), HK2 expression (b), hypoxia signature score (c) and proliferation signature score (d). Linear fits to these plots are shown by the solid black line. All of the correlations shown here passed the Benjamini-Hochberg procedure, and are listed in Table 6.5. Note that the hypoxia and proliferation scores are dimensionless.

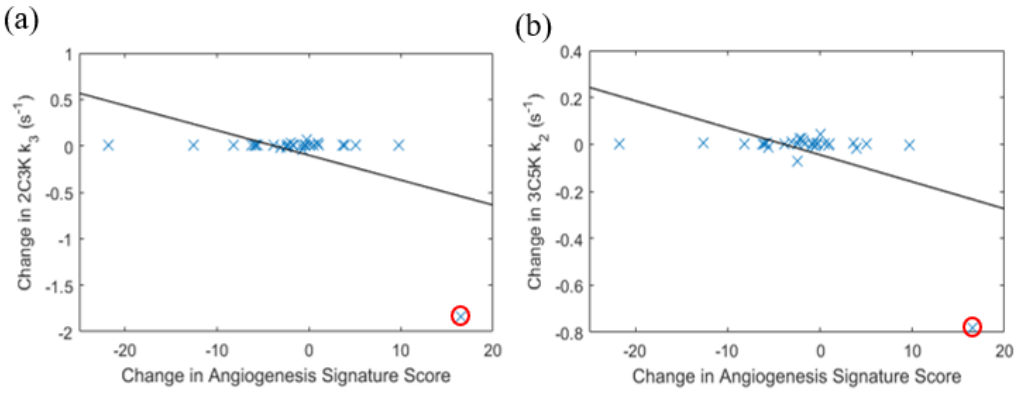


Figure 6.11. Plots of pre- to post-Metformin change in 2C3K k_3 (a) and 3C5K k_2 (b) versus changes in angiogenesis signature score. Linear fits to these plots are shown by the solid black line. One data point on both plots (circled) appears to be an outlier. Both of the correlations shown here passed the Benjamini-Hochberg procedure, and are listed in Table 6.5. Note that the angiogenesis scores are dimensionless.

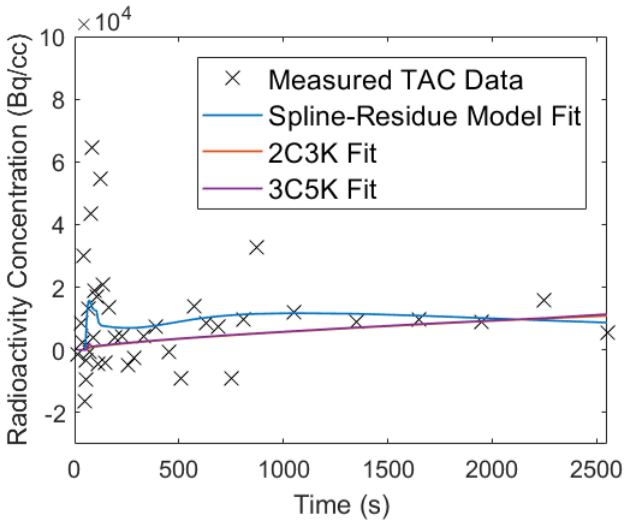


Figure 6.12. Fits of the 2C3K, 3C5K and spline-residue models to a particularly noisy TAC. The 2C3K and 3C5K fits to this TAC give the circled outlier in Figure 6.11

second highest fitting error gave F-test p-values of 1.35×10^{-7}, 1.33×10^{-6} and 5.02×10^{-8}, respectively, implying that the quality of the model fits to the TAC in Figure 6.12 are significantly different to those of the 'next-worst' TAC.

To check whether any of the points in Figures 6.10 and 6.11 are genuinely anomalous, leave-one-out cross-validation was used to calculate the residuals of a linear polynomial fit to each of the plots shown in these figures. The mean and standard deviation (σ) of the residuals for each plot were calculated, and if the residual of a given point lay further than 3σ from the mean residual of that plot, then it was deemed to be anomalous. Using this methodology, none of the points in any of the plots shown in Figure 6.10 were found to be anomalous. For the plots in Figure 6.11, the circled data points were excluded when calculating the mean and σ values of the residuals, as they were such extreme outliers. The residuals of the circled data points in Figure 6.11 (a) and (b) lie 9.47σ and 9.25σ from the mean residual, respectively, indicating that they are anomalous. None of the other points in Figure 6.11 were found to be anomalous.

Repeating the correlations of pre- to post-Metformin change in 2C3K k_3 and 3C5K k_2 with the change in angiogenesis score with the anomalous point removed

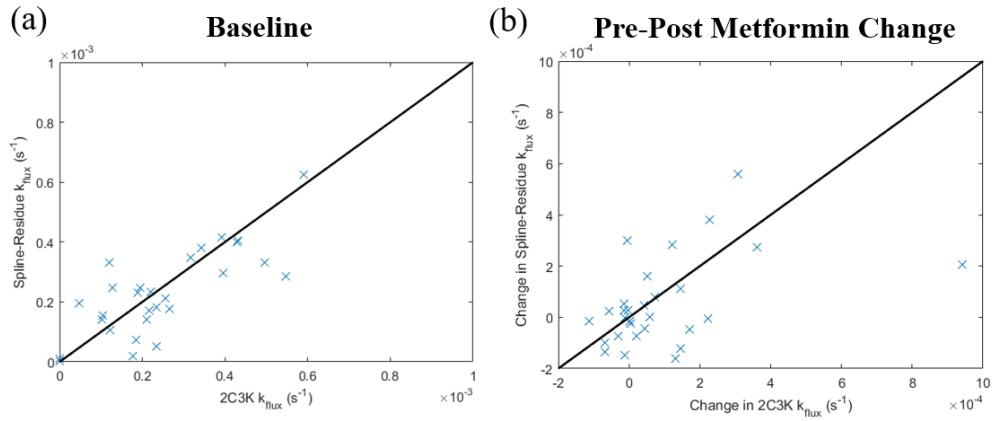


Figure 6.13. Plots of the k_{flux} parameter derived from spline- residue model fits versus the k_{flux} parameter from 2C3K fits for the baseline data (a) and the pre-post Metformin changes (b). The black line is a plot of y = x.

gives $R = 0.038$ (p = 0.846) and $R = 0.030$ (p = 0.876), respectively, indicating that without the anomalous data point these correlations cease to be significant.

The results presented here indicate that, of the PET imaging parameters compared, the k_{flux} parameter from fits of the spline-residue model tended to give the best correlations with the genomic measures. To investigate whether this parameter contains unique information not provided by any of the other parameters, the correlation coefficients of spline-residue k_{flux} with the genomic measures were compared to those of 2C3K k_{flux}, which of all the other kinetic parameters investigated produced the strongest correlations with spline-residue k_{flux}: $R = 0.806$ (p = 7.56×10^{-5}) and $R = 0.465$ (p = 0.0097) for the baseline and pre-post Metformin changes, respectively. These correlations are shown in Figure 6.13.

The results of the Z-transform-based approach for comparing the correlation coefficients between the spline-residue and 2C3K k_{flux} parameters with the genomic measures are shown in Table 6.6. These results show that at baseline there is no significant difference between the correlations of the spline-residue and 2C3K k_{flux} parameters with the genomic measures at baseline, while for the pre- to post-Metformin changes some of the correlation coefficients were significantly different. These results are consistent with the fact that the correlation between the two

Table 6.6. Results from Z-transform-based test for comparing whether the correlation coefficients between k_{flux} parameters derived from the spline-residue and 2C3K models and the genomic measures were significantly different from each other. Results for both the baseline and pre- to post-Metformin changes are shown. Statistically significant results ($p < 0.05$) are highlighted in bold.

Genomic Measure	Baseline	Pre-Post Metformin Change
GLUT1 expression	0.1430	**0.0197**
HK2 expression	0.2876	**0.0067**
Angiogenesis signature score	0.8877	0.5768
Hypoxia signature score	0.7928	0.1850
Proliferation score	0.8956	**0.0325**

sets of k_{flux} parameters was much stronger at baseline than it was for the pre- to post-Metformin changes.

6.5 Discussion

In this chapter a dataset comprising 41 primary breast cancer patients, who underwent static and dynamic FDG-PET scans as well as tumour biopsies before and after the administration of the drug Metformin, was analysed. The dPET data was fitted with the spline-residue, 2C3K and 3C5K models. Of all the models being compared, the spline-residue model generally provided the best fits to the dPET TACs, and the spline-residue k_{flux} parameter provided the best correlations with the patient genomic data after multiple hypothesis testing was accounted for. This implies that the ability of the spline-residue model to provide improved descriptions of the TACs led to the parameters from this model providing a better description of the underlying patient biology.

The correlations between the pre- to post-Metformin changes in the imaging and genomic parameters tended to be stronger than those between the baseline parameters. This may be because there are a wide range of factors which can affect the levels of FDG uptake at baseline, while the pre- to post-Metformin changes in FDG uptake are likely to be driven primarily by the effects of Metformin on the tumour metabolism. This means that the underlying causes of the changes

in FDG uptake may be more consistent across all the patients than those of the baseline uptake levels. This finding is also consistent with that of a study carried out by Mehta *et al.* [106], in which patients underwent dynamic contrast enhanced MRI (DCE-MRI) imaging and tumour biopsies before and after receiving an anti-angiogenic drug. The authors of this study reported stronger correlations between the pre- to post-treatment changes in the DCE-MRI kinetic parameters and gene expression levels than those between the baseline imaging and genomic parameters.

When O'Sullivan *et al.* [111] compared the 2C3K and a non-parametric model representing the residue function with splines, they reported that both models gave similar k_{flux} values. This is consistent with the results shown in Figure 6.13 (a), which shows that at baseline the k_{flux} values from the spline-residue and 2C3K models were strongly correlated ($R = 0.806$, p $= 7.56 \times 10^{-5}$). The pre- to post-Metformin changes in the k_{flux} values derived from 2C3K and spline-residue model fits were less strongly, though still significantly, correlated ($R = 0.465$, p $= 0.0097$). The weaker correlation between the pre- to post-Metformin changes in the k_{flux} values derived from 2C3K and spline-residue model fits is at least partly due to one outlying point in Figure 6.13 (b), which was caused by one post-drug 2C3K k_{flux} value being much higher than all the others.

In general, the k_{flux} constant from the three kinetic models studied correlated better with the patient genomic data than did any of the individual compartment model rate constants, despite these rate constants being conceptually more closely linked to specific biological processes. However, other studies (e.g. [93, 103]) have shown that the k_{flux} parameter typically has lower uncertainties than those of both the individual compartment model rate constants and the V_b parameter. The superior statistical robustness of k_{flux} may explain why it tended to correlate better with the genomic data.

A potential limitation of this study is that the asymptotic (i.e. where $T_{scan} \to \infty$) expressions for the 2C3K and 3C5K k_{flux} parameters were used (see Table 6.1). For the spline-residue model, however, it is not possible to use such an extrapolation, and so k_{flux} must be estimated using the value of the fitted residue function at

T_{scan} [111]. The non-asymptotic compartment model k_{flux} values can be calculated by taking the residue function from the fitted compartment models (which is defined in terms of the fitted rate constants, as shown in [55]), and multiplying the value of this residue function at T_{scan} by the fitted value of K_1. The asymptotic and non-asymptotic k_{flux} values from the 2C3K and 3C5K models were found to be very strongly correlated ($R = 1.000$ to three decimal places for both models, with p = $1.0\times\ 10^{-18}$ and p = $1.3\times\ 10^{-19}$ for the 2C3K and 3C5K k_{flux} values, respectively). The reason that the two sets of k_{flux} values are so similar is that the residue functions were almost constant by the end-time of the scan (see Figure 6.7 (c) and (d) for two specific examples), meaning that further extrapolating the residue functions doesn't substantially change their values. This implies that, for the specific patient dataset studied here, comparing the spline-residue k_{flux} values to the non-asymptotic compartment model k_{flux} values would not make a substantial difference to the overall results.

While the links between FDG-dPET imaging and gene expression data have not been extensively studied, Strauss *et al.* [147] have investigated the link between the expression of glucose transporter and hexokinase genes and the K_1 and k_3 parameters from fits of the 2C3K model to FDG-dPET TACs derived from colorectal cancer tumours. The authors did not find significant correlations between K_1 or k_3 with either GLUT1 or HK2 expression, which is consistent with the results presented here, though they did find that k_3 was weakly correlated with HK1, which was not observed here.

In two additional studies, Strauss *et al.* investigated the links between static PET and 2C3K parameters (including k_{flux}) and genes related to angiogenesis [146] and proliferation [145], all of which were derived from the same colorectal cancer dataset. The authors reported that both SUV and the 2C3K parameters were correlated with a number angiogenesis and proliferation related genes, though in this chapter none of these PET parameters were found to be significantly correlated with the angiogenesis or proliferation scores after multiple hypothesis testing was accounted for. It should be noted that the results published in [145, 146] are not entirely

comparable to the results presented here, as the authors investigated correlations with individual gene expression levels and not signature scores. Furthermore, the authors did not account for multiple hypothesis testing in those studies, meaning that some of their observed correlations could be false positives.

In this study the dPET parameters tended to give better correlations to the genomic measures than those derived from the static PET images, which is consistent with the results presented in [145, 146]. TBR_{mean} tended to produce better correlations than SUV_{mean}, especially for the baseline hypoxia and proliferation signature scores, though none of the correlations between static PET parameters and the genomic measures passed the Benjamini-Hochberg procedure.

6.6 Conclusion

Of the kinetic models that were compared in this work, the spline-residue model provided the best description of the FDG-dPET TACs derived from primary breast cancer tumours, and after accounting for multiple hypothesis testing the k_{flux} parameter obtained from fits of the spline-residue model was the most strongly correlated with the available genomic data. In addition to providing worse fits to the patient TACs, the fits of the compartment models produced more apparently outlying kinetic parameters, which in two cases produced correlations that are likely to be anomalous.

Stronger correlations were observed between the pre- to post-Metformin changes in the imaging and genomic parameters than those between the baseline parameters, which may be due to the fact that the underlying causes of the pre- to post-Metformin changes in FDG uptake are more consistent across all the patients in the dataset.

The correlations between the static PET parameters and the genomic measures were generally worse than those of the dPET parameters, with none of the static PET results passing the Benjamini-Hochberg procedure.

It is possible that, with a dataset containing a larger number of patients, more of the baseline and pre-post Metformin correlations and multiple regression results presented in appendix A may have passed the Benjamini-Hochberg procedure.

Chapter 7

Conclusions and Further Work

7.1 Summary of Results

One of the main challenges in applying 4D-PET reconstruction to cancer imaging is that there can potentially be many different tissues in the scanner FOV, leading to a wide range of TAC shapes to be fitted during the reconstruction. Bias from poorly modelled regions can spatially propagate to well modelled regions during 4D-PET reconstruction, meaning that flexible temporal functions are required. Although many temporal models have been proposed for 4D-PET reconstruction, few studies have inter-compared their performance, meaning that the optimal choice of function remains an open question. This issue forms the primary focus of this thesis.

In chapter 4, a 4D-PET reconstruction algorithm using adaptive-knot cubic B-splines was proposed. This algorithm was compared to 4D-PET reconstruction based on the spectral model, and to the conventional (non-4D) MLEM and MAP algorithms, using Monte-Carlo simulated digital phantom data based on a stage IV NSCLC patient. Overall, the adaptive-knot spline based reconstruction algorithm was found to produce the least biased images of all the algorithms being compared, though 4D-PET reconstruction with the spectral model was found to produce a better bias-noise trade-off in the early time-frames.

It was hypothesised that the reason the spectral model performed best in the early time-frames was because its basis functions are generated by convolving exponential decay functions with the AIF. This causes the shapes of the spectral model basis functions to be constrained at the early time-points. TACs vary most

rapidly early on during imaging, and so short time-frames are used for the early time-points, making them particularly noisy. The greater constraints on the spectral model basis function shapes at the early time-points therefore makes them less likely than the more flexible spline functions to overfit the data from the early time-frames.

In chapter 5, a 'spline-residue' model was proposed. This model describes TACs as convolutions of the AIF with cubic B-spline basis functions. Like the spectral model, convolution with the input function constrains the spline-residue model at early time-points, potentially enhancing noise suppression in the early time-frames, while using cubic B-splines instead of exponential decays allows a wider range of TAC descriptions over the entire imaged time-course, potentially limiting bias. Spline-residue model-based reconstruction was compared to 4D reconstructions based on adaptive-knot cubic B-splines, the spectral model and the non-linear 2C3K model. The 4D reconstructions in this chapter were carried out using a nested-MAP algorithm that incorporated both spatial and temporal roughness penalties into the reconstruction. Of the methods compared in this chapter, the spline-residue model produced images with the best bias-noise trade-off, and compared to the conventional MAP algorithm this model produced reductions of 50% or more in bias or noise for a majority of the combinations of kinetics parameters and bias/noise measures that were analysed, and did not generate notably worse results for any combination.

Another notable result from chapters 4 and 5 was that incorporating spatial and temporal roughness penalties into the reconstruction objective function can provide substantial reductions in noise compared to 4D reconstructions with no such penalty terms, provided that the penalty functions and their associated hyperparameters are well chosen. Temporal roughness penalties are not commonly used in 4D-PET reconstruction studies, possibly because they introduce additional complexity into the reconstruction process, and furthermore the results in chapter 5 show that care must be taken when selecting the penalty function to use for the temporal model being fitted. However, the results in chapter 5 imply that the potential gains from the use of temporal roughness penalty terms make their incorporation into the reconstruction process worthwhile.

The spline-residue model-based 4D reconstruction algorithm introduced in chapter 5 was used to produce dPET image sequences with improved levels of bias and noise. The TACs from these reconstructed dPET images were then analysed post-reconstruction using a 2C3K model. However, the spline-residue model can be considered as a 'nonparametric' kinetic model, which produces an estimate of the residue function by deconvolving the AIF from the measured TAC. It is possible to derive kinetic parameter estimates from the residue function, meaning that physiologically relevant information could potentially be derived directly from fits of the spline-residue model carried out during 4D-PET reconstruction, rather than the TACs being analysed with a different model post-reconstruction. While a number of studies have demonstrated that post-reconstruction kinetics analysis using non-parametric models gives better TAC descriptions than more commonly used compartment models, to our knowledge no published studies have investigated whether the kientic parameter estimates obtained from these models provide an improved description of the underlying patient biology.

In chapter 6, the spline-residue model was used to analyse FDG-dPET data from a clinical dataset comprising 41 primary breast cancer patients. These patients underwent dPET scans and tumour biopsies before and after the administration of the drug Metformin. Fits of the spline-residue model were compared to those of the 2C3K and 3C5K models. Of the models compared, the spline-residue model provided the best description of the patient tumour TACs, and after accounting for multiple hypothesis testing the k_{flux} parameter obtained from fits of the spline-residue model was the most strongly correlated with the genomic data derived from the patient biopsies.

The results presented in this thesis indicate that the choice of temporal model for use in 4D-PET reconstruction has a strong influence on the quality of both the reconstructed images and the estimated kinetic parameters. Of the methods compared in this work, the proposed spline-residue model was found to produce images with the best bias-noise trade-off. While the reconstruction algorithms in chapters 4 and 5 were proposed with a view to providing temporally smoothed image

sequences, which could then be analysed post-reconstruction using a conventional compartment model, the results in chapter 6 suggest that physiologically relevant kinetic parameters can be directly obtained from fits of the spline-residue model during 4D-PET reconstruction, and that (at least for FDG imaging) these parameters could potentially provide an improved description of the underlying patient biology than those derived from a compartment model fit.

7.2 Further Work

The main focus of this thesis was on how the choice of temporal function affects the noise suppressing properties of 4D-PET reconstruction. Because of this, as a simplifying assumption patient motion was not simulated in any of the phantom studies carried out here. However, for cancers such as NSCLC, patient motion can lead to blurring artefacts and misalignment of anatomical features between time-frames. Incorporating motion correction into the 4D-PET image reconstruction process could therefore lead to further improvements in the quality of both the reconstructed images and the estimated kinetic parameters. Potential approaches for accounting for patient motion during 4D-PET reconstruction include:

1. Pre-calculate the deformation fields by first reconstructing the dPET image sequences with a conventional algorithm and performing motion correction on these images using an existing method, such as one of those reviewed in [123, 48]. Finally, use the pre-calculated deformation fields to perform motion correction before each model fitting step during 4D-PET reconstruction.
2. In addition to fitting the kinetic models, perform the motion correction during the 4D-PET reconstruction. This has been proposed in a number of studies [68, 73, 72].

The second of the approaches listed above is appealing, as it allows both the motion correction and the model fitting to be handled during the image reconstruction. However, performing both motion correction and model fitting during the image reconstruction is likely to be computationally intensive, especially

as the motion correction would potentially have to be applied at each reconstruction iteration, while in the first of the above methods it would only have to be applied once. Although a few studies have proposed methods for incorporating motion correction into 4D-PET reconstruction, the optimal approach for doing so remains an open question. A project focused on developing methods for incorporating motion correction into 4D-PET reconstruction would therefore be both interesting and useful.

The 4D-PET reconstruction algorithms proposed in this work were validated using Monte-Carlo simulated data generated from digital phantoms. This approach is useful as it gives a known ground-truth to validate the reconstructed images against, and is the approach used by almost all studies in which 4D-PET reconstruction algorithms are evaluated. However, the ultimate goal of 4D-PET reconstruction is to provide improved estimates of clinically relevant kinetic parameters. To compare how well the estimated kinetic parameters derived from images reconstructed using different methods reflect the underlying patient biology, methods such as those used in chapter 6, in which the kinetic parameters are correlated with biological measurements obtined from biopsies or histological data, could be used. An example of such a study was carried out by Cheng *et al.* [21].

Carrying out clinical and pre-clinical studies is more costly than using digital phantoms, and there are uncertainties associated with biological measurements derived from the imaged subjects. However, evaluating a 4D-PET reconstruction algorithm using both digital phantoms and either clinical or preclinical data (with biological measurements in addition to the scanner data) would allow the algorithm to be validated in both an idealised and a realistic clinical setting, thus increasing its chances of being used in wider clinical practice. The patient dataset used in chapter 6 would have been ideal for this purpose, though unfortunately the sinogram data from the patient scans was not available.

None of the 4D-PET reconstruction algorithms in this work used ordered subsets to accelerate the image reconstruction process. The reason this was not done here is that the choice of the number of subsets to use is an additional tunable

parameter which can affect the quality of the reconstructed images. Furthermore, for research purposes the speed of the reconstruction is less of an issue than it would be in a clinical setting where many PET scans are performed each day. Nested 4D reconstructions have previously been extended to include ordered subsets (e.g. [80]), meaning that the same extensions could probably be applied to the reconstruction algorithms used here.

Wang and Qi [179] have proposed a novel method for performing spatially regularized PET image reconstruction by incorporating a kernel-based image model containing prior information into the system matrix. Such reconstructions can thus be performed using the standard MLEM algorithm with this modified system matrix, as opposed to the more complex MAP algorithm used in this work. Wang and Qi [179] showed that this kernel-based approach out-performed a number of commonly used MAP algorithms. Wang has recently extended the original kernel based approach to provide spatio-temporal regularization [171]. A 4D nested-EM algorithm in which the image update step was performed by the spatially (though not spatio-temporally) regularized kernel-based MLEM update equation has been recenlty proposed by Gong *et al.* [49], and shows promising results. While a MAP framework was used in this thesis to incorporate spatial and temporal roughness penalties into the reconstruction, it is possible that using these kernel-based approaches may provide superior results while simplifying the implementation of the reconstruction algorithms.

Appendices

Appendix A

Supplementary Results for Chapter 6

In chapter 6 a number of genomic measures and PET imaging parameters were compared. These were derived from a dataset containing 41 primary breast cancer patients, who underwent FDG-dPET scans and tumour biopsies before and after the administration of the drug metformin.

The goal of this study was to compare how well each PET imaging parameter matched the patient genomic data, and in particular to investigate the extent to which kinetic parameters from fits of the spline-residue model introduced in chapter 5 reflect the underlying patient biology. Full details of the methods used in this study are available in section 6.3.

When comparing the genomic and imaging measures in chapter 6 a total of 238 hypotheses were tested, and the Benjamini-Hochberg procedure was used to account for multiple hypothesis testing. In chapter 6 only the results that passed this procedure are presented. Here the results of every individual test are presented. It is recommended that the reader refers back to section 6.3 in for an explanation of how the results shown here were obtained.

Table A.1. Correlations between GLUT1 expression and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Pearson Correlation Coefficient (Baseline)	Pearson Correlation Coefficient (Pre-Post Metformin change)
2C3K Model		
K_1	0.053 (p = 0.7801)	0.001 (p = 0.9978)
k_2	0.073 (p = 0.7021)	0.052 (p = 0.7863)
k_3	-0.120 (p = 0.5284)	0.084 (p = 0.6601)
k_{flux}	-0.058 (p = 0.0701)	0.071 (p = 0.7096)
V_b	-0.335 (p = 0.7617)	-0.127(p = 0.5037)
3C5K Model		
K_1	**0.441 (p = 0.0148)**	0.190 (p = 0.3142)
k_2	-0.049 (p = 0.7990)	0.104 (p = 0.5839)
k_3	-0.126 (p = 0.5057)	0.044 (p = 0.8194)
k_4	-0.072 (p = 0.7036)	0.041 (p = 0.8294)
k_5	-0.167 (p = 0.3790)	0.027 (p = 0.8881)
k_{flux}	**0.422 (p = 0.0202)**	0.070 (p = 0.7116)
V_b	0.034 (p = 0.8568)	-0.145 (p = 0.4459)
Spline-Residue Model		
K_1	0.015 (p = 0.9368)	-0.067 (p = 0.7302)
k_{flux}	0.328 (p = 0.0765)	**0.608 (p = 0.0004)**
V_b	-0.143 (p = 0.4521)	-0.287 (p = 0.1248)
Static PET Parameters		
SUV_{mean}	0.212 (p = 0.2598)	0.150 (p = 0.4274)
TBR_{mean}	0.203 (p = 0.2818)	0.179 (p = 0.3435)

Table A.2. Correlations between HK2 expression and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Pearson Correlation Coefficient (Baseline)	Pearson Correlation Coefficient (Pre-Post Metformin change)
2C3K Model		
K_1	-0.204 (p = 0.2787)	-0.234 (p = 0.2039)
k_2	-0.183 (p = 0.3336)	-0.285 (p = 0.1266)
k_3	-0.099 (p = 0.6046)	0.031 (p = 0.8716)
k_{flux}	-0.212 (p = 0.2613)	-0.161 (p = 0.8799)
V_b	0.229 (p = 0.2243)	-0.058 (p = 0.6724)
3C5K Model		
K_1	0.335 (p = 0.070)	0.063 (p = 0.6443)
k_2	-0.039 (p = 0.8384)	0.051 (p = 0.7806)
k_3	0.019 (p = 0.9200)	-0.012 (p = 0.9217)
k_4	-0.181 (p = 0.3382)	-0.207 (p = 0.5243)
k_5	-0.349 (p = 0.0591)	0.018 (p = 0.5258)
k_{flux}	-0.078 (p = 0.6805)	0.073 (p = 0.9779)
V_b	0.259 (p = 0.1674)	0.190 (p = 0.4335)
Spline-Residue Model		
K_1	-0.315 (p = 0.2653)	0.008 (p = 0.1078)
k_{flux}	0.074 (p = 0.6978)	**0.519 (p = 0.0033)**
V_b	-0.049 (p = 0.7959)	0.029 (p = 0.9465)
Static PET Parameters		
SUV_{mean}	-0.246 (p = 0.1895)	-0.073 (p = 0.8999)
TBR_{mean}	-0.251 (p = 0.1811)	-0.149 (p = 0.9983)

Table A.3. Correlations between angiogenesis signature score and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Pearson Correlation Coefficient (Baseline)	Pearson Correlation Coefficient (Pre-Post Metformin change)
2C3K Model		
K_1	-0.055 (p = 0.7721)	-0.146 (p = 0.4401)
k_2	-0.040 (p = 0.8333)	-0.012 (p = 0.9523)
k_3	-0.104 (p = 0.5843)	**-0.520 (p = 0.0032)**
k_{flux}	-0.087 (p = 0.6484)	-0.180 (p = 0.3497)
V_b	0.259 (p = 0.1669)	0.184 (p = 0.3314)
3C5K Model		
K_1	-0.195 (p = 0.3001)	-0.266 (p = 0.1562)
k_2	-0.121 (p = 0.5257)	**-0.521 (p = 0.0031)**
k_3	0.019 -0.067 (p = 0.7237)	**-0.414 (p = 0.0230)**
k_4	0.072 (p = 0.7067)	-0.087 (p = 0.6476)
k_5	-0.267 (p = 0.1530)	-0.145 (p = 0.4433)
k_{flux}	-0.064 (p = 0.3924)	-0.260 (p = 0.1640)
V_b	0.162 (p = 0.7385)	-0.022 (p = 0.9091)
Spline-Residue Model		
K_1	-0.114 (p = 0.5488)	-0.061 (p = 0.7481)
k_{flux}	-0.125 (p = 0.5118)	-0.322 (p = 0.0823)
V_b	0.022 (p = 0.9086)	0.133 (p = 0.4847)
Static PET Parameters		
SUV_{mean}	0.011 (p = 0.9531)	-0.184 (p = 0.3312)
TBR_{mean}	-0.026 (p = 0.8898)	-0.195 (p = 0.3019)

Table A.4. Correlations between hypoxia signature score and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Pearson Correlation Coefficient (Baseline)	Pearson Correlation Coefficient (Pre-Post Metformin change)
2C3K Model		
K_1	-0.011 (p = 0.9544)	-0.041 (p = 0.8396)
k_2	-0.061 (p = 0.7470)	-0.107 (p = 0.5732)
k_3	-0.122 (p = 0.5210)	0.103 (p = 0.5881)
k_{flux}	**0.478 (p = 0.0075)**	0.240 (p = 0.1864)
V_b	-0.158 (p = 0.4040)	-0.248 (p = 0.2014)
3C5K Model		
K_1	0.079 (p = 0.6792)	0.020 (p = 0.9177)
k_2	-0.097 (p = 0.6103)	0.100 (p = 0.5990)
k_3	-0.080 (p = 0.6729)	0.001 (p = 0.9975)
k_4	-0.349 (p = 0.0587)	-0.105 (p = 0.5808)
k_5	0.037 (p = 0.8470)	0.045 (p = 0.8141)
k_{flux}	**0.469 (p = 0.0089)**	0.040 (p = 0.8346)
V_b	0.054 (p = 0.7773)	0.054 (p = 0.7749)
Spline-Residue Model		
K_1	0.227 (p = 0.2274)	0.252 (p = 0.1785)
k_{flux}	**0.421 (p = 0.0204)**	**0.541 (p = 0.0020)**
V_b	0.013 (p = 0.9449)	-0.198 (p = 0.2943)
Static PET Parameters		
SUV_{mean}	0.358 (p = 0.0519)	0.349 (p = 0.0587)
TBR_{mean}	**0.431 (p = 0.0174)**	0.319 (p = 0.0859)

Table A.5. Correlations between proliferation signature score and each kinetic and static PET parameter. Results for baseline and pre-post metformin change data are shown. p-values are given in brackets. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Pearson Correlation Coefficient (Baseline)	Pearson Correlation Coefficient (Pre-Post Metformin change)
2C3K Model		
K_1	0.092 (p = 0.6285)	-0.228 (p = 0.2262)
k_2	-0.068 (p = 0.7225)	-0.016 (p = 0.9333)
k_3	-0.081 (p = 0.6712)	-0.067 (p = 0.7241)
k_{flux}	**0.451 (p = 0.0123)**	0.131 (p = 0.4887)
V_b	0.047 (p = 0.8068)	0.031 (p = 0.8690)
3C5K Model		
K_1	**0.403 (p = 0.0272)**	0.204 (p = 0.2775)
k_2	-0.033 (p = 0.8610)	-0.044 (p = 0.8177)
k_3	-0.132 (p = 0.4868)	-0.253 (p = 0.1769)
k_4	**-0.483 (p = 0.0068)**	-0.120 (p = 0.5291)
k_5	-0.172 (p = 0.3645)	0.051 (p = 0.7901)
k_{flux}	**0.392 (p = 0.0320)**	-0.042 (p = 0.8267)
V_b	0.237 (p = 0.2065)	0.035 (p = 0.8540)
Spline-Residue Model		
K_1	0.207 (p = 0.2716)	0.186 (p = 0.324)
k_{flux}	**0.479 (p = 0.0074)**	**0.613 (p = 0.0003)**
V_b	0.053 (p = 0.7825)	-0.106 (p = 0.5769)
Static PET Parameters		
SUV_{mean}	0.355 (p = 0.0542)	0.252 (p = 0.1794)
TBR_{mean}	**0.455 (p = 0.0115)**	0.233 (p = 0.2154)

Table A.6. Multiple regression results for the linear regression model containing the expression levels of all available glucose transporter genes. F-test results for baseline and pre-post metformin change data are shown. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Baseline p-value	Pre-Post Metformin Change p-value
2C3K Model		
K_1	0.161	0.293
k_2	0.481	0.054
k_3	0.242	0.833
k_{flux}	0.161	**0.024**
V_b	0.841	0.772
3C5K Model		
K_1	0.818	0.377
k_2	0.172	0.841
k_3	0.356	0.473
k_4	0.535	0.510
k_5	0.374	0.207
k_{flux}	0.206	0.829
V_b	0.762	0.236
Spline-Residue Model		
K_1	0.841	0.800
k_{flux}	0.304	0.111
V_b	0.587	0.737
Static PET Parameters		
SUV_{mean}	0.193	0.797
TBR_{mean}	0.080	0.807

Table A.7. Multiple regression results for the linear regression model containing the expression levels of all available hexokinase genes. F-test results for baseline and pre-post metformin change data are shown. Statistically significant ($p < 0.05$) results are highlighted in bold.

Parameter	Baseline p-value	Pre-Post Metformin Change p-value
2C3K Model		
K_1	0.698	0.418
k_2	0.707	0.460
k_3	0.664	0.711
k_{flux}	0.607	0.238
V_b	0.687	0.547
3C5K Model		
K_1	0.158	0.679
k_2	0.653	0.711
k_3	0.712	0.624
k_4	0.681	0.692
k_5	0.713	0.365
k_{flux}	0.711	0.532
V_b	0.629	0.665
Spline-Residue Model		
K_1	**0.039**	0.499
k_{flux}	0.331	0.066
V_b	0.658	0.703
Static PET Parameters		
SUV_{mean}	0.271	0.487
TBR_{mean}	0.107	0.062

References

[1] P. Aguiar, M. Rafecas, J. E. Ortuño, G. Kontaxakis, A. Santos, J. Pavía, and D. Ros. Geometrical and Monte Carlo projectors in 3D PET reconstruction. *Medical Physics*, 37(11):5691–5702, 2010.

[2] S. Ahn and J. A. Fessler. Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms. *IEEE Transactions on Medical Imaging*, 22(5):613–626, 2003.

[3] S. Alenius and U. Ruotsalainen. Bayesian image reconstruction for emission tomography based on median root prior. *European Journal of Nuclear Medicine*, 24(3):258–265, 1997.

[4] R. D. Badawi and P. K. Marsden. Developments in component-based normalization for 3D PET. *Physics in Medicine & Biology*, 44(2):571–594, 1999.

[5] D. L. Bailey and S. R. Meikle. A convolution-subtraction scatter correction method for 3D PET. *Physics in Medicine & Biology*, 39(3):411–424, 1994.

[6] J. Baro, J. Sempau, J. Fernández-Varea, and F. Salvat. PENELOPE: an algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 100(1):31–46, 1995.

[7] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui. Noise properties of the EM algorithm. I. Theory. *Physics in Medicine & Biology*, 39(5):833–846, 1994.

[8] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 289–300, 1995.

[9] C. A. Bouman and K. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, 2(3):296–310, 1993.

[10] C. A. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Transactions on Image Processing*, 5(3):480–492, 1996.

[11] J. Browne and A. De Pierro. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Transactions on Medical Imaging*, 15(5):687–699, 1996.

[12] A. K. Buck, K. Herrmann, T. Stargardt, T. Dechow, B. J. Krause, and J. Schreyögg. Economic evaluation of PET and PET/CT in oncology: evidence and methodologic approaches. *Journal of Nuclear Medicine Technology*, 38(1):6–17, 2010.

[13] F. Buffa, A. Harris, C. West, and C. Miller. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *British Journal of Cancer*, 102(2):428–435, 2010.

[14] I. Buvat, I. Castiglioni, J. Feuardent, and M.-C. Gilardi. Unified description and validation of Monte Carlo simulators in PET. *Physics in Medicine & Biology*, 50(2):329–346, 2005.

[15] C. L. Byrne. Convergent block-iterative algorithms for image reconstruction from inconsistent data. *IEEE Transactions on Image Processing*, 6(9):1296–1304, 1997.

[16] R. Carson and K. Lange. The EM Parametric Image Reconstruction Algorithm. *Journal of the American Statistical Association*, 80(389):20–22, 1985.

[17] M. Casey, H. Gadagkar, and D. Newport. A component based method for normalization in volume PET. In *Three-dimensional image reconstruction in radiology and nuclear medicine, Aix-les-Bains, France*, pages 67–71. 1995.

[18] J. E. Cavanaugh. Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statistics & Probability Letters*, 33(2):201–208, 1997.

[19] C. Chen, S.-Y. Lee, and Z. Cho. Parallelization of the EM algorithm for 3-D PET image reconstruction. *IEEE Transactions on Medical Imaging*, 10(4):513–522, 1991.

[20] K. Chen, S.-C. Huang, and D.-C. Yu. The effects of measurement errors in the plasma radioactivity curve on parameter estimation in positron emission tomography. *Physics in Medicine & Biology*, 36(9):1183–1200, 1991.

[21] X. Cheng, C. Bayer, C.-A. Maftei, S. T. Astner, P. Vaupel, S. I. Ziegler, and K. Shi. Preclinical evaluation of parametric image reconstruction of [18F] FMISO PET: correlation with ex vivo immunohistochemistry. *Physics in Medicine & Biology*, 59(2):347–362, 2013.

[22] X. Cheng, Z. Li, Z. Liu, N. Navab, S.-C. Huang, U. Keller, S. I. Ziegler, and K. Shi. Direct parametric image reconstruction in reduced parameter space for rapid multi-tracer PET imaging. *IEEE Transactions on Medical Imaging*, 34(7):1498–1512, 2015.

[23] S. R. Cherry, M. Dahlbom, and E. J. Hoffman. 3D PET using a conventional multislice tomograph without septa. *Journal of Computer Assisted Tomography*, 15(4):655–668, 1991.

[24] S. R. Cherry, J. A. Sorenson, and M. E. Phelps. *Physics in Nuclear Medicine*. Elsevier, Philadelphia, PA, 4 edition, 2006.

[25] V. Cunningham and T. Jones. Spectral analysis of dynamic PET studies. *Journal of Cerebral Blood Flow and Metabolism*, 13(1):15–23, 1993.

[26] J. Czernin and M. E. Phelps. Positron emission tomography scanning: current and future applications. *Annual Review of Medicine*, 53(1):89–112, 2002.

[27] M. E. Daube-Witherspoon and R. E. Carson. Unified deadtime correction model for PET. *IEEE Transactions on Medical Imaging*, 10(3):267–275, 1991.

[28] M. E. Daube-Witherspoon and G. Muehllehner. Treatment of axial data in three-dimensional PET. *Journal of Nuclear Medicine*, 28(11):1717–1724, 1987.

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

[30] R. De la Prieta, J. Hernandez, E. Schiavi, and N. Malpica. Analytical geometric model for photon coincidence detection in 3D PET. In *Nuclear Science Symposium Conference Record, 2006. IEEE*, volume 4, pages 2229–2232, 2006.

[31] A. R. De Pierro. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Transactions on Medical Imaging*, 14(1):132–137, 1995.

[32] M. Defrise, P. E. Kinahan, D. W. Townsend, C. Michel, M. Sibomana, and D. F. Newport. Exact and Approximate Rebinning Algorithms for 3-D PET Data. *IEEE Transactions on Medical Imaging*, 16(2):145–158, 1997.

[33] G. Delso, S. Fürst, B. Jakoby, R. Ladebeck, C. Ganter, S. G. Nekolla, M. Schwaiger, and S. I. Ziegler. Performance Measurements of the Siemens mMR Integrated Whole-Body PET/MR Scanner. *Journal of Nuclear Medicine*, 52(12):1914–1922, 2011.

[34] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood form incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–38, 1977.

[35] C. Desmedt, B. Haibe-Kains, P. Wirapati, M. Buyse, D. Larsimont, G. Bontempi, M. Delorenzi, M. Piccart, and C. Sotiriou. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clinical Cancer Research*, 14(16):5158–5165, 2008.

[36] L. Eriksson, K. Wienhard, and M. Dahlbom. A simple data loss model for positron camera systems. *IEEE Transactions on Nuclear Science*, 41(4):1566–1570, 1994.

[37] S. España, J. Herraiz, E. Vicente, J. J. Vaquero, M. Desco, and J. M. Udías. PeneloPET, a Monte Carlo PET simulation tool based on PENELOPE: features and validation. *Physics in Medicine & Biology*, 54(6):1723–1742, 2009.

[38] Y.-H. D. Fang and R. F. Muzic. Spillover and partial-volume correction for image-derived input functions for small-animal 18F-FDG PET studies. *Journal of Nuclear Medicine*, 49(4):606–614, 2008.

[39] M. D. Farwell, D. A. Pryma, and D. A. Mankoff. PET/CT imaging in cancer: current applications and future directions. *Cancer*, 120(22):3433–3445, 2014.

[40] D. Feng, S. Huang, and X. Wang. Models for computer simulation studies of input functions for tracer kinetic modelling with positron emission tomography. *International Journal of Bio-Medical Computing*, 32:95–110, 1993.

[41] F. H. Fernandez, D. Hawe, L. O'Suilleabhain, E. Wolsztynski, J. Huang, M. Muzi, and F. O'Sullivan. Kinetic analysis of dynamic 18 F-FDG and 15 OH 2 O PET studies by parametric and nonparametric methods: A statistical analysis. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 3823–3827, 2011.

[42] R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915.

[43] R. A. Fisher. On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, 1:3–32, 1921.

[44] S. Ganan and D. McClure. Bayesian image analysis: An application to single photon emission tomography. In *Proceedings of American Statistical Association*, pages 12–18, 1985.

[45] G. Germano and E. J. Hoffman. Investigation of count rate and deadtime characteristics of a high resolution PET system. *Journal of Computer Assisted Tomography*, 12(5):836–846, 1988.

[46] M. Germino, J.-D. Gallezot, J. Yan, and R. E. Carson. Direct reconstruction of parametric images for brain PET with event-by-event motion correction: evaluation in two tracers across count levels. *Physics in Medicine & Biology*, 62(13):5344–5364, 2017.

[47] R. J. Gillies, I. Robey, and R. A. Gatenby. Causes and consequences of increased glucose metabolism of cancers. *Journal of Nuclear Medicine*, 49(2):24S–42S, 2008.

[48] A. Gillman, S. Rose, J. Smith, P. Thomas, and N. Dowson. PET motion correction in context of integrated PET/MR: current techniques, limitations, and future projections. *Medical Physics*, pages e430–e445, 2017.

[49] K. Gong, J. Cheng-Liao, G. Wang, K. T. Chen, C. Catana, and J. Qi. Direct Patlak reconstruction from dynamic PET data using kernel method with MRI information based on structural similarity. *IEEE transactions on medical imaging*, 37(4):955–965, 2018.

[50] E. Gravier, Y. Yang, M. A. King, and M. Jin. Fully 4D motion-compensated reconstruction of cardiac SPECT images. *Physics in Medicine & Biology*, 51(18):4603–4619, 2006.

[51] P. J. Green. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Transactions on Medical Imaging*, 9(1):84–93, 1990.

[52] L. Grezes-Besset, J. Nuyts, R. Boellard, I. Buvat, C. Michel, C. Pierre, N. Costes, and A. Reilhac. Simulation-based evaluation of neg-ml iterative reconstruction of low count pet data. In *Nuclear Science Symposium Conference Record, 2007. NSS'07. IEEE*, volume 4, pages 3009–3014, 2007.

[53] S. Grootoonk, T. Spinks, D. Sashin, N. Spyrou, and T. Jones. Correction for scatter in 3D brain PET using a dual energy window method. *Physics in Medicine & Biology*, 41(12):2757–2774, 1996.

[54] R. Gunn, S. Gunn, F. Turkheimer, J. Aston, and V. Cunningham. Positron emission tomography compartmental models: a basis pursuit stratergy for kinetic modelling. *Journal of Cerebral Blood Flow and Metabolism*, 22(12):1425–1429, 2002.

[55] R. N. Gunn, S. R. Gunn, and V. J. Cunningham. Positron emission tomography compartmental models. *Journal of Cerebral Blood Flow & Metabolism*, 21(6):635–652, 2001.

[56] S. Hackett, D. Liu, A. Chalkidou, P. Marsden, D. Landau, and J. Fenwick. Estimation of input functions from dynamic [18 F] FLT PET studies of the head and neck with correction for partial volume effects. *European Journal Nuclear Medicine and Molecular Imaging*, 3:84, 2013.

[57] I. Häggström, J. Axelsson, C. R. Schmidtlein, M. Karlsson, A. Garpebring, L. Johansson, J. Sörensen, and A. Larsson. A Monte Carlo study of the dependence of early frame sampling on uncertainty and bias in pharmacokinetic parameters from dynamic PET. *Journal of Nuclear Medicine Technology*, 43(1):53–60, 2015.

[58] R. Harrison, S. Vannoy, D. Haynor, S. Gillispie, M. Kaplan, and T. Lewellen. Preliminary experience with the photon history generator module of a public-domain simulation system for emission tomography. In *IEEE Nuclear Science Symposium Conference Record*, pages 1154–1158, 1993.

[59] D. Hawe, F. R. Hernández Fernández, L. O'Suilleabháin, J. Huang, E. Wolsztynski, and F. O'Sullivan. Kinetic analysis of dynamic positron emission tomography data using open-source image processing and statistical inference tools. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(3):316–322, 2012.

[60] X. He, L. Shen, and Z. Shen. A data-adaptive knot selection scheme for fitting splines. *IEEE Signal Processing Letters*, 8(5):137–139, 2001.

[61] G. V. Heller, D. Calnon, and S. Dorbala. Recent advances in cardiac PET and PET/CT myocardial perfusion imaging. *Journal of Nuclear Cardiology*, 16(6):962–969, 2009.

[62] E. J. Hoffman, T. M. Guerrero, G. Germano, W. Digby, and M. Dahlbom. PET system calibrations and corrections for quantitative and spatially accurate images. *IEEE Transactions on Nuclear Science*, 36(1):1108–1112, 1989.

[63] C. H. Holdsworth, C. S. Levin, T. H. Farquhar, M. Dahlbom, and H. E. J. Investigation of accelerated Monte Carlo techniques for PET simulation and 3D-PET scatter correction. *IEEE Transactions on Nuclear Science*, 48(1):74–81, 2001.

[64] T. Hsiao, A. Rangarajan, P. Khurd, and G. Gindi. An accelerated convergent ordered subsets algorithm for emission tomography. *Physics in Medicine & Biology*, 49(11):2145–2156, 2004.

[65] M. H. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609, 1994.

[66] A. Iriarte, R. Marabini, S. Matej, C. O. S. Sorzano, and R. M. Lewitt. System models for PET statistical iterative reconstruction: A review. *Computerized Medical Imaging and Graphics*, 48:30–48, 2016.

[67] F. Jacobs, E. Sundermann, B. De Sutter, M. Christiaens, and I. Lemahieu. A fast algorithm to calculate the exact radiological path through a pixel or voxel space. *Journal of computing and information technology*, 6(1):89–94, 1998.

[68] M. W. Jacobson and J. A. Fessler. Joint estimation of image and deformation parameters in motion-corrected PET. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 5, pages 3290–3294, 2003.

[69] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assie, D. Autret, S. Avner, R. Barbier, M. Bardies, P. Bloomfield, et al. GATE: a simulation toolkit for PET and SPECT. *Physics in Medicine & Biology*, 49(19):4543–4561, 2004.

[70] Y. Jian, B. Planeta, and R. Carson. Evaluation of bias and variance in low-count OSEM list mode reconstruction. *Physics in Medicine & Biology*, 60(1):15–29, 2014.

[71] C.-R. Jiang, J. A. Aston, and J.-L. Wang. A functional approach to deconvolve dynamic neuroimaging data. *Journal of the American Statistical Association*, 111(513):1–13, 2016.

[72] J. Jiao, A. Bousse, K. Thielemans, N. Burgos, P. S. Weston, J. M. Schott, D. Atkinson, S. R. Arridge, B. F. Hutton, P. Markiewicz, et al. Direct parametric reconstruction with joint motion estimation/correction for dynamic brain PET data. *IEEE Transactions on Medical Imaging*, 36(1):203–213, 2017.

[73] J. Jiao, A. Bousse, K. Thielemans, P. Markiewicz, N. Burgos, D. Atkinson, S. Arridge, B. F. Hutton, and S. Ourselin. Joint parametric reconstruction and motion correction framework for dynamic PET data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 114–121, 2014.

[74] M. Jin, Y. Yang, and M. A. King. Reconstruction of dynamic gated cardiac SPECT. *Medical Physics*, 33(11):4384–4394, 2006.

[75] M. E. Kamasak, C. A. Bouman, and E. D. Morris. Direct Reconstruction of Kinetic Parameter Images From Dynamic PET Data. *IEEE Transactions on Medical Imaging*, 24(5):636–650, 2005.

[76] M. E. Kamasak, B. T. Christian, C. A. Bouman, and E. D. Morris. Quality and precision of parametric images created from PET sinogram data by direct reconstruction: proof of concept. *IEEE Transactions on Medical Imaging*, 33(3):695–707, 2014.

[77] N. A. Karakatsanis, M. E. Casey, M. A. Lodge, A. Rahmim, and H. Zaidi. Whole-body direct 4D parametric PET imaging employing nested generalized Patlak expectation–maximization reconstruction. *Physics in Medicine & Biology*, 61(15):5456–5485, 2016.

[78] P. Khurd, T. Hsiao, A. Rangarajan, and G. Gindi. A globally convergent regularized ordered-subset EM algorithm for list-mode reconstruction. *IEEE Transactions on Nuclear Science*, 51(3):719–725, 2004.

[79] P. E. Kinahan and J. Rogers. Analytic 3D image reconstruction using all detected events. *IEEE Transactions on Nuclear Science*, 36(1):964–968, 1989.

[80] F. Kotasidis, J. C. Matthews, A. J. Reader, G. I. Angelis, and H. Zaidi. Application of adaptive kinetic modelling for bias propagation reduction in direct 4D image reconstruction. *Physics in Medicine & Biology*, 59(20):6061–6084, 2014.

[81] F. A. Kotasidis, J. C. Matthews, G. I. Angelis, P. J. Markiewicz, W. R. Lionheart, and A. J. Reader. Impact of erroneous kinetic model formulation in Direct 4D image reconstruction. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 2366–2367, 2011.

[82] F. Lamarche and C. Leroy. Evaluation of the volume of intersection of a sphere with a cylinder by elliptic integrals. *Computer Physics Communications*, 59(2):359–369, 1990.

[83] K. Lange. Convergence of EM image reconstruction algorithms with Gibbs smoothing. *IEEE Transactions on Medical Imaging*, 9(4):439–446, 1990.

[84] K. Lange, R. Carson, et al. EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr*, 8(2):306–16, 1984.

[85] K. Lange, D. Hunter, and I. Yang. Optimization transfer using surrogate objective functions. *Journal of Computational and Graphical Statistics*, 9(1):1–20, 2000.

[86] A. Langer. A systematic review of PET and PET/CT in oncology: a way to personalize cancer treatment in a cost-effective manner? *BMC health services research*, 10(1):283–299, 2010.

[87] R. M. Leahy and J. Qi. Statistical approaches in quantitative positron emission tomography. *Statistics and Computing*, 10(2):147–165, 2000.

[88] C. S. Levin, M. Dahlbom, and E. J. Hoffman. A Monte Carlo correction for the effect of Compton scattering in 3-D PET brain imaging. *IEEE Transactions on Nuclear Science*, 42(4):1181–1185, 1995.

[89] Q. Li, E. Asma, A. Sangtae, and R. Leahy. A Fast Fully 4-D Incremental Gradient Reconstruction Algorithm for List Mode PET Data. *IEEE Transactions on Medical Imaging*, 26(1):58–67, 2007.

[90] X. Li, D. Feng, and K. Chen. Optimal image sampling schedule: A new effective way to reduce dynamic image storage space and functional image processing time. *IEEE Transactions on Medical Imaging*, 15(5):710–719, 1996.

[91] X. Li, D. Feng, and K. Chen. Optimal image sampling schedule for both image-derived input and output functions in PET cardiac studies. *IEEE Transactions on Medical Imaging*, 19(3):233–242, 2000.

[92] W. H. Liao, K. Lange, M. Bergsneider, and S. C. Huang. Optimal design in dynamic PET data acquisition: a new approach using simulated annealing and component-wise Metropolis updating. *IEEE Transactions on Nuclear Science*, 49(5):2291–2296, 2002.

[93] D. Liu, A. Chalkidou, D. B. Landau, P. K. Marsden, and J. D. Fenwick. 18F-FLT uptake kinetics in head and neck squamous cell carcinoma: A PET imaging study. *Medical Physics*, 41(4), 2014.

[94] S. R. Lord, W.-C. Cheng, D. Liu, E. Gaude, S. Haider, T. Metcalf, N. Patel, E. J. Teoh, F. Gleeson, K. Bradley, S. Wigfield, C. Zois, D. R. McGowan, M.-L. Ah-See, A. M. Thompson, A. Sharma, L. Bidaut, M. Pollak, P. G. Roy, F. Kerpel, T. James, R. English, R. F. Adams, L. Campo, L. Ayers, C. Snell, I. Roxanis, C. Frezza, J. D. Fenwick, F. M. Buffa, and A. L. Harris. Integrated pharmacodynamic analysis identifies two metabolic adaption pathways to metformin in breast cancer. *Cell Metabolism*, 28(5):679–688, 2018.

[95] A. Lougovski, F. Hofheinz, J. Maus, G. Schramm, and J. Van den Hoff. On the relation between Kaiser–Bessel blob and tube of response based modelling of the system matrix in iterative PET image reconstruction. *Physics in Medicine & Biology*, 60(10):4209–4224, 2015.

[96] A. Lougovsky, F. Hofheinz, J. Maus, G. Schramm, E. Will, and J. Van den Hoff. A volume of intersection approach for on-the-fly system matrix calculation in 3d pet image reconstruction. *Physics in Medicine and Biology*, 59:561–577, 2014.

[97] M. Masiero, F. C. Simões, H. D. Han, C. Snell, T. Peterkin, E. Bridges, L. S. Mangala, S. Y.-Y. Wu, S. Pradeep, D. Li, et al. A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor eltd1 as a key regulator of angiogenesis. *Cancer cell*, 24(2):229–241, 2013.

[98] S. Matej and R. M. Lewitt. Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. *IEEE Transactions on Medical Imaging*, 15(1):68–78, 1996.

[99] J. C. Matthews, G. I. Angelis, F. A. Kotasidis, P. J. Markiewicz, and A. J. Reader. Direct reconstruction of parametric images using any spatiotemporal 4D image based model and maximum likelihood expectation maximisation. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*, pages 2435–2441, 2010.

[100] J. C. Matthews, D. Bailey, P. Price, and V. Cunningham. The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction. *Physics in Medicne and Biology*, 42:1155–1173, 1997.

[101] J. C. Matthews, A. J. Reader, G. Angelis, P. Price, P. Markiewicz, and F. Kotasidis. Adaptive parametric kinetic modelling for improved full field of view fitting of PET data. *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 3925–3929, 2012.

[102] B. Mazoyer, M. Roos, and R. Huesman. Dead time correction and counting statistics for positron tomography. *Physics in Medicine & Biology*, 30(5):385–399, 1985.

[103] D. R. McGowan, R. E. Macpherson, S. L. Hackett, D. Liu, F. V. Gleeson, W. G. McKenna, G. S. Higgins, and J. D. Fenwick. 18f-fluoromisonidazole uptake in advanced stage non-small cell lung cancer: a voxel-by-voxel pet kinetics study. *Medical physics*, 40:4465–4676, 2017.

[104] A. McLennan and M. Brady. Spatio-temporal reconstruction of dPET data using complex wavelet regularisation. *Medical Image Computing and Computer Assisted Intervention*, 12(2):398–405, 2009.

[105] A. McLennan, A. Reilhac, and M. Brady. SORTEO: Monte Carlo-based simulator with list-mode capabilities. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 3751–3754, 2009.

[106] S. Mehta, N. P. Hughes, S. Li, A. Jubb, R. Adams, S. Lord, L. Koumakis, R. van Stiphout, A. Padhani, A. Makris, F. M. Buffa, and A. L. Harris. Radiogenomics monitoring in breast cancer identifies metabolism and immune checkpoints as early actionable mechanisms of resistance to anti-angiogenic treatment. *EBioMedicine*, 10:109–116, 2016.

[107] S. Meikle, J. C. Matthews, V. Cunningham, D. Bailey, L. Livieratos, T. Jones, and P. Price. Parametric image reconstruction using spectral analysis of PET projection data. *Physics in Medicine & Biology*, 43(3):651–666, 1998.

[108] E. Ü. Mumcuoglu, R. M. Leahy, and S. R. Cherry. Bayesian reconstruction of PET images: methodology and performance analysis. *Physics in Medicine & Biology*, 41(9):1777–1807, 1996.

[109] T. Nichols, J. Qi, E. Asma, and R. Leahy. Spatiotemporal Reconstruction of List-Mode PET Data. *IEEE Transactions on Medical Imaging*, 21(4):396–404, 2002.

[110] F. O'Sullivan, M. Muzi, D. A. Mankoff, J. F. Eary, A. M. Spence, and K. A. Krohn. Voxel-level mapping of tracer kinetics in pet studies: a statistical approach emphasizing tissue life tables. *The annals of applied statistics*, 8(2):1065, 2014.

[111] F. O'Sullivan, M. Muzi, A. M. Spence, D. M. Mankoff, J. N. O'sullivan, N. Fitzgerald, G. C. Newman, and K. A. Krohn. Nonparametric residue analysis of dynamic PET data with application to cerebral FDG studies in normals. *Journal of the American Statistical Association*, 104(486):556–571, 2009.

[112] M. Pantaleo, M. Nannini, A. Maleddu, S. Fanti, V. Ambrosini, C. Nanni, S. Boschi, and G. Biasco. Conventional and novel PET tracers for imaging in oncology in the era of molecular therapy. *Cancer Treatment Reviews*, 34(2):103–121, 2008.

[113] C. Patlak and R. Blasberg. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. generalizations. *Journal of Cerebral Blood Flow and Metabolism*, 5(4):584–590, 1985.

[114] G. H. Perdew, J. P. V. Heuvel, and J. M. Peters. *Regulation of gene expression*. Springer Science + Business Media, New York, NY, 1 edition, 2008.

[115] M. Phelps, S. Huang, E. Hoffman, C. Selin, L. Sokoloff, and D. Kuhl. Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18) 2-fluoro-2-deoxy-D-glucose: validation of method. *Annals of Neurology*, 6(5):371–388, 1979.

[116] M. E. Phelps. *PET: Physics, Instrumentation and Scanners*. Springer Science + Business Media, New York, NY, 1 edition, 2006.

[117] M. Politis and P. Piccini. Positron emission tomography imaging in neurological disorders. *Journal of Neurology*, 259(9):1769–1780, 2012.

[118] J. Qi. A unified noise analysis for iterative image estimation. *Physics in Medicine & Biology*, 48(21):3505–3519, 2003.

[119] J. Qi and R. M. Leahy. Iterative reconstruction techniques in emission computed tomography. *Physics in Medicine & Biology*, 51(15):R541–R578, 2006.

[120] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Physics in Medicine & Biology*, 43(4):1001–1013, 1998.

[121] M. Rafecas, B. Mosler, M. Dietz, M. Pogl, A. Stamatakis, D. P. McElroy, and S. I. Ziegler. Use of a Monte Carlo-based probability matrix for 3-D iterative reconstruction of MADPET-II data. *IEEE Transactions on Nuclear Science*, 51(5):2597–2605, 2004.

[122] A. Rahmim, J. Qi, and V. Sossi. Resolution modeling in pet imaging: theory, practice, benefits, and pitfalls. *Medical physics*, 40(6):064301, 2013.

[123] A. Rahmim, O. Rousset, and H. Zaidi. Strategies for motion tracking and correction in PET. *PET clinics*, 2(2):251–266, 2007.

[124] A. Rahmim, J. Tang, and H. Zaidi. Four-dimensional (4D) image reconstruction strategies in dynamic PET: beyond conventional independent frame reconstruction. *Medical Physics*, 36(8):3654–3670, 2009.

[125] A. J. Reader, J. C. Matthews, F. C. Sureau, C. Comtat, R. Trebossen, and I. Buvat. Iterative kinetic parameter estimation within fully 4D PET image reconstruction. In *Nuclear Science Symposium Conference Record, 2006. IEEE*, volume 3, pages 1752–1756, 2006.

[126] A. J. Reader, J. C. Matthews, F. C. Sureau, C. Comtat, R. Trebossen, and I. Buvat. Fully 4D image reconstruction by estimation of an input function and spectral coefficients. In *Nuclear Science Symposium Conference Record, 2007. NSS'07. IEEE*, volume 5, pages 3260–3267, 2007.

[127] A. J. Reader, F. Sureau, C. Comtat, R. Trebossen, and I. Buvat. Simultaneous estimation of temporal basis functions and fully 4d pet images. In *Nuclear Science Symposium Conference Record, 2006. IEEE*, volume 4, pages 2219–2223, 2006.

[128] A. J. Reader, F. C. Sureau, C. Comtat, R. Trébossen, and I. Buvat. Joint estimation of dynamic PET images and temporal basis functions using fully 4D ML-EM. *Physics in Medicine & Biology*, 51(21):5455–5474, 2006.

[129] A. J. Reader and J. Verhaeghe. 4D image reconstruction for emission tomography. *Physics in Medicine and Biology*, 59(22):R371–R418, 2014.

[130] A. J. Reader and H. Zaidi. Advances in PET image reconstruction. *PET Clinics*, 2(2):173–190, 2007.

[131] A. Reilhac, G. Batan, C. Michel, C. Grova, J. Tohka, D. L. Collins, N. Costes, and A. C. Evans. PET-SORTEO: validation and development of database of simulated PET volumes. *IEEE Transactions on Nuclear Science*, 52(5):1321–1328, 2005.

[132] A. Reilhac, C. Lartizien, N. Costes, S. Sans, C. Comtat, R. Gunn, and A. Evans. PET-SORTEO: A Monte Carlo-Based Simulator With High Count Rate Capabilities. *IEEE Transactions on Nuclear Science*, 51(1):46–52, 2004.

[133] A. Reilhac, T. Soderlund, B. Thomas, Z. Irace, I. Mérida, M. Villien, J. Redouté, and N. Costes. Validation and application of PET-SORTEO for the geometry of the Siemens mMR scanner. In *PSMR Conference, Cologne*, 2016.

[134] A. Savitzky and M. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964.

[135] J. Scheins, H. Herzog, and N. J. Shah. Fully-3D PET image reconstruction using scanner-independent, adaptive projection data and highly rotation-symmetric voxel assemblies. *IEEE transactions on medical imaging*, 30(3):879–892, 2011.

[136] J. J. Scheins, F. Boschen, and H. Herzog. Analytical calculation of volumes-of-intersection for iterative, fully 3-D PET reconstruction. *IEEE Transactions on Medical Imaging*, 25(10):1363–1369, 2006.

[137] C. Schretter. A fast tube of response ray-tracer. *Medical Physics*, 33(12):4744–4748, 2006.

[138] W. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. Tsui. 4D XCAT phantom for multimodality imaging research. *Medical Physics*, 37(9):4902–4915, 2010.

[139] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122, 1982.

[140] S. Shokouhi, P. Vaska, S. Southekal, D. Schlyer, M. Purschke, V. Dzordzhadze, C. Woody, S. Stoll, D. Alexoff, D. Rubins, et al. Statistical 3D image reconstruction for the RatCAP PET tomograph using a physically accurate, Monte Carlo based system matrix. In *Nuclear Science Symposium Conference Record, 2004 IEEE*, volume 6, pages 3901–3905, 2004.

[141] R. L. Siddon. Fast calculation of the exact radiological path for a three-dimensional CT array. *Medical Physics*, 12(2):252–255, 1985.

[142] D. Snyder. Parameter-estimation for dynamic studies in emission-tomography systems having list-mode data. *IEEE Transactions on Nuclear Science*, 31(2):925–931, 1984.

[143] M. Soret, S. L. Bacharach, and I. Buvat. Partial-volume effect in PET tumor imaging. *Journal of Nuclear Medicine*, 48(6):932–945, 2007.

[144] S. Southekal, M. L. Purschke, D. J. Schlyer, and P. Vaska. Quantitative PET imaging using a comprehensive Monte Carlo system model. *IEEE Transactions on Nuclear Science*, 58(5):2286–2295, 2011.

[145] L. G. Strauss, D. Koczan, S. Klippel, L. Pan, C. Cheng, U. Haberkorn, S. Willis, and A. Dimitrakopoulou-Strauss. Impact of cell-proliferation-associated gene expression on 2-deoxy-2-[18 F] fluoro-D-glucose (FDG) kinetics as measured by dynamic positron emission tomography (dPET) in Colorectal Tumors. *Molecular Imaging and Biology*, 13(6):1290–1300, 2011.

[146] L. G. Strauss, D. Koczan, S. Klippel, L. Pan, C. Cheng, S. Willis, U. Haberkorn, and A. Dimitrakopoulou-Strauss. Impact of angiogenesis-related gene expression on the tracer kinetics of 18F-FDG in colorectal tumors. *Journal of Nuclear Medicine*, 49(8):1238–1244, 2008.

[147] L. G. Strauss, D. Koczan, S. Klippel, L. Pan, S. Willis, C. Sachpekidis, and A. Dimitrakopoulou-Strauss. Dynamic PET with 18F-Deoxyglucose (FDG) and quantitative assessment with a two-tissue compartment model reflect the activity of glucose transporters and hexokinases in patients with colorectal tumors. *American Journal of Nuclear Medicine and Molecular Imaging*, 3(5):417–424, 2013.

[148] L. Taek-Soo, W. P. Segars, and B. Tsui. Study of parameters characterizing space-time Gibbs priors for 4D MAP-RBI-EM in gated myocardial perfusion SPECT. *IEEE Nuclear Science Symposium Conference Record*, 4:2124–2128, 2005.

[149] M. M. Ter-Pogossian, M. E. Phelps, E. J. Hoffman, and N. A. Mullani. A positron-emission transaxial tomograph for nuclear imaging (PETT). *Radiology*, 114(1):89–98, 1975.

[150] A. Terstegge, S. Weber, H. Herzog, H. Muller-Gartner, and H. Halling. High resolution and better quantification by tube of response modelling in 3D PET reconstruction. In *Nuclear Science Symposium, 1996. Conference Record., 1996 IEEE*, volume 3, pages 1603–1607, 1996.

[151] Natinonal Electrical Manufacturers Association. NEMA NU-2-2012 Performance Measurement of Positron emission tomography. *National Electrical Manufacturers Association.*, 2012.

[152] L. Theodorakis, G. Loudos, V. Prassopoulos, C. Kappas, I. Tsougos, and P. Georgoulias. A review of PET normalization: striving for count rate uniformity. *Nuclear Medicine Communications*, 34(11):1033–1045, 2013.

[153] C. Thompson, J. Moreno-Cantu, and Y. Picard. PETSIM: Monte Carlo simulation of all sensitivity and resolution parameters of cylindrical positron imaging systems. *Physics in Medicine & Biology*, 37(3):731–749, 1992.

[154] N. Tikhonov, Andreĭ, A. Goncharsky, V. Stepanov, and A. G. Yagola. *Numerical methods for the solution of ill-posed problems*, volume 328. Springer Science + Business Media, Dordrecht, Netherlands, 1995.

[155] D. Townsend, A. Geissbuhler, M. Defrise, E. Hoffman, T. Spinks, D. Bailey, M.-C. Gilardi, and T. Jones. Fully three-dimensional reconstruction for a PET camera with retractable septa. *IEEE Transactions on Medical Imaging*, 10(4):505–512, 1991.

[156] C. Tsoumpas, F. Turkheimer, and K. Thielemans. Study of direct and indirect parametric estimation methods of linear models in dynamic positron emission tomography. *Medical Physics*, 35(4):1299–1309, 2008.

[157] C. Tsoumpas, F. Turkheimer, and K. Thielemans. A survey of approaches for direct parametric image reconstruction in emission tomography. *Medical Physics*, 35(9):3963–3971, 2008.

[158] J. J. Vaquero and P. Kinahan. Positron emission tomography: current challenges and opportunities for technological advances in clinical and preclinical imaging systems. *Annual Review of Biomedical Engineering*, 17:385–414, 2015.

[159] Y. Vardi, L. Shepp, and L. Kaufman. A statistical model for positron emission tomography. *Journal of the American statistical Association*, 80(389):8–20, 1985.

[160] P. Vaupel, O. Thews, and M. Hoeckel. Treatment resistance of solid tumors. *Medical Oncology*, 18(4):243–259, 2001.

[161] E. Veklerov, J. Llacer, and E. Hoffman. MLE reconstruction of a brain phantom using a Monte Carlo transition matrix and a statistical stopping rule. *IEEE Transactions on Nuclear Science*, 35(1):603–607, 1988.

[162] J. Verhaeghe, Y. D'Asseler, S. Staelens, and I. Lemahieu. Optimization of temporal basis functions in dynamic PET imaging. *Nuclear Instruments and Methods in Physics Research A*, 569:425–428, 2006.

[163] J. Verhaeghe, R. Phlypo, S. Vandenberghe, S. Staelens, Y. D'Asseler, and I. Lemahieu. Adaptive Spatially Variant Optimization of a Temporal Spline Basis for Dynamic PET Reconstruction. *IEEE Nuclear Science Symposium Conference Record*, 6:3397–3403, 2006.

[164] M. Veronese, G. Rizzo, A. Bertoldo, and F. E. Turkheimer. Spectral analysis of dynamic PET studies: a review of 20 years of method developments and applications. *Computational and Mathematical Methods in Medicine*, 2016, 2016.

[165] S. Vivekanandan, D. B. Landau, N. Counsell, D. R. Warren, A. Khwanda, S. D. Rosen, E. Parsons, Y. Ngai, L. Farrelly, L. Hughes, M. A. Hawkins, and J. D. Fenwick. The impact of cardiac radiation dosimetry on survival after radiation therapy for non-small cell lung cancer. *International Journal of Radiation Oncology, Biology, Physics*, 99(1):51–60, 2017.

[166] G. Wagenknecht, H. J. Kaiser, F. M. Mottaghy, and H. Herzog. MRI for attenuation correction in PET: methods and challenges. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 26(1):99–113, 2013.

[167] G. Wahba. *Spline models in statistics (CBMS-NSF Regional Conference Series)*. SIAM, Philadelphia, PA, 1990.

[168] M. D. Walker, M. Asselin, P. J. Julyan, M. Feldmann, P. S. Talbot, T. Jones, and J. C. Matthews. Bias in iterative reconstruction of low-statistics PET data: benefits of a resolution model. *Physics in Medicine & Biology*, 56(4):931–949, 2011.

[169] R. Walledge, R. Manavaki, M. Honer, and A. J. Reader. Inter-frame filtering for list-mode EM reconstruction in high-resolution 4-D PET. *IEEE Transactions on Nuclear Science*, 51(3):705–711, 2004.

[170] E. Wallstén, J. Axelsson, M. Karlsson, K. Riklund, and A. Larsson. A Study of Dynamic PET Frame-Binning on the Reference Logan Binding Potential. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 1(2):128–135, 2017.

[171] G. Wang. High Temporal-Resolution Dynamic PET Image Reconstruction Using A New Spatiotemporal Kernel Method. *bioRxiv*, page 323394, 2018.

[172] G. Wang, L. Fu, and J. Qi. Maximum a posteriori reconstruction of the Patlak parametric image from sinograms in dynamic PET. *Physics in Medicine and Biology*, 53(3):593, 2008.

[173] G. Wang and J. Qi. Generalised Algorithms for Direct Reconstruction of Parametric Images from Dynamic PET Data. *IEEE Transactions on Medical Imaging*, 28(11):1717–1726, 2009.

[174] G. Wang and J. Qi. Generalized algorithms for direct reconstruction of parametric images from dynamic PET data. *IEEE Transactions on Medical Imaging*, 28(11):1717–1726, 2009.

[175] G. Wang and J. Qi. Acceleration of the direct reconstruction of linear parametric images using nested algorithms. *Physics in Medicine & Biology*, 55(5):1505–1517, 2010.

[176] G. Wang and J. Qi. An Optimization Transfer Algorithm for Nonlinear Parametric Image Reconstruction From Dynamic PET Data. *IEEE Transactions on Medical Imaging*, 31(10):1977–1988, 2012.

[177] G. Wang and J. Qi. Penalized Likelihood PET Image Reconstruction using Patch-Based Edge-Preserving Regularization. *IEEE Transactions on Medical Imaging*, 31(12):2194–2204, 2012.

[178] G. Wang and J. Qi. Direct estimation of kinetic parametric images for dynamic PET. *Theranostics*, 3(10):802–815, 2013.

[179] G. Wang and J. Qi. PET image reconstruction using kernel method. *IEEE Transactions on Medical Imaging*, 34(1):61–71, 2015.

[180] W. Wang, J.-C. Georgi, S. A. Nehmeh, M. Narayanan, T. Paulus, M. Bal, J. O'Donoghue, P. B. Zanzonico, C. R. Schmidtlein, N. Y. Lee, et al. Evaluation of a compartmental model for estimating tumor hypoxia via fmiso dynamic pet imaging. *Physics in Medicine & Biology*, 54(10):3083–3099, 2009.

[181] Z. Wang, M. Gerstein, and M. Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57–63, 2009.

[182] C. C. Watson. New, faster, image-based scatter correction for 3D PET. *IEEE Transactions on Nuclear Science*, 47(4):1587–1594, 2000.

[183] C. C. Watson, D. Newport, and M. E. Casey. A single scatter simulation technique for scatter correction in 3D PET. In *Three-dimensional image reconstruction in radiology and nuclear medicine*, pages 255–268. 1996.

[184] R. Wendt. *The Mathematics of Medical Imaging: A Beginner's Guide*. Springer International Publishing AG, Basel, Switzerland, 2 edition, 2015.

[185] A. Werling, O. Bublitz, J. Doll, L.-E. Adam, and G. Brix. Fast implementation of the single scatter simulation algorithm and its use in iterative image reconstruction of PET data. *Physics in Medicine & Biology*, 47(16):2947–2960, 2002.

[186] W. W. Wheaton, S. E. Weinberg, R. B. Hamanaka, S. Soberanes, L. B. Sullivan, E. Anso, A. Glasauer, E. Dufour, G. M. Mutlu, G. R. S. Budigner, and N. S. Chandel. Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife*, 3:e02242, 2014.

[187] J. Yan, B. Planeta-Wilson, and R. E. Carson. Direct 4D list mode parametric reconstruction for PET with a novel EM algorithm. *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*, pages 3625–3628, 2008.

[188] J. Yan, B. Planeta-Wilson, and R. E. Carson. Direct 4-D PET list mode parametric reconstruction with a novel EM algorithm. *IEEE Transactions on Medical Imaging*, 31(12):2213–2223, 2012.

[189] L. Yang, G. Wang, and J. Qi. Theoretical analysis of penalized maximum-likelihood Patlak parametric image reconstruction in dynamic PET for lesion detection. *IEEE Transactions on Medical Imaging*, 35(4):947–956, 2016.

[190] H. Zaidi. Relevance of accurate monte carlo modeling in nuclear medical imaging. *Medical Physics*, 26(4):574–608, 1999.

[191] H. Zhao and A. J. Reader. Fast ray-tracing technique to calculate line integral paths in voxel arrays. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 4, pages 2808–2812, 2003.

[192] S. Zhou and X. Shen. Spatially adaptive regression splines and accurate knot selection schemes. *Journal of the American Statistical Association*, 96(453):247–259, 2001.